

Deep Learning zur Analyse von Mobile Laserscanning Daten im Kontext der kommunalen Straßenerhaltung

Inaugural-Dissertation
in der Fakultät Geistes- und Kulturwissenschaften
der Otto-Friedrich-Universität Bamberg

vorgelegt von
Maximilian Sesselmann
aus
Karl-Marx-Stadt

Bamberg, den 15.09.2023



Bamberg 2025

Dieses Werk ist als freie Onlineversion über das Forschungsinformationssystem (FIS; <https://fis.uni-bamberg.de>) der Universität Bamberg erreichbar.

Das Werk steht unter der CC-Lizenz CC BY.

Lizenzvertrag: Creative Commons Namensnennung 4.0

<https://creativecommons.org/licenses/by/4.0/>



URN: urn:nbn:de:bvb:473-irb-1065284

DOI: <https://doi.org/10.20378/irb-106528>

Diese Arbeit hat der Fakultät Geistes- und Kulturwissenschaften (Institut für Geographie) der Otto-Friedrich-Universität Bamberg als Dissertation vorgelegen.

Erstgutachter: Prof. apl. Dr. Matthias Möller

Zweitgutachter: Prof. Dr.-Ing. Ralf Bill

Tag der mündlichen Prüfung: 14.06.2024

MEINEN ELTERN
UND KINDERN
IN DANKBARKEIT

Vorwort

Die vorliegende Arbeit entstand während meiner Anstellung als Mitarbeiter in der Abteilung Forschung und Entwicklung bei der Firma Lehmann+Partner GmbH in Dresden. Im Zuge der wissenschaftlichen Auseinandersetzung mit den Themen Mobile Mapping, 3D-Punktwolkenverarbeitung und Straßenerhaltung wurden die Forschungsergebnisse in den Jahren von 2017 bis 2023 in Form von mehreren Fachartikeln und Konferenzbeiträgen veröffentlicht. An den entsprechenden Stellen im Text wird auf diese Publikationen verwiesen.

An dieser Stelle möchte ich Prof. apl. Dr. Matthias Möller, Prof. Dr.-Ing. Ralf Bill und Prof. Dr. Gerhard Schellmann für die Betreuung meiner Arbeit danken. Vielen Dank für die bereichernden Gespräche, den kritischen und stets konstruktiven Austausch im Rahmen unserer Treffen und die entscheidenden Motivationsmomente zur rechten Zeit.

Mein Dank gilt außerdem Prof. Dr.-Ing. Andreas Großmann und Jens Morawietz für ihr Vertrauen und ihre Geduld über die letzten Jahre sowie meinen geschätzten Kolleginnen und Kollegen Nadine Stelling und Steffen Scheller, für ihre Unterstützung im Rahmen der Datenerhebung sowie Thorsten Naber und Ronny Stricker für ihre Unterstützung beim Training der neuronalen Netze. Auch möchte ich Michał Maciołek und Matthäus Eliadakis danken. Ohne ihr Engagement wäre es nicht möglich gewesen, Prüfequipment und das I.R.I.S 17 Messsystem für Referenzmessungen zu verwenden.

Zusammenfassung

Für eine systematische und effiziente Erhaltung der kommunalen Straßeninfrastruktur ist eine detaillierte Kenntnis des vorherrschenden Straßenzustandes von zentraler Bedeutung. In Deutschland wurden deshalb Standards etabliert, die ein objektives Erhaltungsmanagement sicherstellen sollen. Trotz innovativer technischer Entwicklungen und hoher Automation im Bereich der mobilen Datenerfassung dominiert im Bereich der Datenauswertung die manuell-visuelle Inspektion von Oberflächenschäden aus Kamerabilddaten. Dieser Prozess ist nicht nur zeit-, personal- und kostenintensiv, sondern nutzt darüber hinaus ausschließlich fotometrische Informationen für den Inspektionsvorgang. Mobil erfasste Laserscans werden lediglich zur Erfassung und Bewertung der Ebenheit einer Straße genutzt, obwohl ausreichend dichte Oberflächenscans eine gleichzeitige dreidimensionale Beschreibung der Oberfläche und ihrer Schädigungen zuließen. Eine solche Datenbasis wäre sowohl für eine manuelle Schadenserfassung als auch für eine Automatisierung des Inspektionsvorgangs von Vorteil. Studien zeigen, dass die manuelle Erfassung von Oberflächenschäden aus Kamerabilddaten die größte Fehlerquelle im Prozess der Zustandserfassung und -bewertung von Straßen ist und das Ziel gefährdet, eine objektive und vergleichbare Datengrundlage für das Erhaltungsmanagement zu schaffen.

Ziel der vorliegenden Arbeit ist es deshalb, eine automatische Erkennung, Klassifikation und Kartierung von Straßenoberflächenschäden aus den Daten eines speziellen Mobilen Laserscanning Systems zu realisieren sowie deren Potentiale und Grenzen im Kontext der kommunalen Straßenerhaltung zu untersuchen. Dabei kommen sowohl Methoden der Geoinformatik und der digitalen Geländeanalyse sowie Methoden aus dem Bereich der Mustererkennung und des maschinellen Lernens zum Einsatz. Um tiefe neuronale Netzwerkmodelle zur semantischen Segmentierung von Oberflächenmodellen trainieren zu können, werden umfangreiche Datensätze aufgebaut und eine Transfer Learning Strategie angewendet. Die Leistungsfähigkeit des entwickelten Analysesystems wird mithilfe von Vergleichsmessungen durch zwei Referenzsysteme evaluiert, die auf nationaler und internationaler Ebene den Stand der Technik im Bereich Straßenzustandserfassung repräsentieren.

Die Ergebnisse zeigen, dass es möglich ist, mit dem untersuchten LiDAR-System, eine ganzheitliche Datengrundlage (Ebenheit und Oberflächenschäden) für die kommunale Straßenzustandserfassung aufzunehmen und mit dem entwickelten Deep-Learning-Ansatz eine Zustandsbewertung zu erzeugen, die dasselbe Qualitätsniveau einer durch geschultes Personal durchgeführten Inspektion aufweist und dabei deutlich weniger Prozesszeit in Anspruch nimmt.

Abstract

For a systematic and efficient maintenance of municipal road infrastructure, a detailed knowledge of the prevailing road condition is of central importance. Therefore, standards have been established in Germany to ensure objectivity in pavement management. Despite innovative technical developments and a high degree of automation in the field of mobile data acquisition, the manual-visual inspection of surface damage from camera image data dominates in the field of data evaluation. This process is not only time-, labour- and cost-intensive, but also uses exclusively photometric information for the inspection process. Mobile laser scans are only used to record and evaluate the evenness of a road, although sufficiently dense surface scans would allow a simultaneously three-dimensional description of the surface and its damage. Such a database would be beneficial for both manual damage detection and automation of the inspection process. Studies show that the manual recording of surface damage from camera image data is the greatest source of error in the process of road condition assessment and endangers the goal of creating an objective and comparable data basis for pavement management.

Therefore, the aim of the present work is to realise an automatic detection, classification and mapping of road surface damage from the data of a special mobile laser scanning system and to investigate its potentials and limitations in the context of municipal road maintenance. For this purpose, methods of geoinformatics and digital terrain analysis as well as methods from the field of pattern recognition and machine learning are used. In order to train deep neural network models for semantic segmentation of surface models, extensive data sets are built and a transfer learning strategy is applied. The performance of the developed analysis system is evaluated with the help of comparative measurements by two reference systems that represent the state of the art in the field of road condition monitoring on a national and international level. The results show that it is possible to record a holistic data basis (evenness and surface damage) for the municipal road condition monitoring with the investigated LiDAR system and to generate a condition assessment with the presented deep learning approach, which has the same quality level as a conventional inspection carried out by trained operators and requires significantly less process time.

Inhaltsverzeichnis

Vorwort	I
Zusammenfassung.....	II
Abstract.....	III
Inhaltsverzeichnis.....	IV
Abbildungsverzeichnis.....	VIII
Tabellenverzeichnis.....	XIV
Formelverzeichnis.....	XV
Abkürzungsverzeichnis.....	XVI
1 Einleitung	1
1.1 Fragestellungen und Ziele	3
1.2 Aufbau der Arbeit.....	5
2 Einführung in Themen und Technologien	8
2.1 Systematische Straßenerhaltung	8
2.1.1 Organisationen	10
2.1.2 Bundesfernstraßennetz	10
2.1.2.1 Ebenheit	11
2.1.2.2 Griffigkeit	14
2.1.2.3 Substanzmerkmale.....	16
2.1.2.4 Zustandsbewertung.....	18
2.1.3 Kommunales Straßennetz.....	20
2.1.3.1 Ebenheit	22
2.1.3.2 Substanzmerkmale.....	23
2.1.3.3 Zustandsbewertung.....	25
2.1.4 Zuverlässigkeit der manuellen Erfassung von Oberflächenschäden.....	27
2.2 Mobile Laserscanning.....	31
2.2.1 Eigenschaften des Laserlichts	33
2.2.2 Elektrooptische Entfernungsmessung	37
2.2.3 Scansysteme	40
2.2.4 Satellitengestützte Positionsbestimmung	41

2.2.4.1	Global Positioning System	42
2.2.4.2	Differentielle Korrekturverfahren.....	44
2.2.4.3	Positionierungssysteme mit Inertialsystem	46
2.2.5	Georeferenzierung von Sekundärsensordaten	50
2.2.6	Fehlereinflüsse	52
2.2.6.1	Laserscanner	52
2.2.6.2	Satellitengestützte Positionsbestimmung	54
2.2.6.3	Inertialsystem.....	56
2.3	Digitale Oberflächenmodellanalyse in Geo-Informationssystemen.....	57
2.3.1	Oberflächenmodellierung.....	57
2.3.1.1	Arten von Oberflächenmodellen	59
2.3.1.2	Räumliche Interpolationsverfahren	62
2.3.2	Geomorphometrie.....	64
2.4	Künstliche Intelligenz, maschinelles Lernen und Deep Learning.....	67
2.4.1	Künstliche neuronale Netze	69
2.4.1.1	Aufbau und Funktionsweise	70
2.4.1.2	Lernen eines neuronalen Netzwerkmodells.....	73
2.4.1.3	Generalisierungsfehler und Regularisierungsstrategien	78
2.4.2	Semantische Bildsegmentierung mithilfe von Deep Learning.....	82
2.4.2.1	Convolutional Neural Networks.....	84
2.4.2.2	Evolution von CNN Architekturen.....	88
2.4.2.3	Fully Convolutional Networks	92
2.4.2.4	Encoder/Decoder Networks.....	95
2.4.3	Transfer Learning	96
2.4.4	Geospatial Artificial Intelligence	99
2.5	Forschungsstand zur Automatisierung im Bereich Straßenzustandserfassung	101
2.5.1	Ebenheitsindikatoren aus Mobile Laserscanning Daten	102
2.5.2	Automatisierte Erfassung von Oberflächenschäden.....	104
3	Datengrundlage	110
3.1	Datenerfassung	110
3.1.1	Mobiles Laserscanning System I.R.I.S	110
3.1.1.1	Inertiales Positionierungssystem	112
3.1.1.2	Laserscanner	115
3.1.1.3	Kalibrierung und Boresight-Alignment.....	118

3.1.2	Handgeführte Satellitenvermessung.....	122
3.2	Datensätze.....	125
3.2.1	Aspekt: absolute Positionierungsgenauigkeit.....	125
3.2.2	Aspekt: Präzision des Messsystems	128
3.2.3	Aspekt: Oberflächenmodellierung	131
3.2.4	Aspekt: Erkennbarkeit von Oberflächendiskontinuitäten	134
3.2.5	Aspekt: Trainings- und Testdaten für Deep-Learning-Modelle.....	137
3.2.6	Aspekt: Evaluierung des Deep Learning basierten Analysesystems.....	138
4	Methodik	143
4.1	Arbeitsumgebung.....	143
4.1.1	Trajektorienoptimierung und 3D-Punktwolkenerzeugung.....	144
4.1.2	3D-Prozessierung und Visualisierung	145
4.1.3	Geoprozessierungen	145
4.1.4	Training und Evaluierung von Deep-Learning-Modellen.....	149
4.1.5	Erzeugung regelwerkskonformer TP3-Daten.....	152
4.2	Deterministische räumliche Interpolationsverfahren.....	154
4.2.1	Delaunay-Triangulation	157
4.2.2	Natural Neighbor.....	159
4.2.3	Inverse Distance Weighted.....	162
4.2.4	Modified Quadratic Shepard	164
4.2.5	Thin Plate Spline	167
4.3	Geomorphometrie	169
4.3.1	Analytical Hillshading.....	172
4.3.2	Topographic Openness	176
4.3.3	Black Top Hat Transformation	179
4.3.4	Topographic Position Index	182
4.4	Das DeepLab System.....	185
4.4.1	DeepLabv1 und DeepLabv2.....	186
4.4.2	DeepLabv3 und DeepLabv3+	190
4.5	Evaluation der Ergebnisse	193
4.5.1	Nutzwertanalyse	193
4.5.2	Kreuzvalidierung.....	194
4.5.3	Evaluation von semantischen Segmentierungsmodellen	195
4.5.4	Evaluation der automatischen Zustandserfassungsqualität	198

5	Ergebnisse	201
5.1	Absolute Positionierungsgenauigkeit	201
5.1.1	Analyse der Abweichung zwischen Passpunkten und Laserscan	201
5.1.2	Erkenntnisse zur absoluten Positionierungsgenauigkeit	203
5.2	Präzision des Messsystems	205
5.2.1	Analyse des Prüfkörpers.....	205
5.2.2	Erkenntnisse zur Präzision des Messsystems	208
5.3	Oberflächenmodellierung	209
5.3.1	Parametrisierung der Interpolationsverfahren	210
5.3.2	Vergleich der Interpolationsverfahren	212
5.3.3	Erkenntnisse zur Oberflächenmodellierung	215
5.4	Erkennbarkeit von Oberflächendiskontinuitäten	217
5.4.1	Analysekonzept	217
5.4.2	Erkenntnisse zur Erkennbarkeit von Oberflächendiskontinuitäten	218
5.5	Deep Learning basierte Oberflächenanalyse	219
5.5.1	Erstellung eines Trainings- und Evaluationsdatensatzes	220
5.5.2	Training und Evaluation von DeepLabv3+ Modellen.....	225
5.5.2.1	Voruntersuchungen.....	226
5.5.2.2	Finale Trainingspipeline	232
5.5.3	Automatische Kartierung von Oberflächenschäden.....	238
5.5.4	Vergleich zu Referenzsystemen der Zustandserfassung	241
5.5.5	Erkenntnisse zur Oberflächenanalyse mithilfe von Deep Learning	253
5.6	Zusammenfassung der Ergebnisse.....	254
6	Schlussbetrachtung	257
	Literaturverzeichnis.....	260
	Anhang	293

Abbildungsverzeichnis

Abbildung 1: Ablaufschema des systematischen Erhaltungsmanagements	9
Abbildung 2: Beispiele für schnellfahrende Multisensorsysteme	11
Abbildung 3: Lasteintrag und Schädigung der Straße durch Längsunebenheiten	12
Abbildung 4: Schematische Darstellung des Wasserrückhalts in Spurrinnen	12
Abbildung 5: Messprinzip der Ebenheit im Längsprofil und Querprofil	13
Abbildung 6: Texturbereiche in Abhängigkeit zur Wellenlänge	14
Abbildung 7: Bestimmung der Griffigkeit durch Seitenkraftmessung	15
Abbildung 8: Oberflächenbilder eines BAST-zertifizierten Messfahrzeuges	17
Abbildung 9: Systematik der Erfassung von Oberflächenschäden	18
Abbildung 10: Beispiel einer Normierungsfunktion zur Berechnung von Zustandswerten	19
Abbildung 11: Bewertungsschema zur Ermittlung des Gesamtwertes	20
Abbildung 12: Visuelle Zustandserfassung durch Begehung	21
Abbildung 13: Relevante Ebenheitsindikatoren für kommunale Straßen	23
Abbildung 14: Relevante Oberflächenschäden für kommunale Straßen	24
Abbildung 15: Schema der messtechnischen Zustandsbewertung kommunaler Straßen	26
Abbildung 16: Qualität der manuellen Auswertung von Rissen	30
Abbildung 17: Trägerfahrzeugunabhängiges Mobile-Mapping-System	32
Abbildung 18: Mobile Laserscanning Systeme für die Straßenzustandserfassung	32
Abbildung 19: Laserstrahlausbreitung mit entfernungsabhängiger Strahldivergenz	35
Abbildung 20: Einfluss von Oberflächenformen auf die Signalform einer Lasermessung	35
Abbildung 21: Reflexionen an Oberflächen unterschiedlicher Rauheit	36
Abbildung 22: Prinzip der laserbasierten optischen Triangulation	37
Abbildung 23: Prinzip der Distanzmessung mittels Impulslaufzeitverfahren	38
Abbildung 24: Prinzip der Amplitudenmodulation	39
Abbildung 25: Prinzip des Phasenvergleichsverfahrens	40
Abbildung 26: Systemtypen und Abtastmuster von Laserscannern	41
Abbildung 27: Prinzip der Pseudostreckenmessung	43
Abbildung 28: Prinzip des differentiellen GNSS	45
Abbildung 29: Beispiel einer SAPOS-Referenzstationsantenne	46
Abbildung 30: Aufbau von Strapdown-Systemen	48
Abbildung 31: Prinzip der Korrektur driftender IMU Daten durch GNSS Messungen	48
Abbildung 32: Aufbau eines integrierten INS/GNSS-Positionierungssystems	49
Abbildung 33: Transformationen im Zuge der direkten Georeferenzierung	50

Abbildung 34: Eingeschränkte Erfassbarkeit von Kanten	53
Abbildung 35: Auszug aus dem Datenblatt eines Laserscanners.....	54
Abbildung 36: GNSS Abschattung und Mehrwegeausbreitung	55
Abbildung 37: Gitter- und Dreiecksmodell im Vergleich.....	60
Abbildung 38: Aufbau eines Curved Regular Grids.	61
Abbildung 39: Prinzipdarstellung eines Variogramms	63
Abbildung 40: Quelldisziplinen und Anwendungsdomänen der Geomorphometrie.	64
Abbildung 41: Ablaufschema einer geomorphometrischen Analyse.....	65
Abbildung 42: Künstliche Intelligenz und Deep Learning	67
Abbildung 43: Übertragung von Reizen in Nervenzellen.....	70
Abbildung 44: Schematische Darstellung eines künstlichen neuronalen Netzes.....	71
Abbildung 45: Innere Struktur eines Neurons.....	72
Abbildung 46: Aktivierungsfunktionen in künstlichen Neuronalen Netzen.....	73
Abbildung 47: Ablaufschema zum Lernen eines neuronalen Netzwerkmodells	74
Abbildung 48: Gradientenabstieg auf der Fehleroberfläche einer Kostenfunktion	76
Abbildung 49: Einfluss der Lernrate auf das Training eines Modells	77
Abbildung 50: Loss Kurven auf Trainings- und Validierungsdaten	79
Abbildung 51: Schema der Regularisierungstechnik Dropout.....	81
Abbildung 52: Einfluss der Regularisierungstechnik Batch-Normalization.....	82
Abbildung 53: Faltung als zentrale Operation in einer Convolution Schicht	85
Abbildung 54: Schema von Zero Padding	86
Abbildung 55: Beispiel für eine Pooling Operation.....	87
Abbildung 56: Architektur und das Training eines Convolutional Neural Network	88
Abbildung 57: Schematische Darstellung der AlexNet CNN Architektur	89
Abbildung 58: Trainings- und Testfehler unterschiedlich tiefer Netze.....	90
Abbildung 59: Struktur eines ResNets	90
Abbildung 60: Prinzip der Depthwise Separable Convolution	92
Abbildung 61: Schematische Darstellung eines Fully Convolutional Networks.....	93
Abbildung 62: Prinzip des Upsamplings.....	93
Abbildung 63: Kombination von Ausgaben an unterschiedlich tiefen Stellen im Netz	94
Abbildung 64: Schematische Darstellung der SegNet-Architektur	96
Abbildung 65: Faltungskernels in der ersten Schicht der AlexNet Architektur.....	97
Abbildung 66: Transfer Learning Arten.....	98
Abbildung 67: Übersicht GeoAI	99

Abbildung 68: Querebenheitsindikatoren auf Basis eines CRG Querprofils.....	103
Abbildung 69: Risserkennung durch schwellenwertbasierte Bildanalyse	106
Abbildung 70: Risserkennung durch merkmalsbasiertes Machine Learning	107
Abbildung 71: Risserkennung mithilfe eines Bildausschnittsbasierten Faltungsnetzes	108
Abbildung 72: Externe Einflüsse auf Kamerabildern von Straßenoberflächen.....	109
Abbildung 73: Das Mobile Laserscanning System I.R.I.S 16	111
Abbildung 74: Übersicht zum modularen Aufbau des I.R.I.S Systems.	112
Abbildung 75: Komponenten des POS LV Systems.....	113
Abbildung 76: Prinzip der Applix IN-Fusion Architektur	114
Abbildung 77: Laserscanner Pavement Profile Scanner Plus	115
Abbildung 78: Rotierendes Polygonprisma als Ablenkeinheit des PPS	116
Abbildung 79: Kalibrierfeld zur Bestimmung des Boresight Vektors	118
Abbildung 80: Kalibrierung des PPS mittels eines Stahl-Prüfkörpers.....	119
Abbildung 81: Auszug aus dem Kalibrierschein des Stahl-Prüfkörpers.....	120
Abbildung 82: Streuung der Höhenwertmessungen auf dem Prüflineal	121
Abbildung 83: Intensitätswerte der Lasermessungen auf dem Prüflineal	122
Abbildung 84: GNSS-RTK Rover Leica GS18 T und Bedienelement Leica CS20	123
Abbildung 85: Passpunktaufnahme mit einem Leica GNSS-RTK Rover	124
Abbildung 86: Lagekarte der Passpunkte.....	125
Abbildung 87: Beispielbilder von Passpunktmarkierungen.....	126
Abbildung 88: Messfahrzeug I.R.I.S 13 bei der Messung der Passpunkt Strecken	127
Abbildung 89: Passpunktmarkierung im Frontkamerabild und PPS-Laserscan	127
Abbildung 90: 3D-Rekonstruktion des realen Prüfkörpers	128
Abbildung 91: Datenerfassung zur Untersuchung der Präzision des Messsystems	129
Abbildung 92: Aufnahmesituation während der Messungen des Prüfkörpers.....	130
Abbildung 93: Positionierungsfehler der Trajektorie während der Prüfkörpermessung	130
Abbildung 94: Längsschnitt durch die Punktwolken der acht Prüfkörpermessungen	131
Abbildung 95: Testareal 1 für die Untersuchungen zur Oberflächenmodellierung	132
Abbildung 96: Testareal 2 für die Untersuchungen zur Oberflächenmodellierung	133
Abbildung 97: Histogramme der Punktdichte für Testareal 1 und Testareal 2.....	134
Abbildung 98: Prüfkörper zur Bestimmung der Erkennbarkeit von Spalten	135
Abbildung 99: Messsystem I.R.I.S. 16 während der Prüfkörpermessung	136
Abbildung 100: 3D-Punktwolke des Prüfkörpers bei 30 km/h	136
Abbildung 101: Karte der räumlichen Verteilung von Trainings- und Testdaten	137

Abbildung 102: Beispiel eines TP3-Oberflächenbildes	139
Abbildung 103: I.R.I.S 16 während der Aufnahme der Evaluierungsstrecke	140
Abbildung 104: Beispiel für einen Laserscan der Evaluierungsstrecke	140
Abbildung 105: I.R.I.S. 17 während der Aufnahme der Evaluierungsstrecke	141
Abbildung 106: Oberflächenbild aus den LCMS2 Sensoren des I.R.I.S 17 Systems	142
Abbildung 107: User Interface der Software Applanix POSpac MMS 8.6	144
Abbildung 108: User Interface der Software CloudCompare.	145
Abbildung 109: User Interface der Software SAGA GIS.	146
Abbildung 110: User Interface der Software QGIS	147
Abbildung 111: GeoPackage	148
Abbildung 112: User Interface der Software ScaLabel	149
Abbildung 113: Beispiel für ein Jupyter Notebook	150
Abbildung 114: User Interface von TensorBoard	152
Abbildung 115: Auswertearbeitsplatz zur Erfassung von Oberflächenschäden	153
Abbildung 116: User Interface der Software TP3-Websolution.	154
Abbildung 117: Räumliche Verteilungsmuster von Messwerten	154
Abbildung 118: Prinzip der Delaunay-Triangulation	157
Abbildung 119: Regelmäßiges Gittermodell aus Delaunay-Triangulation	158
Abbildung 120: Prinzip der Sibson Natural Neighbors Interpolationsmethode	159
Abbildung 121: Prinzip der Non-Sibson Natural Neighbors Interpolationsmethode	161
Abbildung 122: Prinzip der Inverse Distance Weighted Interpolationsmethode	163
Abbildung 123: Beispiel für den "bull's-eye" Effekt bei der IDW Interpolation	164
Abbildung 124: Beispiel einer durch MQS Interpolation erzeugte Oberfläche.	167
Abbildung 125: Einfluss des Regularisierungsparameters bei der TPS Interpolation.	169
Abbildung 126: Analogie zwischen Formelementen des Georeliefs und Straßenschäden	170
Abbildung 127: Profilschnitt durch ein Digitales Geländemodell	171
Abbildung 128: Profilschnitt durch ein PPS Oberflächenmodell	172
Abbildung 129: Prinzip der Berechnung eines Hillshades	173
Abbildung 130: Hillshade für ein Airborne LiDAR DGM	174
Abbildung 131: Hillshade für ein PPS Oberflächenmodell	175
Abbildung 132: Kommandozeilenaufruf des Analytical Hillshading Tools	175
Abbildung 133: Prinzip der Topographic Openness Berechnung	176
Abbildung 134: Positive Openness für ein Airborne LiDAR DGM	177
Abbildung 135: Positive Openness für ein PPS Oberflächenmodell	178

Abbildung 136: Kommandozeilenaufruf des Topographic Openness Tools	179
Abbildung 137: Funktionsweise der Black Top Hat Transformation	179
Abbildung 138: Black Top Hat Transformation für ein Airborne LiDAR DGM	180
Abbildung 139: Black Top Hat Transformation für ein PPS Oberflächenmodell	181
Abbildung 140: Kommandozeilenaufruf des Black Top Hat Approach Tool	181
Abbildung 141: Prinzip des Topographic Position Index	182
Abbildung 142: Topographic Position Index für ein Airborne LiDAR DGM.....	183
Abbildung 143: Topographic Position Index für ein PPS Oberflächenmodell	184
Abbildung 144: Kommandozeilenaufruf des Topographic Position Index Tools	185
Abbildung 145: Beispiel für Kernel-Formen zur TPI Berechnung.....	185
Abbildung 146: Prinzip der Atrous Convolution	187
Abbildung 147: Funktionsprinzip von DeepLabv1	188
Abbildung 148: Prinzipskizze des Atrous Spatial Pyramid Pooling	190
Abbildung 149: Atrous Separable Convolution in DeepLabv3+	191
Abbildung 150: Encoder/Decoder Struktur von DeepLabv3+	192
Abbildung 151: Segmentierungsergebnisse vortrainierter DeepLabv3+ Modelle.....	193
Abbildung 152: Beispiel für Kreuzvalidierung bei der Oberflächeninterpolation.....	195
Abbildung 153: Interpretationsprinzip einer Konfusionsmatrix	196
Abbildung 154: Prinzip der Evaluation über ein gemeinsames Erfassungsraster.....	200
Abbildung 155: Post-Processing der Trajektorien	202
Abbildung 156: Bestimmung des 3D-Abstands zwischen Passpunkt und Laserscan.....	203
Abbildung 157: Absolute Positionierungsqualität der Laserscans.....	204
Abbildung 158: Vergleich von Echtzeit- und post-prozessierter Trajektorie	204
Abbildung 159: Laserscans der acht Wiederholungsmessungen des Prüfkörpers	206
Abbildung 160: ICP Matching der Scandaten auf das 3D-Modell des Prüfkörpers	206
Abbildung 161: Profil durch das Prüfkörpermodell und die gematchten Laserscans.....	208
Abbildung 162: Ausrundungseffekt an den Kanten des Prüfkörpers.....	209
Abbildung 163: Fehlerverteilung der Laserscans auf einer Prüfkörperfläche	213
Abbildung 164: Interpolationsmethoden zur Erzeugung von Intensitätsbildern	216
Abbildung 165: Erkennbarkeit von Spalten auf Basis eines PPS Höhenmodells	217
Abbildung 166: Profilschnitte durch PPS Höhenmodelle des Prüfkörpers.....	218
Abbildung 167: Detailgrad der Labelmasken für alle relevanten Oberflächenschäden.....	221
Abbildung 168: Zusammensetzung des Z-A/B Datensatzes	222
Abbildung 169: Zusammensetzung des Z-P Datensatzes	222

Abbildung 170: Zusammensetzung des I-A/B Datensatzes	223
Abbildung 171: Labelgenauigkeit mehrfach manuell annotierter Schadstellen	224
Abbildung 172: Einfluss der Trainingsdauer auf die Leistungsfähigkeit eines Modells	229
Abbildung 173: Vergleich geomorphometrischer Raster als Eingabe für DeepLabv3+	230
Abbildung 174: Stretch Werte für die Konvertierung von TPI Rastern in 8 bit Daten.....	232
Abbildung 175: Konfusionsmatrix des Z-A/B Modells	234
Abbildung 176: Konfusionsmatrix des I-A/B Modells	235
Abbildung 177: Konfusionsmatrix des Z-P Modells	237
Abbildung 178: Automatische Zustandsbewertung auf Basis von MLS-Daten	239
Abbildung 179: Georeferenzierte Ergebnisdatei mit Zustandsgrößen und -werten.....	240
Abbildung 180: Manuelle Auswertung von begradigten PPS Oberflächenmodellen	242
Abbildung 181: Automatische Auswertung von begradigten PPS Oberflächenmodellen.....	243
Abbildung 182: Vergleich von automatischer und manueller Schadenserfassung.	245
Abbildung 183: Evaluierungsstrecke der drei verglichenen Messsysteme	246
Abbildung 184: User Interface der Software LcmsRoadInspect	247
Abbildung 185: Vergleich zu Referenzsystemen auf der Evaluierungsstrecke	249
Abbildung 186: Aufgelegte Flickstelle im Oberflächenbild des TP3-Systems	250
Abbildung 187: LCMS2 Intensitätsbild mit überlagerten Schadstellen und Objekten	251

Tabellenverzeichnis

Tabelle 1: Oberflächenschäden im kommunalen Straßennetz	25
Tabelle 2: Zustandsklassen zur grafischen Darstellung in Karten und Diagrammen	27
Tabelle 3: Menschliche Fehlerwahrscheinlichkeit	28
Tabelle 4: Geforderte Auswertegenauigkeit im Rahmen der Eigenüberwachung	29
Tabelle 5: Fehlereinflüsse der satellitengestützten Positionsbestimmung	55
Tabelle 6: Größenordnung IMU-spezifischer Fehler im Mobile-Mapping-Bereich.....	57
Tabelle 7: Wichtige Begriffe der Geomorphometrie	66
Tabelle 8: Ansätze zur automatischen Erkennung von Straßenschäden	105
Tabelle 9: Technische Spezifikation des Pavement Profile Scanners	117
Tabelle 10: Technische Daten der verwendeten GPU.....	143
Tabelle 11: Überblick über die untersuchten Interpolationsverfahren.....	156
Tabelle 12: Ergebnis zur Untersuchung der Präzision des Messsystems.....	207
Tabelle 13: Überblick über wichtige Charakteristika der beiden Testareale.	210
Tabelle 14: Ergebnis des Vergleichs von räumlichen Interpolationsmethoden	214
Tabelle 15: Ergebnis der Untersuchung zur Erkennbarkeit von Spalten	219
Tabelle 16: Hyperparameter zum Training von DeepLabV3+ Modellen	227
Tabelle 17: Streuung von Trainingsprozessen	228
Tabelle 18: Einfluss der geomorphometrischen Vorverarbeitungen von Höhenmodellen	231
Tabelle 19: Globale Bewertung der trainierten DeepLabV3+ Modelle	233
Tabelle 20: Klassenweise IoU Werte der trainierten Modelle	238
Tabelle 21: Vergleich von manueller und automatischer Auswertung	244
Tabelle 22: Übereinstimmung der verglichenen Systeme für einzelne Schadensarten.....	252
Tabelle 23: Zusammenfassung der Ergebnisse (Teil 1 von 2)	254
Tabelle 24: Zusammenfassung der Ergebnisse (Teil 2 von 3)	255
Tabelle 25: Zusammenfassung der Ergebnisse (Teil 3 von 3)	256

Formelverzeichnis

Formel 1: Georeferenzierung von Laserscannermessungen	51
Formel 2: Geostatistische Grundformel	62
Formel 3: Polynom zur Korrektur des lateralen Intensitätsabfalls.....	121
Formel 4: Lineare Interpolation bei der Delaunay Triangulation	158
Formel 5: Natural Neighbors Interpolation (Sibson Methode).....	160
Formel 6: Gewichtungskoeffizient (Sibson Methode)	160
Formel 7: Natural Neighbors Interpolation (Non-Sibson Methode).....	160
Formel 8: Gewichtungskoeffizient (Non-Sibson Methode).....	161
Formel 9: Inverse Distance Weighted Interpolation	162
Formel 10: Modified Quadratic Shepard Interpolation.....	165
Formel 11: Gewichtungswert und Knotenfunktion (Modified Quadratic Shepard)	165
Formel 12: Koeffizienten der Knotenfunktion (Modified Quadratic Shepard)	165
Formel 13: Gewichtung der Koeffizienten (Modified Quadratic Shepard)	166
Formel 14: Kreisradien (Modified Quadratic Shepard)	166
Formel 15: Interpolation mit radialer Basisfunktion.....	167
Formel 16: Thin Plate Splines als Basisfunktion	168
Formel 17: Thin Plate Splines als Minimierungsproblem	168
Formel 18: F1	197
Formel 19: Intersection over Union	197
Formel 20: Matthews Correlation Coefficient	198
Formel 21: Root Mean Squared Error.....	212

Abkürzungsverzeichnis

A	Asphalt		
AFLI	Aufgelegte Flickstelle	DGM	Digitales Geländemodell
AMA	Ausmagerungen	DGPS	Differential GPS
		DGNSS	Differential GNSS
API	Application Programming Interface	DHM	Digitales Höhenmodell
ASPP	Atrous Spatial Pyramid Pooling	DL	Deep Learning
ASTM	American Society for Testing and Materials	DMI	Distance Measurement Instrument
AUN	Unebenheitsmaß	DOP	Digitales Orthophoto
AUS	Ausbruch	DTM	Digital Terrain Model
B	Beton	E EMI	Empfehlungen für das Erhaltungsmanagement von Innerortsstraßen
		EAB	Eckabbrüche
BASSt	Bundesanstalt für Straßenwesen	EDM	elektronisches Distanzmessgerät
BDS	chinesisches Satellitennavigationssystem		
BIN	Bindemittelanreicherung	EFLI	Eingelegte Flickstelle
BLP	Bewertetes Längsprofil	EMV	Elektromagnetische Verträglichkeit
BMVI	Bundesministerium für Verkehr und digitale Infrastruktur	F1	harmonisches Mittel aus Precision und Recall
BN	Batch Normalization	FC	Fully Connected
BRU	schadhafte/fehlende Pflasterelemente	FCN	Fully Convolutional Network
BSD	Berkeley Software Distribution	FGSV	Forschungsgesellschaft für Straßen- und Verkehrswesen e. V.
BTE	bituminöser Teilersatz	FLI	Flickstelle
BTHT	Black Top Hat Transform	FN	False Negatives
C/A	Coarse/Acquisition	FOG	faseroptische Kreisel
CCD	Charge-Coupled Device	FP	False Positives
CNN	Convolutional Neural Networks	GeoAI	Geospatial Artificial Intelligence
COCO	Common Objects in Context	GEW	Gesamtwert
CPU	Central Processing Unit	GIS	Geo-Informationssystem
CRF	Conditional Random Field	GLCM	Gray-Level Co-Occurrence Matrix
CRG	Curved Regular Grid	GLONASS	russisches Satellitennavigationssystem
CVPR	Computer Vision and Pattern Recognition Conference	GNSS	Globales Navigationssatellitensystem
CW	Continous Wave	GPL	General Public License
DACH	Deutschland, Österreich und Schweiz	GPRS	General Packet Radio Service

GPS	Global Positioning System	MEMS	mikroelektronisch-mechanische Systeme
GPU	Graphics Processing Unit	mIoU	mean Intersection over Union
GRI	Griffigkeitswerte	ML	Machine Learning
GVL	Lagerungsschäden	MLS	Mobile Laser Scanning
HOG	Histogram of Oriented Gradients	MQS	Modified Quadratic Shepard
HRM	High-Speed Road Monitoring	MSE	Mean Squared Error
HS	Hillshade	MSPH	Maximum der Mittelwerte der rechten und linken fiktiven Wassertiefe
I	Reflexionsintensität	MSPT	Maximum der Mittelwerte der rechten und linken Spurrinnentiefe
I.R.I.S	Integrated Road Information System	NN	Natural Neighbors
ICP	Iterative Closest Point	NTR	Nester und Abplatzungen
IDW	Inverse Distance Weighted	NTRIP	Networked Transport of RTCM via Internet Protocol
IGS	International GNSS Service	OBIA	Object-based Image Analysis
ILSVRC	ImageNet Large-Scale Visual Recognition Challenge	OGC	Open Geospatial Consortium
IMU	inertiale Messeinheit	ONA /	Offene Arbeitsnaht
INS	Inertialsystem	ONF	
IoU	Intersection over Union	OPF	offene Pflasterfugen
JSON	JavaScript Object Notation	P	Pflaster
KAS	Kantenschäden	PAMA	Gefügebrauflösung auf Pflasterelementen
KI	Künstliche Intelligenz	PASCAL	Visual Object Classes Challenge
KNN	Künstliches Neuronales Netz	VOC	
L04i	Simulation der gleitenden 4 m Richtlatte	PFLI	Flickstellen auf Pflaster mit Ersatzmaterial
L1 / L2	GNSS Trägerwellen bzw. Frequenzbänder	PGR	Simulation gleitende 4 m Richtlatte
LBP	Local Binary Patterns	PO	Positive Openness
LCMS	Laser Crack Measurement System	ppm	Parts Per Million
LGPL	Lesser General Public Licence	PPS	Pavement Profile Scanner
LiDAR	Light Detection and Ranging	PSP	Pyramid Scene Parsing
LQR	Längs- und Querrisse	QN	Querneigung
LWI	Längsebenheitswirkindex	RBF	Radiale Basisfunktion
MAE	Mean Absolute Error	ResNet	Residual Neural Network
MCC	Matthews Correlation Coefficient	RGB	Rot, Grün und Blau
		RISS	Risse
		RLG	Ringlasergyroskope

RMS	Root Mean Square	TP3	Teilprojekt Substanzmerkmale
RMSE	Root Mean Squared Error	TP4	Bewertung und standardisierte Auswertung
RSFA	Restschadensfläche Asphalt	TPI	Topographic Position Index
RSFB	Restschadensfläche Beton	TPS	Thin Plate Spline
RTK	Real Time Kinematic	TPU	Tensor Processing Unit
S01 / S03 / S10 / S30	Standardabweichung der Differenzen zum gleitenden Mittelwert bei Mittelungslänge 1, 3, 10 und 30 m	TWGEB	Teilzielwert Gebrauch
SAPOS	Satellitenpositionierungsdienst der deutschen Landesvermessung	TWSUB	Teilzielwert Substanz
SE	Strukturelement	VGG	Visual Geometry Group
SegNet	semantic segmentation network	VRS	Virtual Reference Station
SGD	stochastische Gradientenabstieg	ZW	Zustandswert (Note)
SKM	Seitenkraftmessung mittels schräggestelltem Rad	ZWAUN	Zustandswert für AUN
SM4L	Simulation der gleitenden 4 m Richtlatte	ZWRISS	Zustandswert für RISS
SM4L-A	Mittelwert der Simulation der gleitenden 4 m Richtlatte	ZWRSFA	Zustandswert für RSFA
SM4L-M	Maximalwert der Simulation der gleitenden 4 m Richtlatte		
SPH	fiktive Wassertiefe		
SPT	Spurrinnentiefe		
SQL	Structured Query Language		
SVM	Support Vector Machine		
TIN	Triangulated Irregular Network		
TN	True Negatives		
TOF	Time Of Flight		
TP	True Positives		
TP Eben	Technische Prüfvorschriften für Ebenheitsmessungen auf Fahrbahnoberflächen in Längs- und Querrichtung		
TP0	Teilprojekt Vorbereitung der Straßennetzdaten		
TP1	Teilprojekt Ebenheit		
TP2	Teilprojekt Griffigkeit		

1 Einleitung

Die Erhaltung des bestehenden Straßennetzes ist von zentraler Bedeutung, um eine leistungs- und zukunftsfähige Straßeninfrastruktur zu gewährleisten. Dies gilt insbesondere für ein Land wie Deutschland, das im Zentrum des europäischen Personen- und Gütertransports liegt. Mit der wachsenden Bedeutung des Online-Handels ist auch eine Zunahme des Warentransports über die Straße verbunden. Über 70 % der Beförderungsleistung im Güterverkehr entfällt auf den Straßenverkehr (RADKE 2019). Die Verkehrszunahme, vor allem der wachsende Anteil des Schwerverkehrs, ist dabei ein treibender Faktor für die Verschlechterung der baulichen Substanz von Fahrbahnoberflächen. Gerade der Zustand der öffentlichen Straßen im kommunalen Netz verschlechtert sich entsprechend ihrer Altersstruktur und der Verkehrsbelastung zunehmend.

Um auf Zustandsveränderungen zeitnah und angemessen reagieren zu können, ist eine genaue Erfassung des Straßenzustands nötig und damit essentieller Teil im Prozess der Straßenerhaltung. Eine turnusmäßige Erfassung und Bewertung des Straßenzustands führt dazu, dass Mängel an der Straßenbefestigung frühzeitig entdeckt und oft mit verhältnismäßig geringem Aufwand beseitigt, sowie die verfügbaren Erhaltungsmittel bedarfsgerecht eingesetzt werden können. Im Geschäftsbericht des Bayerischen Kommunalen Prüfungsverbandes wird ausgeführt, dass die Kosten einer verspäteten Sanierung erheblich und offensichtlich gewordener Mängel schnell die Kosten für eine turnusmäßige Zustandsüberwachung übersteigen (SCHMID 2019).

Für das Bundesfernstraßennetz wurde seit der deutschen Wiedervereinigung eine einheitliche und messdatenbasierte Systematik etabliert, um alle vier Jahre qualitätsgesicherte Daten für das Erhaltungsmanagement der ca. 13000 Bundesautobahnkilometer und 38000 Bundesstraßenkilometer bereitzustellen, die sogenannte Zustandserfassung und -bewertung von Straßen (STÖCKERT, SCHMERBECK & LAU 2019). Dabei gelten höchste Qualitätsansprüche im Hinblick auf die Datenaufnahme und -auswertung (FGSV 2006). Für die restlichen Landes-, Staats- und Kreisstraßen mit ca. 178000 km Netzlänge sowie für die Gemeindestraßen mit ca. 600000 km Netzlänge ist dieses durchaus aufwändige Instrument nicht verpflichtend vorgeschrieben (RADKE 2019). Für das nachgeordnete Netz wurden aus diesem Grund offizielle Empfehlungen zur systematischen Zustandserfassung und -bewertung erarbeitet, die die Besonderheiten kommunaler Straßen berücksichtigen (FGSV 2012). In der Praxis ist es jedoch unterschiedlich, wie diese Richtlinien vom jeweiligen Baulastträger umgesetzt werden. Dementsprechend stark

variiert die Qualität, mit der Daten zum Straßenzustand erfasst und in das kommunale Erhaltungsmanagement integriert werden (STEINAUER, UECKERMAN & MAERSCHALK 2006). Besondere Bedeutung für die Erhaltung von Straßen haben Schäden an der Straßenoberfläche. Sie sind der häufigste Grund für bauliche Maßnahmen und Deckenerneuerungen (STASCHKIEWICZ & RADENBERG 2022). Trotz innovativer technischer Entwicklungen und hoher Automation im Bereich der mobilen Datenerfassung (Sensoren, Messfahrzeuge) wird die Datenauswertung von Oberflächenschäden aus Bilddaten der Straßenoberfläche in der Praxis zumeist immer noch von speziell geschultem Personal visuell-manuell durchgeführt. Studien zeigen, dass dies die größte Fehlerquelle im Prozess der Zustandserfassung und -bewertung ist und das Ziel gefährdet, eine objektive und vergleichbare Datengrundlage für das Erhaltungsmanagement zu schaffen (BERTHOLD et al. 2023).

Bei den im kommunalen Bereich zur Erfassung von Oberflächenschäden eingesetzten Sensoren handelt es sich typischerweise um RGB-Kameras. Als passive Sensoren sind Kameras gegenüber aktiver Sensorik mit einigen Nachteilen versehen: Zunächst ist die Qualität der Kamerabilddaten maßgeblich von den vorherrschenden Lichtverhältnissen abhängig (SKAKUJ et al. 2019). Über- bzw. Unterbelichtung infolge von Gegenlicht oder extremen Schattenbereichen können zu einem Informationsverlust in weiten Bildbereichen führen. Ferner steht neben den fotometrischen Informationen der RGB-Kameras keine präzise Tiefeninformation zur Verfügung. Ein Riss oder ein Schlagloch kann also nur erkannt werden, wenn sich der Schaden aufgrund von Farbunterschieden im Bild abzeichnet (Feuchtigkeitsunterschiede, Materialunterschiede, Schattenbildung). Die Verwendung aktiver Sensoren, wie beispielsweise Laserscannern, ist derzeit nur für die Erfassung von Ebenheitsdaten gebräuchlich, obwohl Laserscans zusätzlich zur Informationen über die Reflexionseigenschaften des Zielobjektes auch eine geometrische Information über dessen Oberflächenform enthalten. Risse oder Schlaglöcher können demnach auf Basis ihrer charakteristischen dreidimensionalen Morphologie als solche erkannt und klassifiziert werden. Hierzu bietet es sich an, Oberflächenmodelle auf Basis von Laserscans zu berechnen und mithilfe von Methoden der Geomorphometrie zu analysieren (SESSELMANN, STRICKER & EISENBACH 2019). Da alle aus dem mobil erfassten Laserscan abgeleiteten Datenderivate direkt georeferenziert sind, können sie in Form von Geoobjekten für weiterführende räumliche Analysen genutzt werden.

In vielen Wirtschaftsbereichen und Anwendungsfeldern wird Künstliche Intelligenz (KI) als Werkzeug zur Automatisierung und Optimierung von Abläufen eingesetzt (DESTATIS 2023). Meist übernimmt eine KI dabei arbeitsaufwändige Routineaufgaben, wie beispielsweise die

Überwachung von Produktionsabläufen und die Qualitätssicherung. Darüber hinaus ist KI gut geeignet, um Muster in großen Datenmengen zu finden und für nachgelagerte, höherwertige Anwendungen nutzbar zu machen. Als eine Methode innerhalb des Themenfeldes KI hat das sogenannte “Deep Learning“ die automatisierte Datenanalyse im Bereich Computer Vision revolutioniert (KRIZHEVSKY, SUTSKEVER & HINTON 2012). Seit den 2010er-Jahren konnten nicht zuletzt aufgrund der sich rasant entwickelnden Rechentechnik, insbesondere der Leistungssteigerung im Bereich der Prozessierung von Grafikdaten, erstmals tiefe Convolutional Neural Networks (CNN) trainiert werden, die bei praxisrelevanten visuellen Erkennungsaufgaben ähnliche Fehlerraten wie menschliche Experten erreichen und Entscheidungen dabei wesentlich schneller und objektiver treffen können. Innerhalb eines Jahrzehnts hat sich Deep Learning als wissenschaftlich fundierte Methode im Bereich automatisierter Datenauswertung neben klassischen Ansätzen des Machine Learnings durchgesetzt. Auch in den Bereichen Straßenraumerfassung, Geodäsie und Geoinformatik zählen Deep-Learning-Verfahren und insbesondere der Einsatz von CNN heute zu den etablierten Analysemethoden (NABER et al. 2021; GRUNAU 2022). Da Laserscan-basierte Oberflächenmodelle dieselbe grundlegende Matrixstruktur aufweisen wie Kamerabilder, können sie als Eingabedaten für leistungsfähige Deep Learning Architekturen eingesetzt werden, auch wenn diese ursprünglich für visuelle Erkennungsaufgaben auf Basis anderer Eingabebilddaten konzipiert wurden.

Wenn es möglich wäre, alle für die Straßenerhaltung relevanten Oberflächenschäden aus den Daten eines Mobile Laserscanning Systems zu erkennen und automatisiert auszuwerten, könnte dies einen entscheidenden Beitrag dazu leisten, für die kommunalen Baulastträger eine messtechnisch erfasste und vom menschlichen, subjektiven Fehlereinfluss bereinigte Datenbasis für das Erhaltungsmanagement zur Verfügung zu stellen. Objektivere und schneller verfügbare Ergebnisse einer Straßenzustandserfassung könnten Planungsprozesse optimieren, einen bedarfsgerechten und wirtschaftlichen Einsatz der verfügbaren Mittel zur Straßenerhaltung unterstützen und insgesamt zur Zielerreichung einer besseren Verfügbarkeit der kommunalen Straßeninfrastruktur beitragen.

1.1 Fragestellungen und Ziele

Gegenstand der vorliegenden Arbeit ist es, eine automatische Klassifikation und Kartierung von Oberflächenschäden aus den Daten eines speziellen Mobile Laserscanning Systems zu realisieren, sowie deren Grenzen und Potentiale im Kontext des kommunalen Erhaltungs-

managements zu untersuchen. Dabei kommen sowohl Methoden der digitalen Geländeanalyse und der Geoinformatik als auch Methoden aus dem Bereich der Mustererkennung und des maschinellen Lernens zum Einsatz.

Folgende Forschungsfragen sind für die Arbeit von zentraler Bedeutung:

- 1) Welche Typen von Fahrbahnoberflächenschäden müssen in welcher Art erkannt werden, um eine Zustandserfassung gemäß der derzeit in Deutschland für das kommunale Straßennetz geltenden Richtlinien zu realisieren und können diese aus den LiDAR-Daten eines Mobile Laserscanning Systems detektiert werden?
- 2) Wie hoch ist die absolute Positionsgenauigkeit der mit dem Mobile Laserscanning System erfassten Laserscans der Fahrbahnoberfläche und implizit der daraus abgeleiteten Schadstellen?
- 3) Mit welcher Präzision kann die Geometrie eines Messobjekts abgebildet werden, wenn es wiederholt mit dem Mobile Laserscanning System erfasst wird?
- 4) Welche räumliche Interpolationsmethode ist für die Erzeugung von digitalen Oberflächenmodellen aus 3D-Straßenoberflächenscans im Rahmen der gegebenen Problemstellung und der verfügbaren Daten am besten geeignet?
- 5) Bis zu welchem Ausmaß können Oberflächendiskontinuitäten bzw. Schadstellenstrukturen in Laserscan-basierten Oberflächenmodellen räumlich aufgelöst und als solche erkannt werden?
- 6) Welches Hyperparameter-Setting ist für ein erfolgreiches Training von DeepLabv3+ Modellen auf Basis der vorliegenden Eingabedaten geeignet?
- 7) Welcher Streuung aufgrund von zufälligen Initialisierungsprozessen innerhalb der DeepLabv3+ Netzwerkstruktur unterliegt die Leistungsfähigkeit trainierter Modelle?
- 8) Welche Algorithmen der digitalen Geländeanalyse eignen sich zur Erkennung von Straßenschäden aus Oberflächenmodellen und welchen Einfluss hat die Wahl einer geomorphometrischen Vorverarbeitung der Höheninformation auf die Leistungsfähigkeit eines darauf trainierten DeepLabv3+ Modells?
- 9) Mit welcher Qualität ist eine vollautomatische Erfassung von Oberflächenschäden unter Nutzung des semantischen Segmentierungsnetzwerkes DeepLabv3+ möglich?
- 10) Wie leistungsfähig ist ein auf Mobile Laserscanning und Deep Learning basierender, automatischer Inspektionsprozess im Vergleich zu einer konventionellen Inspektion durch menschliche Auswerter im kommunalen Straßennetz?

1.2 Aufbau der Arbeit

In diesem Kapitel wird der Aufbau der vorliegenden Arbeit beschrieben, um einen Überblick über die inhaltliche Ausrichtung der sechs Kapitel zu ermöglichen.

Nachdem in einem ersten Kapitel die Motivation und Zielsetzung dieser Arbeit dargestellt werden, folgt das Kapitel zu den theoretischen Grundlagen.

Im zweiten Kapitel werden alle Aspekte beleuchtet, die zum Verständnis der praktischen Untersuchungen dieser Arbeit nötig sind. Dabei wird zunächst der Themenkomplex der systematischen Straßenerhaltung erläutert. Der Fokus liegt hierbei auf den Richtlinien und Regelwerken, die für das deutsche Straßennetz maßgeblich sind. Es wird beschrieben, wie die Zustandserfassung und -bewertung sowohl im Bundesfernstraßennetz als auch im kommunalen Straßennetz abläuft und welche Faktoren die Qualität einer manuellen Erfassung von Straßenschäden beeinflussen. Anschließend folgt ein Überblick über die Kerntechnologien des mobilen Laserscannings, wobei schwerpunktmäßig die Themenkomplexe LiDAR und satellitengestützte Positionierung erläutert werden. Danach werden die räumliche Oberflächenmodellierung in Geo-Informationssystemen sowie die Geomorphometrie als Teildisziplin der physischen Geographie vorgestellt. Im Anschluss daran folgt die schrittweise Einführung von Deep Learning als wesentliche Datenanalysetechnologie der vorliegenden Arbeit. Zunächst wird Deep Learning in das Themenfeld der künstlichen Intelligenz und des maschinellen Lernens eingeordnet. Anschließend werden die grundlegenden Strukturen und Prozesse innerhalb künstlicher neuronaler Netze erläutert. Ein Block widmet sich jenen Aspekten, die für das Problem der semantischen Bildsegmentierung mithilfe tiefer neuronaler Faltungsnetze von Bedeutung sind. Am Ende des zweiten Kapitels wird aufgezeigt, welche Entwicklungen in Richtung Automatisierung auf nationaler, wie auch internationaler Ebene im Themenfeld der Straßenzustandserfassung bisher stattfanden und die Rolle der Untersuchungen der vorliegenden Arbeit dahingehend eingeordnet.

Das dritte Kapitel soll ein Verständnis für die in der Arbeit verwendeten Daten schaffen. Im ersten Teil liegt der Fokus auf der Datenerfassung. Hierzu wird das Mobile Laserscanning System I.R.I.S inklusive aller relevanten Komponenten sowie sonstige Vermessungsinstrumente und Systeme vorgestellt, die zur Aufnahme von Messdaten für die praktischen Untersuchungen der Arbeit dienen. Im Anschluss folgt eine Beschreibung aller im Rahmen der vorliegenden Arbeit erfassten und analysierten Datensätze. Zur Bearbeitung und Beantwortung der Forschungsfragen (siehe Kapitel 1.1) wurden insgesamt sechs verschiedene Datensätze aufgenommen und aufbereitet.

Im vierten Kapitel werden die Methoden, Werkzeuge und Verfahren erläutert, mit deren Hilfe die praktischen Untersuchungen der Arbeit umgesetzt wurden. Zunächst wird die Arbeitsumgebung im Hinblick auf die verwendete Hard- und Software beschrieben. Anschließend widmet sich ein Unterkapitel der Erzeugung von Oberflächenmodellen aus den rohen Laserscans mittels räumlicher Interpolation. Hierbei werden ausgewählte deterministische Interpolationsverfahren im Detail erläutert. An die Darstellung der Erzeugung von Oberflächenmodellen schließt sich ein Themenblock zur digitalen Analyse von Oberflächenmodellen mittels geomorphometrischer Algorithmen an. Hier werden Ansätze vorgestellt, die es ermöglichen, relevante Strukturen für die Schadenserkennung im Oberflächenmodell zu extrahieren und zu visualisieren, sodass die im Höhenmodell inhärente Information sowohl von der menschlichen Intelligenz (erfassende Person) wie auch von der künstlichen Intelligenz (Deep-Learning-Ansatz) interpretiert werden kann. Im Kontext dessen wird auch auf die Analogie zwischen Formen des Georeliefs und Fahrbahnschäden als Elemente des Piko- und Nanoreliefs eingegangen. Danach folgt die Erläuterung des im Rahmen dieser Arbeit verwendeten Deep-Learning-Ansatzes. Dabei wird die Entwicklung des DeepLab-Systems nachvollzogen und wesentliche Innovationen und Merkmale der DeepLabV3+ Architektur beschrieben. Im letzten Teil des Methodikkapitels wird dargelegt, wie die Untersuchungen der vorliegenden Arbeit konkret evaluiert werden.

Im fünften Kapitel werden die praktischen Untersuchungen beschrieben und deren Ergebnisse in Bezug auf die zentralen Fragestellungen der Arbeit vorgestellt. Dabei werden zunächst die absolute Positionierungsgenauigkeit des Mobile Laserscanning Systems I.R.I.S. und die Wiederholungsgenauigkeit des Messsystems anhand eines Prüfkörpers evaluiert. Anschließend folgen die Untersuchungen zur Oberflächenmodellierung. Hierbei wird zunächst ein für das vorliegende Problem und Datenmaterial optimiertes Interpolationsverfahren zur Erzeugung von Oberflächenmodellen identifiziert. Danach wird anhand eines zweiten Prüfkörpers evaluiert, bei welchen Aufnahmeszenarien (Geschwindigkeit und Ausrichtung) welche Spaltbreiten im Oberflächenmodell des gescannten Prüfkörpers visuell erkennbar sind.

Im Anschluss daran folgen die Untersuchungen im Themenfeld Deep Learning. Dabei wird zunächst ein umfangreicher Trainings- und Testdatensatz aufgebaut, der in reduzierter Form dazu genutzt wird, im Rahmen von Voruntersuchungen eine optimale Trainings-Pipeline zu entwickeln und ein geeignetes Hyperparameter-Setting für die Trainingsprozesse zu finden. Anschließend werden die finalen DeepLabV3+ Modelle auf Basis des gesamten Datensatzes mit den gewonnenen Erkenntnissen trainiert und evaluiert. Diese Evaluation geschieht auf Pixel-Ebene unter Nutzung eines Teils des Datensatzes, der nicht im Training genutzt wurde.

Danach wird erläutert, wie die Ergebnisse der semantischen Bildsegmentierung, die durch das Deep-Learning-Modell erzeugt werden, vom Bildkoordinatensystem wieder zurück in das globale Koordinatensystem transformiert werden. Am Ende dieses Prozesses entsteht eine automatische Kartierung der Fahrbahnoberfläche (Schäden, Objekte und Deckschichtarten), bei der die Analyseergebnisse als Geoobjekte für weiterführende räumliche Analysen zur Verfügung stehen. Um das entwickelte Verfahren dahingehend zu evaluieren, ob die Analysequalität den Anforderungen der maßgeblichen Regelwerke entspricht, findet ein Vergleich mit einer manuellen Referenzauswertung und einem internationalen Referenzsystem im Bereich der automatischen Erfassung von Straßenschäden statt.

Das letzte Kapitel der Arbeit fasst die Ergebnisse der Untersuchungen vor dem Hintergrund der Forschungsfragen zusammen und gibt einen Ausblick auf weitergehende Forschungsfragen.

2 Einführung in Themen und Technologien

In den folgenden Unterkapiteln werden die zum Verständnis der vorliegenden Arbeit grundlegenden Themenfelder und Technologien aus den Bereichen Straßenerhaltung, Mobiles Laserscanning, Geo-Informationssysteme sowie künstliche Intelligenz vorgestellt.

2.1 Systematische Straßenerhaltung

Das systematische Erhaltungsmanagement von Straßen bildet den thematischen Rahmen der Untersuchungen der vorliegenden Arbeit und wird daher in den nachfolgenden Unterkapiteln erläutert. Eine detaillierte Auseinandersetzung mit den Grundlagen des Erhaltungsmanagements ist u. a. bei SCHMUCK (1987) und BLUMENFELD (2020) zu finden.

Unter dem Begriff Straßenerhaltung werden alle Maßnahmen zur Kontrolle, Wartung, Instandhaltung, Instandsetzung und Erneuerung von Straßenverkehrsflächen subsummiert (BLEßMANN et al. 2010). Die Ziele der Straßenerhaltung adressieren die Aspekte Sicherheit, Verfügbarkeit, Wirtschaftlichkeit und Nachhaltigkeit. So soll die Straßenerhaltung zuvorderst eine sichere Nutzung der Straßeninfrastruktur gewährleisten. Dabei sind die Erhaltungsmaßnahmen so zu planen, dass die Leistungsfähigkeit der Straßeninfrastruktur sichergestellt bleibt. Beeinträchtigungen des Verkehrs sind auf ein notwendiges Maß zu beschränken. Ferner sind die Aspekte Wirtschaftlichkeit und Umwelt wesentliche Zielkriterien, da die Straßenerhaltung einerseits das Anlagevermögen der Straßeninfrastruktur unter wirtschaftlichem Mitteleinsatz sichern und andererseits natürliche Ressourcen im Zuge von Erhaltungsmaßnahmen schonen soll. Als operatives Instrumentarium unterstützt eine systematische Straßenerhaltung die jeweiligen Straßenbaulastträger dabei, Erhaltungsmaßnahmen auf Projektebene nach einheitlichen Verfahrensabläufen und Bewertungsalgorithmen zu planen und strategische Zielvorgaben in Form von Qualitäts- oder Finanzszenarien auf Straßennetzebene umzusetzen. Auch wenn die Begriffe Straßenerhaltung und Erhaltungsmanagement oft synonym verwendet werden, beschreibt Letzteres eher eine spezielle, koordinierte Abfolge von Planungsaufgaben und Aktivitäten (BLUMENFELD 2020). Basierend auf der Grundidee des prozessorientierten Qualitätsmanagements, ist es die Aufgabe des Erhaltungsmanagements, zu gewährleisten, dass notwendige Entscheidungen der Straßenerhaltung rational und systematisch auf Basis zielbezogener Kriterien getroffen und optimale Ergebnisse erzielt werden können (SCHMUCK 1987). In Abbildung 1 ist der Prozesskreislauf des Erhaltungsmanagements dargestellt. Neben politischen Zielvorgaben bilden möglichst objektiv und

systematisch erfasste Daten die Grundlage für Entscheidungsprozesse sowohl auf operativer wie auch strategischer Ebene.

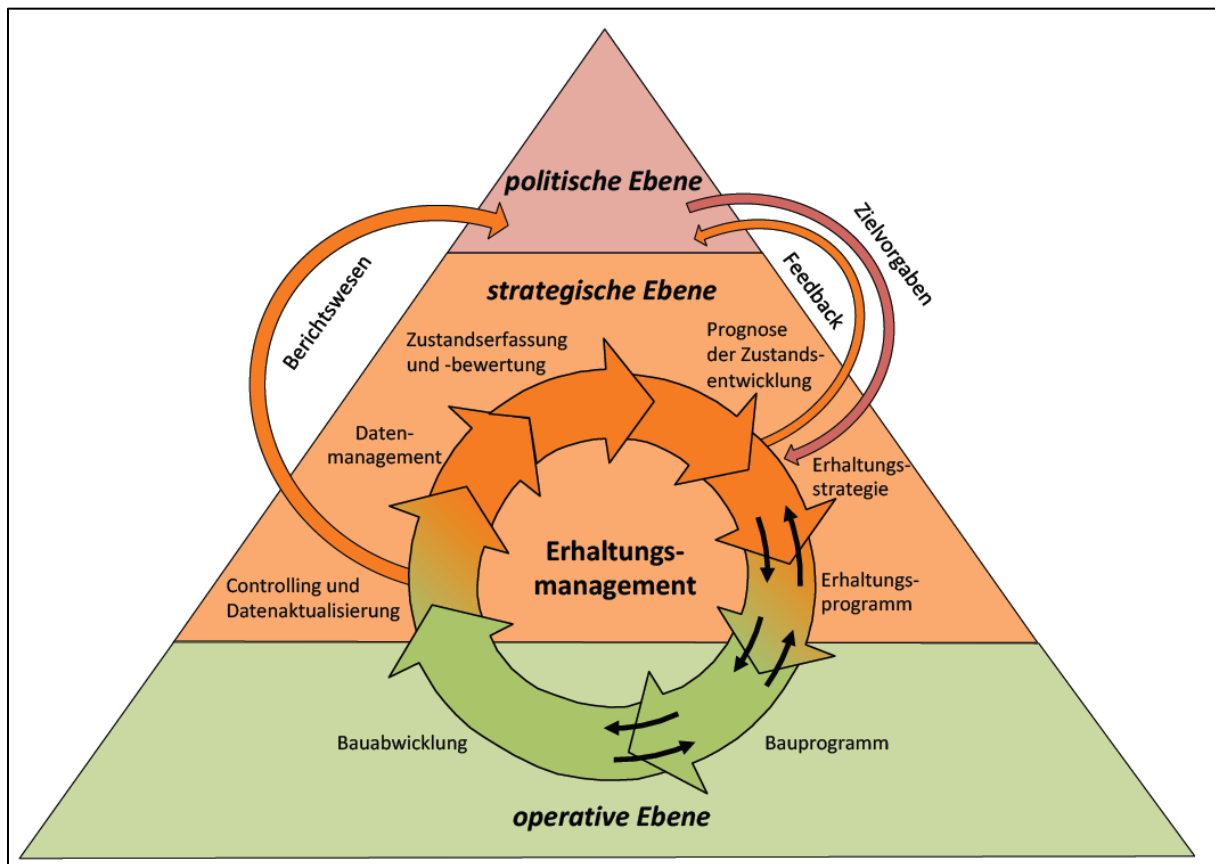


Abbildung 1: Ablaufschema des systematischen Erhaltungsmanagements (BLUMENFELD 2020).

Im Kontext der Straßenerhaltung finden verschiedene Daten Berücksichtigung. So fließen beispielsweise Informationen zum Aufbau der Straßenbefestigung (Dicke, Material, Alter, Eigenschaften) und zur Verkehrsbelastung ein, da diese Aspekte die Zustandsentwicklung maßgeblich beeinflussen. Zusammen mit Netz-, Verkehrs- und Aufbaudaten bilden Daten zum tatsächlichen Straßenzustand die wesentlichen Eingangsgrößen für eine systematische Straßenerhaltung innerhalb eines ganzheitlichen Erhaltungsmanagements. Der Prozess der Zustandserfassung und -bewertung liefert dabei die wichtigste Datenbasis (BLUMENFELD 2020). Daher wird in den folgenden Unterkapiteln näher auf die Systematik der Straßenzustands-erfassung eingegangen.

2.1.1 Organisationen

Zunächst seien an dieser Stelle zwei Organisationen bzw. Gremien erwähnt, die für die Straßenerhaltung in Deutschland eine wichtige Rolle spielen: die Bundesanstalt für Straßenwesen (BASt) und die Forschungsgesellschaft für Straßen- und Verkehrswesen e. V. (FGSV). Die BASt ist eine Forschungseinrichtung im Ressort des Bundesverkehrsministeriums, die den Auftrag hat, Sicherheit, Umweltverträglichkeit, Wirtschaftlichkeit und Leistungsfähigkeit des deutschen Straßennetzes zu verbessern. Im Bereich der Erhaltung beteiligt sie sich an der Entwicklung und Optimierung technischer Mess- und Aufnahmeverfahren sowie der Zulassung von Messsystemen zur Straßenzustandserfassung. Die FGSV ist ein Kompetenznetzwerk für Forschung und Wissenstransfer und verantwortet die Erstellung aller technischen Regelwerke für das gesamte Straßen- und Verkehrswesen in Deutschland.

2.1.2 Bundesfernstraßennetz

Im Hinblick auf die Ziele der Straßenerhaltung ist eine lückenlose Erfassung des Straßenzustandes nötig. Hierzu wurde vom Bundesministerium für Verkehr und digitale Infrastruktur (BMVI) zusammen mit den Bundesländern für das deutsche Fernstraßennetz ein genormtes Verfahren eingerichtet – die sogenannte Zustandserfassung und -bewertung von Straßen (ZEB) (BASt 2018a; STÖCKERT, SCHMERBECK & LAU 2019). Seit 1991 werden in einem vierjährigen Turnus im Zuge der ZEB Kriterien erfasst und bewertet, die den baulichen Zustand und Sicherheitsaspekte des Bundesfernstraßennetzes charakterisieren. Das maßgebliche Regelwerk hierfür ist die ZTV ZEB-StB (FGSV 2006). Oberstes Ziel der ZEB ist es, qualitätsgesicherte Daten für das Erhaltungsmanagement der Bundesautobahnen und Bundesstraßen bereitzustellen.

Die messtechnische Zustandserfassung erfolgt im deutschen Straßennetz mit schnellfahrenden Erfassungssystemen (siehe Abbildung 2). Mithilfe unterschiedlicher Sensoren werden dabei straßenzustandsrelevante Daten erfasst, deren Einzelmessungen auf 10 m genau georeferenziert werden müssen. Aus den Messdaten werden zunächst Indikatoren ermittelt – die sogenannten Zustandsgrößen. Dies können beispielsweise die Tiefe einer Spurrinne in Millimeter oder die Anzahl von beschädigten Ecken einer Betonplatte sein. Mittels Normierungsfunktionen werden die Zustandsgrößen in dimensionslose Zustandswerte überführt. Die Notenskala der Zustandswerte reicht dabei von 1 bis 5 (sehr gut bis sehr schlecht).

Die ZEB ist in fünf sogenannte Teilprojekte (TP) gegliedert: Vorbereitung der Straßennetzdaten (TP0), Ebenheit (TP1), Griffigkeit (TP2), Substanzmerkmale (TP3) sowie Bewertung und

standardisierte Auswertung (TP4). Entsprechend dieser thematischen Unterteilung werden in den folgenden Unterkapiteln zunächst die Zustandsmerkmale Ebenheit, Griffigkeit und Oberflächenschäden erläutert. Anschließend wird auf die Bewertung eingegangen.



Abbildung 2: Beispiele für schnellfahrende Multisensorsysteme, die im Rahmen der Straßenzustandserfassung eingesetzt werden. Links: Messfahrzeug MEFA der BAST (I aus SKAKUJ et al. 2019). Rechts: Messfahrzeug ARGUS der Schniering GmbH (II aus GROßMANN et al. 2023).

2.1.2.1 Ebenheit

Zu den Funktionen einer Fahrbahnoberfläche zählt nicht zuletzt das Auffangen von Kräften, die bei der Überrollung durch Fahrzeuge in die Straßenkonstruktion eingetragen werden. Da Unebenheiten in der Fahrbahnoberfläche das sie überrollende Fahrzeug in Schwingung versetzen, können sie den Lasteintrag in die Straße in Form dynamischer Radlasten verstärken, was wiederum zu einer beschleunigten Schädigung der Fahrbahnoberfläche führt (UECKERMANN & OESER 2015). Deshalb ist die Ebenheit einer Straße im Längs- und Querprofil ein entscheidendes Qualitätsmerkmal im Straßenbau. Die Ursachen der Entstehung von Unebenheiten sind vielfältig und oftmals Resultat einer Kombination von mehreren Faktoren (KRAUSE 2000). Teilweise können Ebenheitsmängel aus dem Fertigungsprozess der Fahrbahndecke resultieren. Andere Unebenheiten entstehen aufgrund von übermäßiger Druckbelastung, verminderter Tragfähigkeit oder sind thermisch induziert.

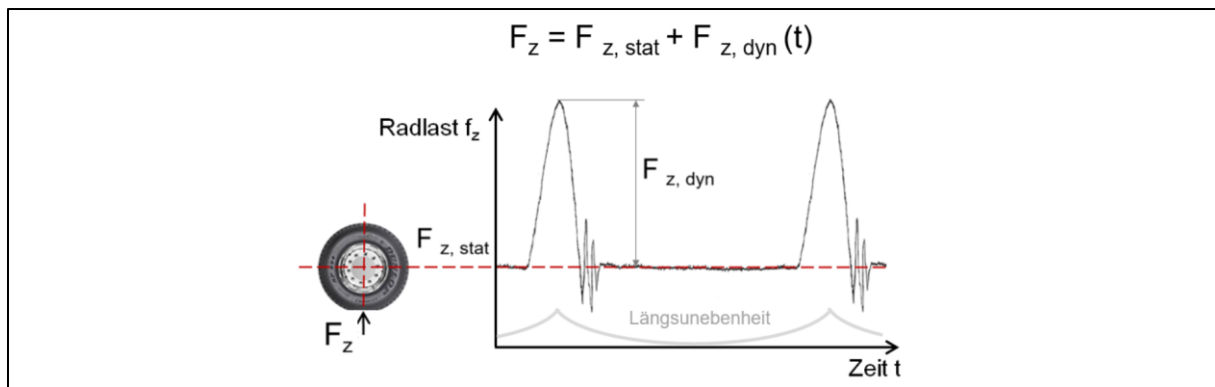


Abbildung 3: Schematische Darstellung der statischen Radlast $F_{z, \text{stat}}$ und dynamischen Radlast $F_{z, \text{dyn}}$ bei Überrollung einer Unebenheit in Fahrtrichtung. Durch die Unebenheit der Fahrbahnoberfläche erfährt das überrollende Fahrzeug eine Anregung und in der Folge erhöht sich die Kraft, die in die Fahrbahn eingetragen wird, um den Anteil der dynamischen Radlast (WIELAND & SESSELMANN 2018).

Während Unebenheiten im Längsprofil einer Straße vor allem den Fahrkomfort beeinflussen und aufgrund der angesprochenen dynamischen Radlasten zu einer stärkeren Beanspruchung der Bausubstanz führen (siehe Abbildung 3), wirken sich Unebenheiten im Querprofil einer Straße vor allem auf die Fahrsicherheit aus, da sie beispielsweise im Fall von Spurrinnen das Unfallrisiko bei Nässe erhöhen können (siehe Abbildung 4). Im Zuge der Zustandsbewertung liegt daher ein besonderer Fokus auf der genauen messtechnischen Erfassung der Fahrbahnoberflächengeometrie in Längs- und Querrichtung.

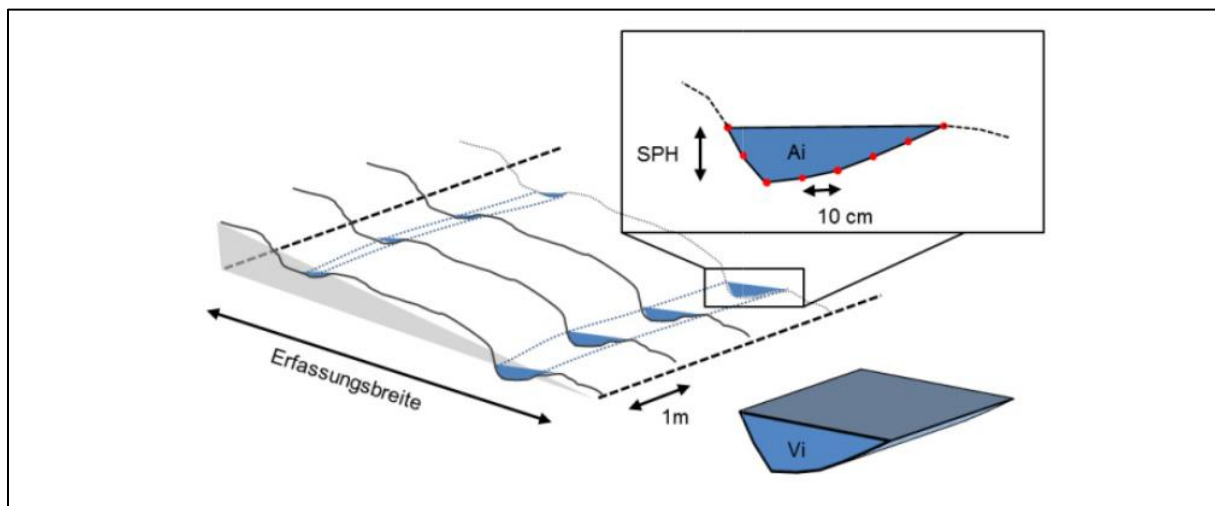


Abbildung 4: Schematische Darstellung des Wasserrückhalts in Spurrinnen (BALCK et al. 2017).

Das für die berührungslose Erfassung und Bewertung der Ebenheit maßgebliche technische Regelwerk TP Eben (FGSV 2009a) definiert die Erfassungstechnologien, die Messdurchführung und Auswerteverfahren. Die zur Erfassung eingesetzten Messfahrzeuge müssen demnach mit einer geeigneten Sensorik ausgestattet sein. Das Regelwerk sieht hierfür

prinzipiell Laserabstandssensoren vor, die nach dem Triangulationsprinzip arbeiten. Sofern eine Gleichwertigkeit mit den im technischen Regelwerk beschriebenen Anforderungen durch die BASt nachgewiesen ist, können auch abweichende Messprinzipien, wie beispielsweise Laserscanning, zum Einsatz kommen. Wie in Abbildung 5 links dargestellt, werden zur Erfassung des Längsprofils vier Laserabstandssensoren in einer bestimmten Konfiguration auf einem starren, längs zur Fahrtrichtung am Messfahrzeug montierten Balken, eingesetzt. Nach dem Prinzip der Mehrfachabtastung kann so ein Höhenlängsprofil entlang der Fahrlinie des Messfahrzeuges erfasst werden. Diese sogenannte High-Speed Road Monitoring (HRM) Methode nach STILL & JORDAN (1980) bildet derzeit das Standardverfahren in Deutschland, um Höhenlängsprofile einer Straße für eine Längsebenheitsanalyse zu generieren. Eine detaillierte Beschreibung des Verfahrens ist auch bei NEUBECK & WIESEBROCK (2015) zu finden. Zur Erfassung der Querebenheit wird, wie in Abbildung 5 rechts dargestellt, ein Quermessbalken eingesetzt, der typischerweise vorne am Messfahrzeug verbaut wird (siehe Abbildung 2 rechts). Aus dem geforderten Messpunktabstand von 0,1 m und den Mindestfahrstreifenbreiten auf Autobahnen von 3,2 m ergibt sich eine Mindestzahl von 32 Abstandssensoren, die auf einem Quermessbalken zu integrieren sind. Sowohl für die Ableitung der Längs- als auch der Querebenheitsindikatoren (Zustandsgrößen) werden Höhenprofile der Straßenoberfläche in einem 0,1 m Intervall zugrunde gelegt.

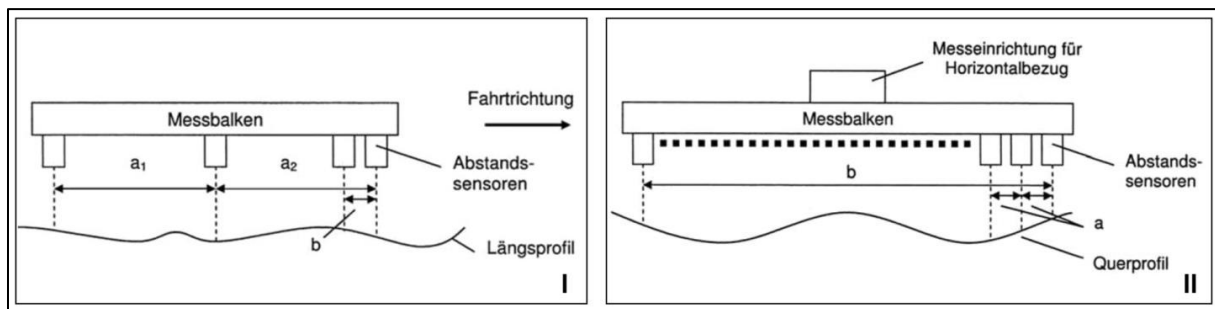


Abbildung 5: Messprinzip der Ebenheit im Längsprofil (I) und Querprofil (II) auf Basis von Punktlasersensoren (FGSV 2009a).

Zur Analyse der Längsebenheit auf Basis eines berührungslos gemessenen Höhenlängsprofils einer Straße sind verschiedene Kennwerte vorgesehen. Darunter sind geometrische Indikatoren, aber auch komplexere Kennwerte, die Analysen im Frequenzraum erfordern. Zur ersten Kategorie zählen die Standardabweichung der Differenz zum gleitenden Mittelwert über verschiedene Mittelungslängen von 3 m bis 30 m (S03, S10 und S30) und die Simulation der gleitenden 4 m Richtlatte (PGR). In die zweite Gruppe fallen der Längsebenheitswirkindex (LWI), die Allgemeine Unebenheit, gekennzeichnet durch das Unebenheitsmaß (AUN) und die

Welligkeit (W), sowie das Bewertete Längsprofil (BLP). Langwellige Unebenheiten werden mindestens bis zu einer Wellenlänge von 50 m erfasst, um auch fahrdynamische Auswirkungen auf die Straßennutzer zu berücksichtigen. Hinsichtlich der Querebenheit sind als Indikatoren gemäß FGSV (2009a) die Spurrinntentiefe nach dem 2 m-Lattenprinzip (SPT), die fiktive Wassertiefe (SPH) sowie die Querneigung (QN) zu berechnen. Die konkreten Berechnungsvorschriften der einzelnen Indikatoren sind im Anhang des technischen Regelwerkes aufgeführt (FGSV 2009a).

2.1.2.2 Griffigkeit

Unter dem Begriff Griffigkeit wird im Kontext der Straßenzustandsbewertung eine Eigenschaft der Fahrbahnoberfläche verstanden, die im Zusammenhang mit ihrer Mikro- und Makrotextur steht (siehe Abbildung 6). Während die Ebenheit Wellenlängen im Zentimeter- und Meterbereich charakterisiert, adressiert die Griffigkeit eher Oberflächenunstetigkeiten mit Wellenlängen im Bereich Submillimeter und Millimeter. Griffigkeit ist definiert als ein Wirkindex, der die Wirkung der Rauheit auf den Reibungswiderstand zwischen Fahrzeugreifen und nasser Fahrbahn kennzeichnet. Dementsprechend ist es das Ziel von Griffigkeitsmessungen, die Wirkung der Textureigenschaften einer Fahrbahnoberfläche bei nasser Fahrbahn in Form des sogenannten Reibungsbeiwertes zu beschreiben (JACOT, LINDENMANN & SEILER 2007).

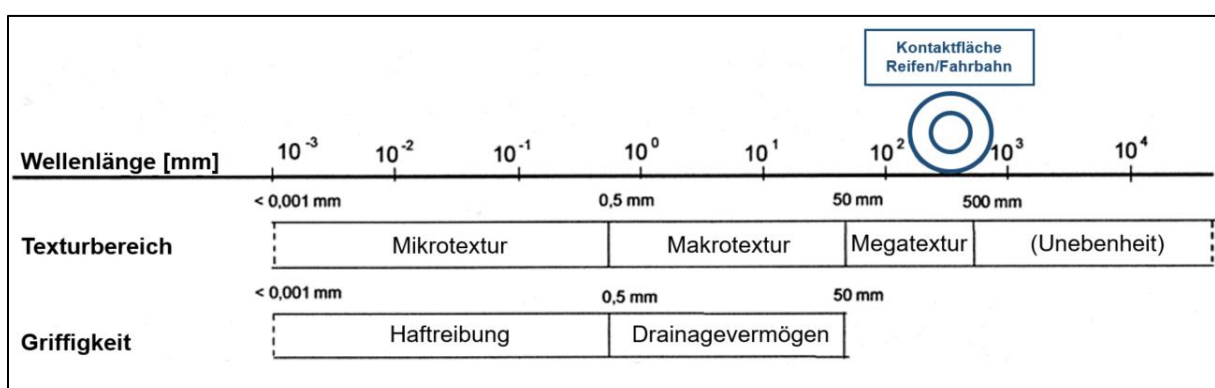


Abbildung 6: Untergliederung der Texturbereiche in Abhängigkeit zur Wellenlänge (nach FGSV 2009b).

Die Griffigkeit steht unter dem Einfluss diverser Faktoren. So wirkt die Witterung, wie beispielsweise Frost, in der Regel aufräufend auf eine Fahrbahnoberfläche ein. Dagegen kommt es infolge starker Verkehrsbeanspruchung eher zu einem Polieren der Oberfläche

(ROHLER et al. 2010). Die Erfassung von Griffigkeitsdaten einer Fahrbahnoberfläche adressiert im Kontext der Straßenzustandsbewertung vor allem fahrsicherheitsrelevante Aspekte. Geringe Griffigkeitswerte implizieren eine verringerte Haftung durch nutzungsbedingtes Polieren der Fahrbahnoberfläche und somit ein erhöhtes Unfallrisiko vor allem bei Nässe (MAYER, VAN DER SLUIS & STEINAUER 2002).

In Deutschland erfolgt die Erfassung der Griffigkeit gemäß technischem Regelwerk TP Griff-Stb (FGSV 2007) mit speziellen Messfahrzeugen nach dem Prinzip der Seitenkraftmessung mittels schräggestelltem Rad – kurz SKM. Wie in Abbildung 7 dargestellt, ist das Kernelement von SKM-Messfahrzeugen ein Rad mit einem von der BASt zertifizierten Reifen, welches in der rechten Rollspur in einem Winkel $\alpha = 20^\circ$ zur Fahrtrichtung schräg gestellt ist. Die bei der Vorwärtsbewegung V des Messfahrzeuges wirkende Seitenkraft F_y wird ins Verhältnis zur Normalkraft F_z gesetzt und als Seitenkraftbeiwert bezeichnet. Die Fahrbahnoberfläche wird während der Griffigkeitsmessung definiert angenässt, sodass ein möglichst gleichmäßiger Wasserfilm entsteht. Die Messung erfolgt bei einer definierten Geschwindigkeit, die auf Autobahnen 80 km/h, auf Landesstraßen 60 km/h und Innerorts 40 km/h beträgt. Zusätzlich werden die Temperaturen von Reifen, Wasser, Fahrbahn und Luft während der Messung aufgezeichnet, da sie den Seitenkraftbeiwert beeinflussen. Abschließend werden die kontinuierlich aufgezeichneten temperatur- und geschwindigkeitskorrigierten Griffigkeitswerte (GRI) zu Mittelwerten im 1 m-Intervall verrechnet.

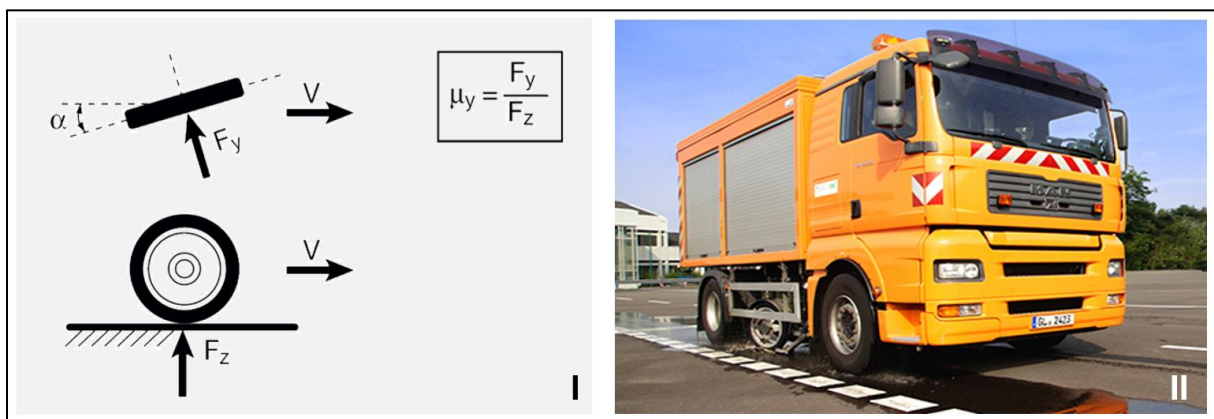


Abbildung 7: Links: Schematische Darstellung der Seitenkraftmessung mit schräggestelltem Rad (I aus FGSV 2007). Rechts: Beispiel für ein SKM Messsystem (II aus BASt 2022).

2.1.2.3 Substanzmerkmale

Unter dem Begriff Substanzmerkmale werden sichtbare Schäden an der Fahrbahnoberfläche verstanden. Ob und wenn ja wann sich Substanzmerkmale nach der Verkehrsfreigabe einer Fahrbahn ausbilden, hängt von verschiedenen Einflussfaktoren ab. Dazu zählen die Dimensionierung des Oberbaus einer Straße, Eigenschaften der verbauten Materialien, die Einbauqualität, klimatische Randbedingungen und nicht zuletzt die Verkehrsbelastung vor allem durch den Schwerlastverkehr (WACKER 2021). „Ohne die Durchführung einer geeigneten Erhaltungsmaßnahme nehmen das Ausmaß und die Schwere eines Oberflächenschadens zu und er kann sich zu einem anderen, eventuell schwerwiegenderen Schaden weiterentwickeln“ (WACKER 2021, S. 8).

Die Erfassung von Oberflächenschäden ist gemäß dem technischen Regelwerk ZTV ZEB-StB (FGSV 2006) auf Grundlage einer videotechnischen Aufzeichnung der Straßenoberfläche durchzuführen. Technische Details zur Bildaufnahme und Auswertetechnik sind in der Prüfvorschrift TP Oberflächenbild-StB (FGSV 2020) erläutert. Als zulässige Sensoren sind dort neben Zeilen- und Flächenkameras auch Laserscanner aufgeführt. Um bei passiver Sensorik eine homogene Beleuchtungssituation auf den Oberflächenbildern zu gewährleisten, werden die Erfassungssysteme mit entsprechenden Beleuchtungseinheiten (z. B. Stroboskoplampen) ausgestattet. Die aufgenommenen Bilddaten müssen eine Breite von 4,5 m abdecken und es muss gewährleistet sein, dass Risse ab einer Breite von 1 mm sicher erkannt werden können. Standardmäßig bilden Oberflächenbilder 10 m Straßenoberfläche in Fahrtrichtung ab. Dabei wird die reale Straßengeometrie nicht berücksichtigt, weshalb die Straße gerade verlaufend im Bildkoordinatensystem abgebildet wird. Bei sehr engen Kurvenradien können daher Verzerrungen im Bild entstehen.

Primär erfolgt die Substanzmerkmalserfassung auf Basis hochauflösender Oberflächenbilder. Als sekundäre Datenquelle können Frontkamerabilder unterstützend eingesetzt werden. Im Regelwerk ZTV ZEB-StB (FGSV 2006) wird von einer manuellen Bildauswertung durch explizit geschultes Personal ausgegangen. Abbildung 8 zeigt beispielhaft Ausschnitte aus Oberflächenbildern eines Zeilenkamerasystems.

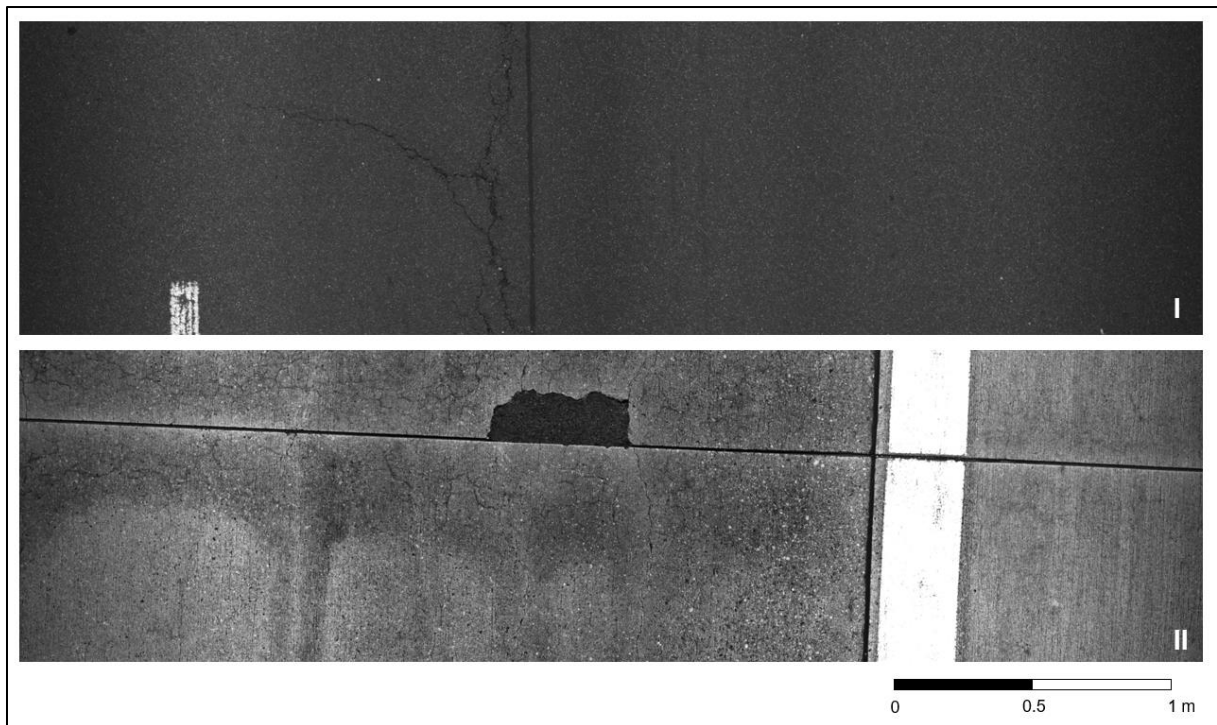


Abbildung 8: Ausschnitte aus Oberflächenbildern des BAST-zertifizierten Messfahrzeuges S.T.I.E.R der Firma Lehmann+Partner GmbH (LEHMANN+PARTNER GMBH 2022). Oben: Deutliche Risse auf einer Asphaltdecke (I). Unten: Feine Risse sowie ein Kantenschaden auf einer Betondecke (II).

Die zu erfassenden Schadensarten (Zustandsgrößen) unterscheiden sich nach der vorliegenden Bauweise der Fahrbahndecke in Asphalt- und Betonschäden. Für Asphaltbauweisen sind Risse (RISS), aufgelegte und eingesetzte Flickstellen (AFLI und EFLI), Ausbrüche (AUS), offene Arbeitsnähte (ONA) sowie Bindemittelanreicherungen (BIN) zu erfassen. Wie in Abbildung 9 links dargestellt, werden Asphaltbeschäden mithilfe eines in die Fahrstreifengrenzen eingepassten, dreispaltigen Rasters erfasst. Eine Rasterzelle ist dabei 1 m hoch, sodass das Raster für ein einzelnes Oberflächenbild 3 x 10 Zellen umfasst. Bis auf wenige Ausnahmen, bei denen ein sogenannter Flächenausgleich vorgenommen werden muss, ist eine Rasterzelle für den jeweiligen Schaden als betroffen zu kodieren, sobald sie diesen enthält. Die Zustandsgrößen ergeben sich aus den aktivierten Zellen des Rasters: sind beispielsweise 15 Zellen auf einem Oberflächenbild von Rissen betroffen, so entspricht dies einem Indikatorwert von 50 % für die Zustandsgröße RISS. Für ONA ist die Zustandsgröße als Länge der offenen Naht in Metern anzugeben, was sich aus der Anzahl der betroffenen Rasterzellen ableiten lässt. Vereinfachend wird angenommen, dass eine mit dem Merkmal ONA aktivierte Zelle 1 m offener Arbeitsnaht entspricht. Für Betonfahrbahnen werden Schäden grundsätzlich in einer anderen Art und Weise erfasst. Wie in Abbildung 9 rechts dargestellt, werden Betonschäden plattenbezogen aufgenommen. Anders als bei Asphaltbauweisen bilden die Indikatoren hier vornehmlich Anzahl und Länge der Schäden sowie den Anteil der betroffenen Platten ab. Die

relevanten Schadensarten sind Längs- und Querrisse (LQR), Eckabbrüche (EAB), Kantenschäden (KAS), Abplatzungen (NTR) und bituminöser Teilersatz (BTE). Auf eine ausführliche Beschreibung der Schadensbilder und Auswerteregeln sowie der vielfältigen Ausnahmeregelungen wird an dieser Stelle verzichtet und auf den Anhang des technischen Regelwerkes (FGSV 2006) verwiesen.

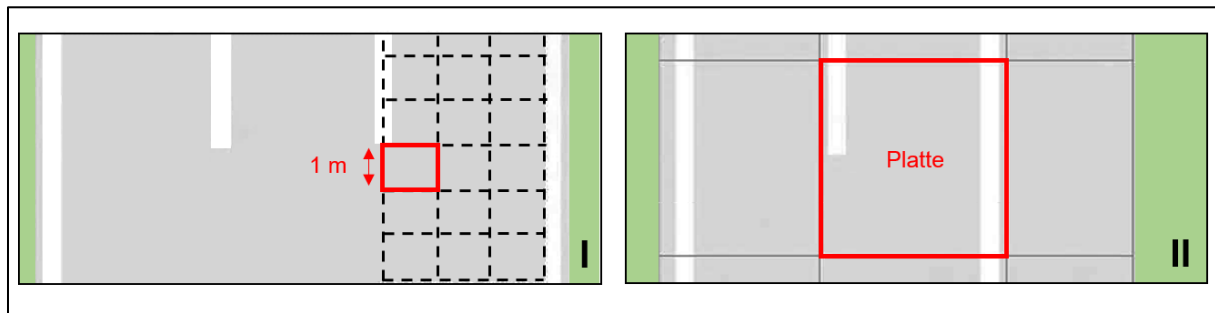


Abbildung 9: Prinzipskizze zur Veranschaulichung der Erfassungssystematik für Asphaltfahrbahnen (I) und Betonfahrbahnen (II). Asphalt Schäden müssen einer Rasterzelle von 1 m Länge und 1/3 der Fahrstreifenbreite zugeordnet werden. Die Fahrstreifen sind in der Abbildung durch Markierungslinien begrenzt. Betonschäden werden plattenscharf erfasst. Eine Platte ist dabei durch den Verlauf der Längs- und Querfugen begrenzt (nach FGSV 2006).

2.1.2.4 Zustandsbewertung

Inhalt des Teilprojektes TP4 ist die Bewertung der in den einzelnen Teilprojekten TP1 (Ebenheit), TP2 (Griffigkeit) und TP3 (Substanzmerkmale) erfassten Zustandsgrößen, die auf 100 m lange Auswerteabschnitte des Straßennetzes gemittelt wurden. Im Zuge des Bewertungsprozesses werden die Zustandsdaten dahingehend beurteilt, welche Wirkung sie auf die Erhaltungsziele Befahrbarkeit, Verkehrssicherheit, Substanzerhalt und Nachhaltigkeit haben (FGSV 2006). Aus diesem Grund werden die Zustandsgrößen mittels Normierung in ein einheitliches Notensystem überführt (Zustandswerte). Abbildung 10 zeigt den generalisierten Verlauf einer Normierungsfunktion zur Berechnung von Zustandswerten. Der Zustandswert 1,0 repräsentiert dabei den Idealzustand, der Zustandswert 3,5 den Warnwert und der Zustandswert 4,5 den Schwellenwert. Bei einer Zustandsnote zwischen 3,5 und 4,5 ist eine intensive Beobachtung nötig, um potentielle Schadensursachen zu identifizieren. Auch sollte bereits mit der Planung geeigneter Maßnahmen begonnen werden. Ab dem Schwellenwert (Note 4,5) ist die Einleitung baulicher oder verkehrsbeschränkender Maßnahmen angeraten. Die Fixpunkte, die in die Normierungsfunktion der jeweiligen Zustandsgröße eingesetzt werden müssen, werden auch als Anforderungsniveau bezeichnet. Die im Regelwerk ZTV ZEB-StB (FGSV 2006) definierten, von der Straßenkategorie abhängigen Anforderungsniveaus sind das

Ergebnis von Erfahrungswerten aus der Praxis und wurden zuletzt im Jahr 2015 grundlegend angepasst. Für die Zustandsgröße RISS ist der Warnwert beispielsweise bei 5 % Flächenanteil und der Schwellenwert bei 10 % Flächenanteil erreicht.

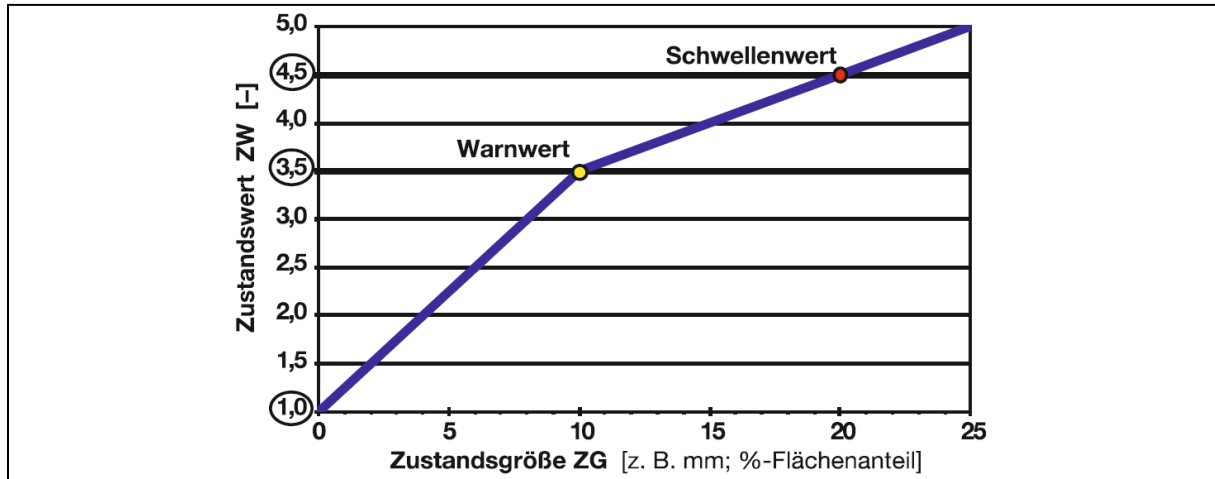


Abbildung 10: Prinzipskizze zur Überführung von Zustandsgrößen (ZG) in Zustandswerte (ZW). Die Normierungsfunktion ist definiert durch die Fixpunkte Warn- und Schwellenwert (FGSV 2006).

Anzumerken ist, dass nicht alle Zustandsgrößen individuell normiert werden. Beispielsweise sieht das technische Regelwerk ZTV ZEB-StB (FGSV 2006) für die Bewertung von Substanzmerkmalen auf Asphaltdecken zwei Normierungsfunktionen vor. Während der Flächenanteilindikator für Risse (RISS) explizit zum Zustandswert ZWRISS normiert wird, werden alle von anderen Schäden betroffenen Rasterzellen im 100 m langen Auswerteabschnitt, die ihrerseits nicht von Rissen betroffen sind, zur sogenannten Restschadensfläche (RSFA) zusammengefasst und gemeinsam zum Zustandswert ZWRSFA normiert. Auf Betonoberflächen wird analog dazu vorgegangen, nur bezieht sich der zu normierende Indikator nicht auf den prozentualen Flächenanteil, sondern den Anteil schadhafter Platten im Auswerteabschnitt. Hinsichtlich der Längsebene sind ebenfalls nicht alle Indikatoren bewertungsrelevant. So wird gemäß dem technischen Regelwerk derzeit nur das Unebenheitsmaß (AUN) in einen entsprechenden Zustandswert ZWAUN überführt.

Auf die Normierung folgt als letzter Schritt die sogenannte Wertesynthese. Dabei findet eine logische oder gewichtet-additive Verknüpfung der Zustandswerte zu den sogenannten Teilzielwerten (Gebrauchswert und Substanzwert) statt. Abschließend werden die beiden Teilzielwerte zum Gesamtwert verknüpft. Dabei ist das Maximum maßgeblich. In Abbildung 11 ist das Bewertungsschema bestehend aus Normierung und Wertesynthese schematisch dargestellt.

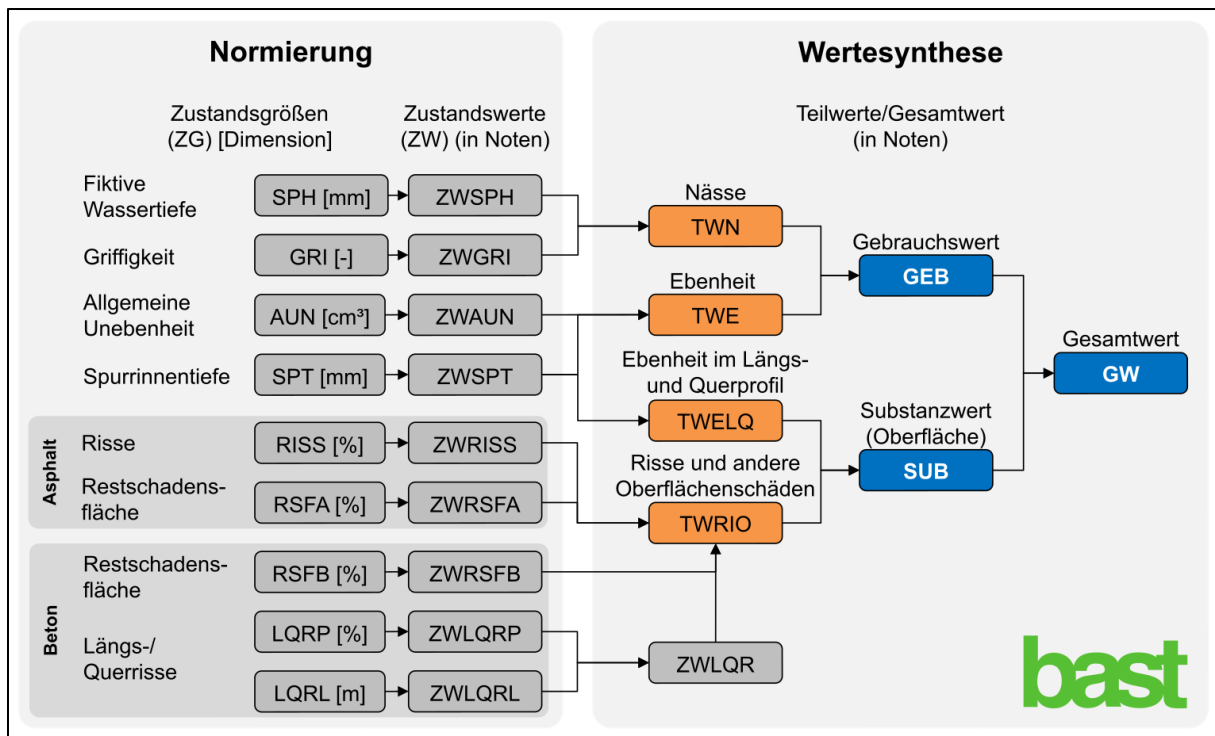


Abbildung 11: Bewertungsschema zur Ermittlung des Gesamtwertes eines Untersuchungsabschnittes (HELLER INGENIEURGESELLSCHAFT MBH 2022).

Das Bewertungsschema deutet bereits die herausgehobene Stellung von Rissen im Rahmen der deutschen Straßenerhaltung an. Risse sind „mit großem Abstand der häufigste Auslöser für die Erneuerung einer Asphaltdecke in Deutschland“ (STASCHKIEWICZ & RADENBERG 2022, S. 286). In knapp 89 % der Fälle sind Risse die maßgebende Schadensart, die einen Gesamtwert schlechter als 4,5 bedingen und damit eine unverzügliche Erhaltungsmaßnahme erforderlich machen. In etwa 4 % der Fälle waren andere Oberflächenschäden ausschlaggebend. Der Rest entfällt auf die Zustandswerte der Ebenheit und Griffigkeit. Daran wird deutlich, wie wichtig eine qualitativ hochwertige Erfassung von Oberflächenschäden im Kontext des Erhaltungsmanagements ist.

2.1.3 Kommunales Straßennetz

Auch zur Straßenerhaltung des sogenannten nachgeordneten Netzes (auch: kommunale Straßen oder Innerortsstraßen) wurde in Deutschland eine Systematik entwickelt, die im Gegensatz zum Bundesfernstraßennetz aber nicht verpflichtend ist, sondern eher den Charakter von Empfehlungen bzw. Richtlinien hat. Maßgeblich sind hier die Empfehlungen für das Erhaltungsmanagement von Innerortsstraßen (E EMI) sowie die zugehörigen Arbeitspapiere der Reihe K (FGSV 2012, 2016). Die übergeordneten Ziele einer systematischen kommunalen

Straßenerhaltung sind das Gewährleisten einer sicheren Befahrbarkeit des Straßennetzes und das Sicherstellen des Substanzerhalts. Die physische Beanspruchung von Straßennutzer, Fahrzeug und Ladung soll also möglichst gering sein und das Vermögen, welches in die Verkehrsflächen investiert wurde, soll wirtschaftlich erhalten werden. Der Erfüllungsgrad dieser Teilziele wird durch die sogenannten Teilzielwerte Gebrauchs- und Substanzwert repräsentiert, deren Ermittlung Gegenstand der kommunalen ZEB ist. Während der Gebrauchswert einer Straße hauptsächlich durch die Ebenheit in Längs- und Querrichtung charakterisiert wird, gehen in den Substanzwert in erster Linie die Oberflächenschäden einer Fahrbahn ein. Die Integration von Daten zur Griffigkeit sind im nachgeordneten Netz nicht vorgesehen.

Abhängig von der Struktur des Straßennetzes und den Zielsetzungen des jeweiligen kommunalen Baulastträgers kann eine visuelle oder eine messtechnische Zustandserfassung erfolgen. Auch Mischformen werden in der Praxis angewandt. Die visuelle Zustandserfassung ist im Arbeitspapier AP 9 K 2.2 beschrieben (FGSV 2016). Die Erfassung des baulichen Zustandes wird hierbei durch manuell-visuelle Inaugenscheinnahme im Rahmen von Begehungen und einer Dokumentation mittels analoger oder digitaler Formblätter realisiert (siehe Abbildung 12). Die Erfassung kann aber auch auf Basis von Bilddaten erfolgen, die beispielsweise mithilfe geeigneter mobiler Kamerasysteme aufgenommen wurden. Solche Erfassungssysteme müssen aber keinen messtechnischen Ansprüchen genügen. So ist es hierbei beispielsweise üblich, dass die Spurrinnentiefe als ein Indikator des Gebrauchswertes rein auf Basis des visuellen Erscheinungsbildes einer Oberfläche aus Smartphone-Kamerabildern geschätzt wird.

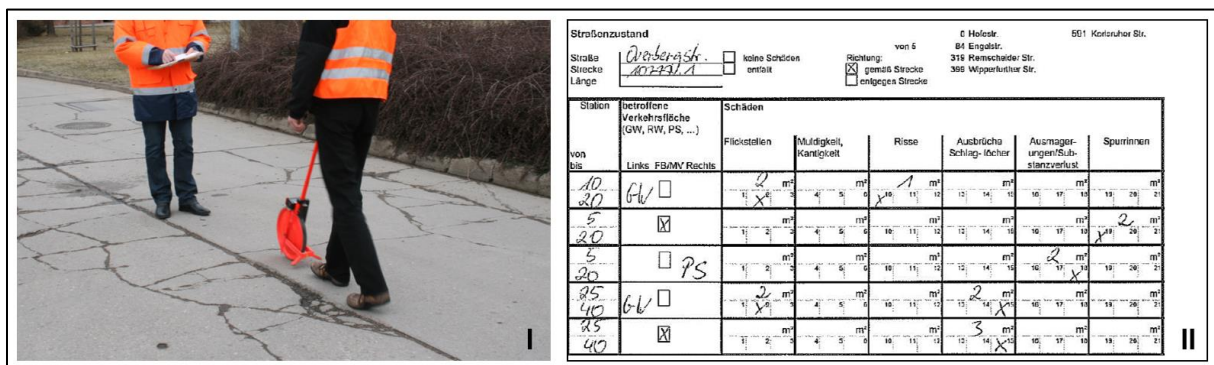


Abbildung 12: Visuelle Zustandserfassung durch Begehung (I) und Dokumentation mithilfe eines Erhebungsbogens (II nach FGSV 2016).

Im Unterschied dazu werden bei einer messtechnischen Zustandserfassung, die im Arbeitspapier AP 9 K 2.1 beschrieben ist, kinematische Multisensorsysteme eingesetzt, die in der Lage sind, mittels spezieller Sensorik (Kameras, Triangulationslaser, Laserlichtschnittsensoren, Laserscanner) Oberflächeneigenschaften einer Straße zu quantifizieren (FGSV 2016). Ziel der

messtechnischen Zustandserfassung ist es, Zustandsdaten im kommunalen Straßennetz objektiv und standardisiert zu erheben – in Anlehnung an die für das Bundesfernstraßennetz geltende Systematik (FGSV 2006). Während auf Autobahnen 100 m lange Auswerteabschnitte verwendet werden, betragen die Abschnittslängen im kommunalen Straßennetz typischerweise 10 m oder 20 m. Die Charakteristika dieser Art der Zustandserfassung, wie sie im kommunalen Straßennetz Anwendung findet, wird in den folgenden Unterkapiteln beschrieben.

2.1.3.1 Ebenheit

Während im Rahmen einer visuellen Zustandserfassung die Ebenheit einer Straße im nachgeordneten Netz anhand des Auftretens und der Ausprägung der Zustandsmerkmale Muldigkeit, Kantigkeit und Spurrinnen subjektiv beurteilt wird (schwach, deutlich, stark), richtet sich die messtechnische Zustandserfassung der Ebenheit nach der Systematik des Regelwerks ZTV ZEB-StB (FGSV 2006) und TP Eben (FGSV 2009a). Demzufolge werden berührungslos erfasste Längs- und Querprofile der Straßenoberfläche objektiv ausgewertet.

Gemäß der Richtlinie E EMI (FGSV 2012) wird die Ebenheit nur mithilfe geometrischer Indikatoren quantifiziert. Analysen im Spektralbereich sind nicht vorgesehen. Als Indikator für die Längsebenheit dient die Simulation der gleitenden 4 m Richtlatte. In der Literatur werden hierfür unterschiedliche Abkürzungen verwendet (PGR, L04_i, SM4L). Nachfolgend wird die Abkürzung SM4L genutzt. Eine virtuelle Latte von 4 m Länge wird schrittweise über ein digitales Höhenlängsprofil mit 0,1 m Punktabstand bewegt und an jeder Position der Abstand unter der Lattenmitte zum entsprechenden Profilpunkt berechnet (siehe Abbildung 13 links). Im Ergebnis liegt ein Signal der Stichmaße unter der 4 m Latte im 0,1 m Intervall vor. Darauf basierend wird für die Länge des Auswerteabschnitts ein mittlerer SM4L-Wert (SM4L-A) und ein maximaler SM4L-Wert (SM4L-M) berechnet und in Millimeter angegeben. Singuläre Unebenheiten im Wellenlängenbereich bis zu 4 m Länge können somit adäquat abgebildet werden. Ergänzend sind auf Basis des erfassten Höhenprofils die Standardabweichungen der Differenzen zum gleitenden Mittelwert bei unterschiedlichen Mittelungslängen zu berechnen (S01, S03, S10 und S30).

Die Querebenheit ist durch die Zustandsgrößen Spurrinntiefe und fiktive Wassertiefe zu charakterisieren. Querprofile müssen in Fahrtrichtung in einem gleichbleibenden Abstand von maximal 1 m erfasst und ausgewertet werden. Zur Berechnung der Spurrinntiefe wird eine virtuelle 2 m Latte schrittweise über die Höhenwerte eines Querprofils bewegt (siehe Abbildung 13 rechts oben). Getrennt nach linker und rechter Rollspur werden die Stichmaße unter der 2 m

Latte berechnet. Dafür wird das Querprofil in der Mitte geteilt. Das jeweils größte Stichmaß unter der Latte definiert die linke bzw. rechte Spurrinne und deren Tiefe. Die fiktive Wassertiefe gibt die theoretische Tiefe des Wassers an, welches bedingt durch die Spurrinne nicht abfließen kann und leitet sich dementsprechend direkt aus der Spurrinntiefen-Berechnung ab. Maßgeblich sind dabei die Auflagepunkte der 2 m Latte, die die Lage der Spurrinne im Querprofil definieren (siehe Abbildung 13 rechts unten). Auch hier wird je Querprofilhälfte ein Wert ermittelt. Auf der Grundlage dieser Einzelmesswerte werden für die Länge des Auswerteabschnittes Mittelwerte, Standardabweichungen und Maximalwerte berechnet: die Zustandsgröße MSPT beschreibt dann das Maximum der Mittelwerte der rechten und linken Spurrinntiefe und die Zustandsgröße MSPH das Maximum der Mittelwerte der rechten und linken fiktiven Wassertiefe jeweils in Millimeter.

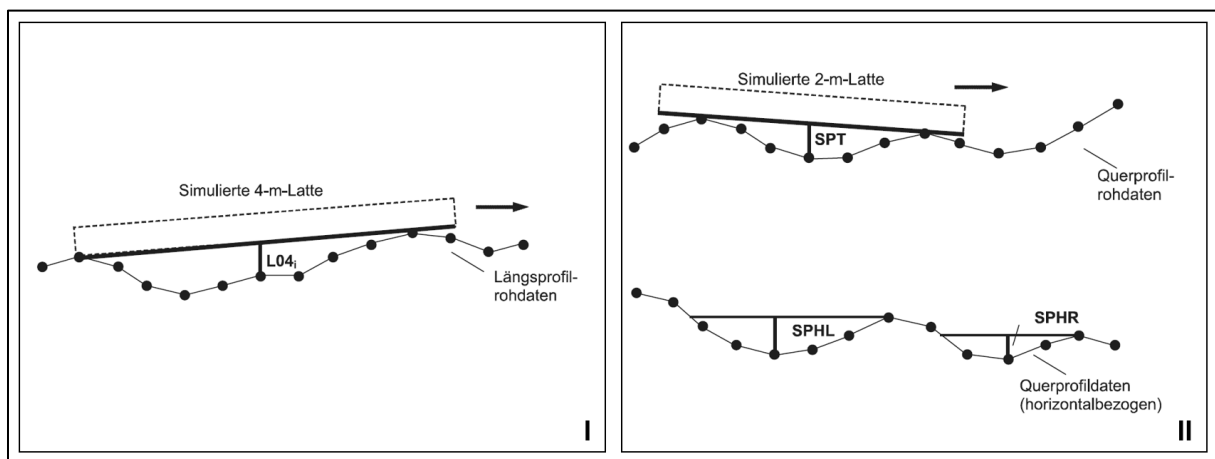


Abbildung 13: Prinzipdarstellung der für das kommunale Straßennetz bei messtechnischer Erfassung relevanten Ebenheitsindikatoren (FGSV 2006). Links (I): Simulation der gleitenden 4 m Richtlatte. Rechts oben (II): Spurrinntiefe. Rechts unten (II): fiktive Wassertiefe.

2.1.3.2 Substanzmerkmale

Im Arbeitspapier AP 9 K 2.1 (FGSV 2016) sind die zu erfassenden, an der Fahrbahnoberfläche erkennbaren strukturellen Schäden der Fahrbahnbefestigung beschrieben. Die relevanten Oberflächenschäden für Asphaltbauweisen sind demnach offene und sanierte Risse, aufgelegte und eingelegte Flickstellen, offene Nähte, Ausbrüche, Ausmagerungen und Bindemittelanreicherungen. Die relevanten Schäden für Pflasterbauweisen sind offene Pflasterfugen, schadhafte Pflasterelemente, Lagerungsschäden bzw. Störungen im Verband, Flickstellen und Gefügelösungen. Da Betondecken im kommunalen Bereich eher selten sind, wurden bislang keine expliziten Regelungen zur Analyse dieser Bauweise definiert (FGSV 2012, 2016). Im

Unterschied zu Bundesfernstraßen, sind die Plattengeometrien bei Betonbauweisen im kommunalen Netz sehr heterogen. Deshalb wird in der Praxis keine Platten-bezogene Schadenserfassung durchgeführt.

Oberflächenschäden werden im kommunalen Straßennetz aus Foto- oder Videodaten erfasst. Typischerweise werden dazu RGB-Kameras genutzt, die alle 1 m bis 5 m Bilder aus einer oder mehreren Perspektiven aufnehmen. In Abbildung 14 sind die zu erfassenden Schäden beispielhaft dargestellt.



Abbildung 14: Beispiele für die auf kommunalen Straßen relevanten Schadensarten Riss (I), offene Naht (II), Ausbruch (III), Ausmagerung (IV), eingelegte Flickstelle (V), aufgelegte Flickstelle (VI), Bindemittelanreicherung (VII), schadhaftes Pflasterelement (VIII), Flickstellen (IX), Lagerungsschaden (X), offene Pflasterfuge (XI) und Gefügelösungen (XII).

In Anlehnung an die Zustandserfassung auf Bundesfernstraßen werden Oberflächenschäden fahrstreifenweise über Rasterfelder von 1 m Länge und gedrittelter Fahrstreifenbreite erfasst. In der Praxis wird häufig ein 1 m² Raster für die Schadenserfassung zugrunde gelegt. Aus der Anzahl betroffener Rasterzellen kann dann pro Auswerteabschnitt die prozentuale Schädigung (Flächenanteil in %) ermittelt sowie die Länge in Meter approximiert werden. Tabelle 1 bietet

einen Überblick über die relevanten Schadensarten und die zugehörigen Zustandsgrößen sowie Indikatoren.

Tabelle 1: Übersicht über die relevanten Oberflächenschäden im kommunalen Straßennetz (nach FGSV 2016).

Zustandsgröße	Beschreibung	Zustandsindikator für Auswerteabschnitt
RISS	offene Risse, sanierte Risse, Einzelrisse, Netzerisse	Anteil der betroffenen Fläche [%]
ONF	offene Nähte, Nahtrisse	Länge [m]
AFLI	aufgelegte Flickstellen	Anteil der Schadensfläche [%]
EFLI	eingelegte Flickstellen	Anteil der Schadensfläche [%]
AUS	Ausbrüche, Abplatzungen, Schlaglöcher, flächenhafter Materialverlust	Anteil der Schadensfläche ohne RISS [%]
AMA	Ausmagerungen, Kornverluste	Anteil der Schadensfläche ohne RISS [%]
BIN	Austritt und Anreicherung von Bindemittel/Bitumen	Anteil der Schadensfläche ohne RISS [%]
OPF	offene Platten-/Pflasterfugen	Anteil der Schadensfläche [%]
BRU	schadhafte/fehlende Pflasterelemente	Anteil der Schadensfläche [%]
GVL	Störungen im Pflasterverband und Pflasterfugenverlauf, Lagerungsschäden	Anteil der Schadensfläche [%]
PFLI	Flickstellen auf Pflaster mit Ersatzmaterial	Anteil der Schadensfläche [%]
PAMA	Gefügebrauflösung auf Pflasterelementen	Anteil der Schadensfläche [%]

2.1.3.3 Zustandsbewertung

Je nach gewähltem Erfassungsansatz (messtechnisch oder visuell) wird ein entsprechendes Bewertungsverfahren empfohlen. Für die messtechnische Zustandserfassung ist dies in Arbeitspapier AP 9 K 3.1 definiert (FGSV 2016). Auf die Zustandsbewertung visuell erfasster Zustandsdaten, die im Arbeitspapier AP 9 K 3.2 beschrieben ist (FGSV 2016), wird hier nicht weiter eingegangen. Im Zuge der Bewertung werden die messtechnisch erfassten Zustandsgrößen (Merkmale bzw. Indikatoren) über Normierungsfunktionen in dimensionslose Zustandswerte (Notenwerte von 1 bis 5) überführt und zu den Teilzielwerten verknüpft, deren Maximalwert schließlich den Gesamtwert einer Straße ergibt (siehe Abbildung 15). Der Gebrauchswert wird demnach aus den normierten Zustandsgrößen Spurrinnentiefe (MSPT), fiktiver Wassertiefe (MSPH) sowie dem Stichmaß unter einer simulierten 4m-Latte gebildet (SM4L-M und SM4L-A). Der Substanzwert hingegen setzt sich aus den Zustandsgrößen Risse (RISS), aufgelegte Flickstellen (AFLI) und sonstige Oberflächenschäden (OFS) zusammen. Unter OFS werden Ausbrüche, Ausmagerungen und Bindemittelaustritte zusammengefasst. Offene Nähte sind nicht bewertungsrelevant. Analog zur Vorgehensweise auf Bundes-

fernstraßen bildet die Normierung und Wertesynthese den Kern der Zustandsbewertung und erlaubt nicht nur eine direkte Vergleichbarkeit der verschiedenen Merkmale, sondern die Funktionen sind auch so gewählt, dass sie die Messergebnisse fachlich beurteilen (FGSV 2012). Ein Notenwert von besser als 1,5 charakterisiert auch im kommunalen Straßennetz einen ordnungs-gemäßen Neubau. Ab einer Note von 3,5 ist eine intensive Beobachtung und Analyse der Schadensursache erforderlich. Bei Noten schlechter als 4,5 sollten unverzüglich bauliche Maßnahmen eingeleitet oder der Verkehr entsprechend eingeschränkt werden. Die unterschiedlich hohen Anforderungen und Belastungen von Hauptverkehrsstraßen und Anlieger- bzw. Wohnstraßen werden im Zuge der Bewertung insofern berücksichtigt, als dass je nach Funktionsklasse unterschiedliche Warn- und Schwellenwerte eingesetzt werden. In Abbildung 15 sind dementsprechend unterschiedliche Anforderungsniveaus für Funktionsklasse A und B mit drei in die Normierungsfunktion einzusetzenden Fixpunkt-Werten aufgeführt.

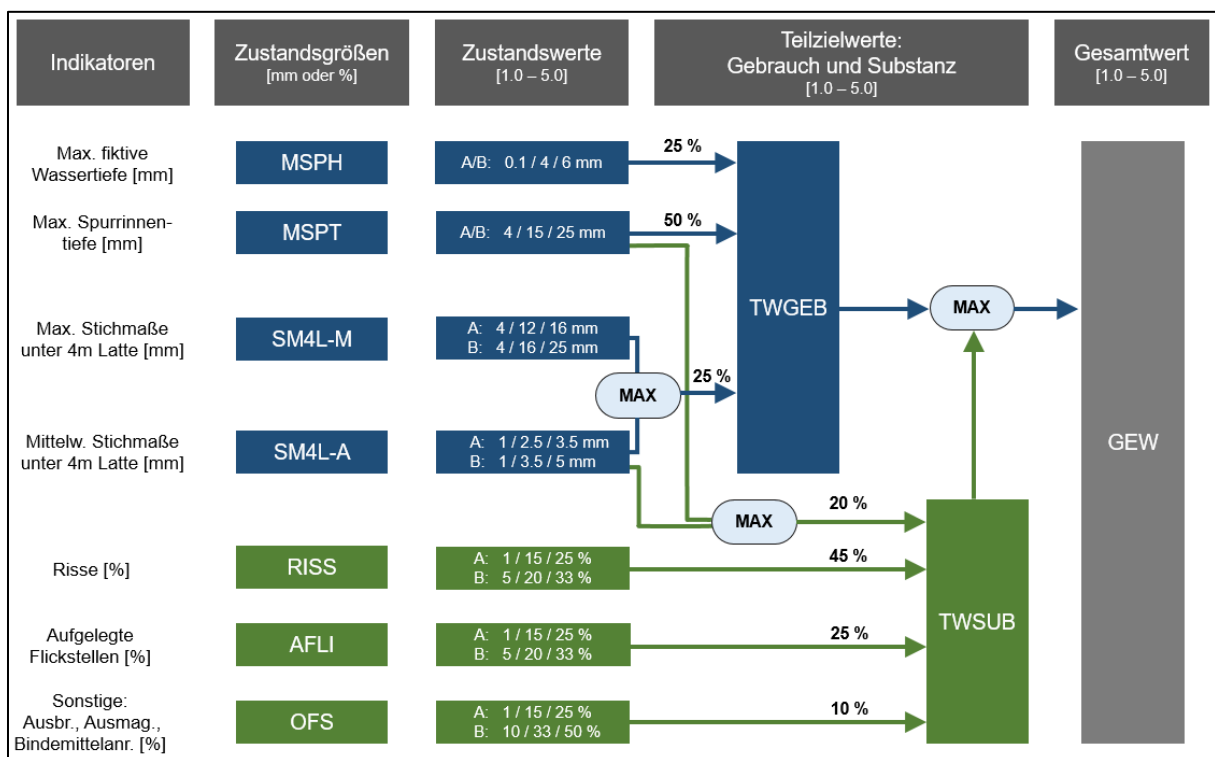


Abbildung 15: Schematische Darstellung der messtechnischen Zustandsbewertung kommunaler Straßen (nach FGSV 2016).

Am Ende einer messtechnischen Zustandserfassung und -bewertung steht typischerweise eine georeferenzierte Ergebnisdatei, die für Straßenabschnitte von 10 m oder 20 m Länge neben diversen Netz- und Metainformationen die Zustandsgrößen, Zustandswerte, Teilzielwerte sowie den Gesamtwert beinhaltet (FGSV 2016). Zum Zwecke einer einheitlichen grafischen Darstellung von Zustandswerten, Teilzielwerten oder des Gesamtwertes in Karten oder

Diagrammen sind die in Tabelle 2 aufgeführten Zustandsklassen sowie die ihnen zugeordneten Farben zu verwenden. Diese Daten können im Rahmen des kommunalen Erhaltungsmanagements strategisch oder operativ genutzt werden, um Erhaltungsabschnitte zu bilden, Maßnahmen zu priorisieren und Mitteleinsätze zu planen.

Tabelle 2: Zustandsklassen zur grafischen Darstellung in Karten und Diagrammen (nach FGSV 2012).

Zustandsklasse	Wertebereich (Zustandswert ZW)	Farbe
1	$ZW < 1,5$	Blau
2	$1,5 \leq ZW < 3,5$	Grün
3	$3,5 \leq ZW < 4,5$	Gelb
4	$4,5 \leq ZW < 5,0$	Rot

2.1.4 Zuverlässigkeit der manuellen Erfassung von Oberflächenschäden

Wie in den vorherigen Kapiteln dargelegt, geht die Straßenzustandserfassung und -bewertung in Deutschland bislang von einer manuellen Auswertung von Oberflächenschäden aus. Der Mensch als potentielle Fehlerquelle kann in diesem Prozess nicht quantifiziert werden, wie es beispielsweise für Messsensoren möglich ist. Einen allgemeinen Überblick über Techniken zur Bewertung der menschlichen Zuverlässigkeit bieten BOLT et al. (2010).

Für den Bereich Schienenverkehr ist die Fehlerrate menschlicher Handlungen Gegenstand expliziter Untersuchungen, da hier der Mensch trotz hoher Prozessautomation auf einer Rückfallebene, beispielsweise beim Ausfall kritischer Technik, eine sicherheitsrelevante Rolle spielt (MASCHEK 2018). Wie in Tabelle 3 dargestellt, können menschliche Handlungen gemäß RASMUSSEN (1983) in drei Kategorien unterteilt werden: fertigkeitsbasierte, regelbasierte und wissensbasierte Verhaltensweisen. HINZEN (1993) schlägt in Abhängigkeit dieser Kategorien und bestimmten Randbedingungen Wahrscheinlichkeiten für menschliche Fehler vor. Eine Fehlerrate von 0,01 ist hierbei als eine fehlerhafte Handlung pro 100 Handlungen zu verstehen. Die von HINZEN (1993) theoretisch hergeleiteten Fehlerwahrscheinlichkeiten beziehen sich auf den Bereich Schienenverkehr und können je nach Art der Aufgabe variieren, trotzdem ermöglichen sie es, die Größenordnung menschlicher Fehler für eine bestimmte Handlung, wie beispielsweise das Erfassen von Oberflächenschäden aus Bilddaten in einem Massendatenprozess, einzuschätzen. Die Handlungen, die das Auswertepersonal im Bereich der Straßenzustandserfassung ausführen, können mehrheitlich der Kategorie regelbasiertes

Verhalten zugeordnet werden, da die zu erfassenden Schadenstypen sowie das Erfassungsprinzip in Form von Regelwerken und Leitfäden explizit definiert sind. Nur in Grenzfällen muss auf Wissen und Erfahrung zurückgegriffen werden, um eine Entscheidung zu treffen und diese in eine konkrete Handlung zu überführen. Es kann demnach eine menschliche Fehlerwahrscheinlichkeit bzw. Fehlerrate im Sinne von HINZEN (1993) zwischen 0,01 und 0,1 angenommen werden.

Tabelle 3: Menschliche Fehlerwahrscheinlichkeit in Abhängigkeit von Umweltbedingungen und Stressniveau (nach HINZEN 1993).

Handlung	günstige Umweltbedingungen			Ungünstige Umweltbedingungen		
	Unterforderung	Mittel	Überforderung	Unterforderung	Mittel	Überforderung
fertigkeitsbasiert	0,002	0,001	0,002	0,01	0,005	0,01
regelbasiert	0,02	0,01	0,02	0,1	0,05	0,1
wissensbasiert	0,2	0,1	0,5	1,0	0,5	1,0

Um die Qualität der manuellen Auswertung von Oberflächenschäden im Zuge einer ZEB-Kampagne zu sichern, sehen die Regelwerke bzw. Richtlinien zur Zustandserfassung von Straßen die sogenannte Eigenüberwachung vor. Sie muss vom jeweiligen Dienstleister über Wiederholungsmessungen bzw. Wiederholungsauswertungen in einem zweiwöchigen Turnus realisiert werden. Diese Maßnahme dient dem Nachweis, dass die Messsysteme fehlerfrei funktionieren und der Auswerteprozess korrekt abläuft. Im Regelwerk ZTV ZEB-Stb (FGSV 2006) ist dieses Kontrollinstrument für den Anwendungsfall im Bundesfernstraßennetz beschrieben. Darauf beziehen sich auch die Ausführungen im Arbeitspapier AP 9 K 2.1 (FGSV 2016), welches für das kommunale Straßennetz maßgeblich ist. Für das Teilprojekt TP3 (Oberflächenschäden) läuft die Eigenüberwachung folgendermaßen ab: Zunächst wird vom Dienstleister ein Streckenabschnitt von 2 km Länge ausgewählt, dessen mittlerer Zustandswert bei mindestens einem Zustandsmerkmal (z. B. Riss oder Flickstelle) im Mittel zwischen Note 2,0 und 4,0 liegt. Dadurch soll sichergestellt werden, dass keine vollkommen intakten oder vollkommen geschädigten Abschnitte für die Prüfung genutzt werden. Ist der Streckenabschnitt gewählt, wird dieser von unterschiedlichem Personal ausgewertet, sodass im Ergebnis zwei Auswertungen für dieselbe Datengrundlage existieren. Für beide Erfassungen werden die abschnittswiseen Zustandswerte für Risse (ZWRIS) und Flickstellen (ZWFLI) berechnet, sodass 20 aufeinanderfolgende Auswerteabschnitte von 100 m Länge mit jeweils zwei Zustandswerten ZWRIS und ZWFLI vorliegen. Nun wird die Differenz der Zweit- zur Erstauswertung berechnet (Absolutwert). Überschreitet der Mittelwert oder die Standard-

abweichung, berechnet über die Differenzwerte der 20 zweifach ausgewerteten Abschnitte, eine Toleranz von 0,5 Notenpunkten, so gilt die Eigenüberwachung als nicht bestanden (siehe Tabelle 4). Da die Eigenüberwachung für den Bereich der Oberflächenschäden von einer manuell-visuellen Inspektion durch geschultes Auswertepersonal ausgeht, kann die Toleranz von 0,5 Notenpunkten demnach als die Fehlerquote bezeichnet werden, die für eine manuelle Auswertung von Oberflächenschäden als „Faktor Mensch“ akzeptiert wird.

Tabelle 4: Geforderte Auswertegenauigkeit beim Vergleich zweier manueller Auswertungen im Rahmen der Eigenüberwachung gemäß ZTV ZEB-Stb¹ (nach FGSV 2006).

	σ Differenzen der Zustandswerte	σ Differenzen der Zustandswerte
Risse	0,5	0,5
Flickstellen	0,5	0,5

Die Frage, welche Faktoren die manuelle Auswertung von Oberflächenschäden aus Kamerabilddaten und in welchem Maße beeinflussen, stand im Fokus verschiedener Studien, die von der BAST betreut wurden. Beispielsweise zeigen die Untersuchungen von SKAKUJ et al. (2019), dass der Feuchtegrad der Fahrbahnoberfläche, die Beleuchtung, der Bewölkungsgrad sowie der Sonnenstand Einfluss auf die mittels Oberflächenkameras aufgenommenen Bilder und damit die Sichtbarkeit von Oberflächenschäden haben. Auf diese Erkenntnisse aufbauend wurden im Rahmen der Studie von BERTHOLD et al. (2023) die Einflussfaktoren Mensch, Technik und Umwelt auf die Qualität der Zustandserfassung im Teilprojekt TP3 (Oberfläche) vertieft systematisch analysiert. Neben dem Messsystem (Bildmaterial) und unterschiedlichen Wetterszenarien wurde auch der Einfluss des Auswertenden auf die Erfassungsqualität von Substanzmerkmalen gemäß ZTV ZEB-Stb (FGSV 2006) im Hinblick auf die Reproduzierbarkeit untersucht. Eines der Experimente dieser Studie untersuchte die Fehlerrate, die bei der manuellen Auswertung von Rissen zu erwarten ist. Den Teilnehmern, von denen 10 Experten und 24 Laien waren, wurden 400 Bildausschnitte von 1 m² großen Straßenoberflächen präsentiert. Die Hälfte der Bilder zeigte ungeschädigte Oberflächen, die andere Hälfte zeigte Oberflächen mit diversen Schädstellen. Die Auswerter mussten entscheiden, ob der Bildausschnitt einen Riss zeigt oder nicht. Wie aus dem linken Diagramm in Abbildung 16 zu entnehmen ist, lag die Fehlerrate insgesamt in einem Bereich von 0,03 bis 0,2 (Mittelwert: 0,08). Eine Auswertezeit von 1 bis 3 Sekunden war nötig, um die

¹ Die Normierungsfunktion für Flickstellen ist aktuell nur in den vor dem Jahr 2018 erschienenen Ausgaben des Regelwerkes aufgeführt, da eine explizite Bewertung von Flickstellen zugunsten der sogenannten Restschadensfläche verworfen wurde.

Entscheidung „Riss vorhanden“ oder „kein Riss vorhanden“ für ein 1 m² großes Oberflächenelement zu treffen. Während sich die Fehlerrate insgesamt zwischen Experten und Laien nur wenig unterscheidet, zeigt sich in der Art der Fehler ein Unterschied, wie aus dem rechten Diagramm der Abbildung 16 hervorgeht: so tendieren Experten eher dazu Fehler zweiter Art zu machen, also Risse dort zu erkennen, wo keine vorhanden sind (False Positive). Bei Laien überwiegt der Fehler erster Art. Sie neigen eher dazu, Risse zu übersehen (False Negative). In beiden Fällen traten die Fehler vor allem bei sehr feinen Oberflächenstrukturen auf. Eine wesentliche Erkenntnis der Studie ist, dass der Faktor Mensch im Auswerteprozess von Oberflächenschäden den größten Einfluss auf die Wiederholbarkeit der Erfassungsergebnisse hat und zwar in einem solchen Maße, dass er Technik- und Umweltaspekte deutlich überprägt. Darüber hinaus ermittelten die Autoren der Studie mithilfe von qualitativen Interviews Ursachen für die festgestellte Spannbreite der Qualität einer manuellen Schadenserfassung (BERTHOLD et al. 2023). Als hauptsächliche Fehlerquelle wurde dabei der monotone Arbeitsprozess identifiziert, der akribische Konzentration trotz streckenweiser sehr geringer Schadensdichte und -intensität unter nicht unerheblichem Zeitdruck erfordert. Die Auswirkungen sind ein allgemeines Nachlassen der Konzentrationsfähigkeit im Tagesgang sowie asthenopische Beschwerden, die auch für andere Formen intensiver Bildschirmarbeit typisch sind (DEGLE 2006). Eine weitere im Rahmen der Befragung ermittelte Fehlerursache ist die sporadisch auftretende Unsicherheit in der Auslegung der Auswertevorschriften der Regelwerke.

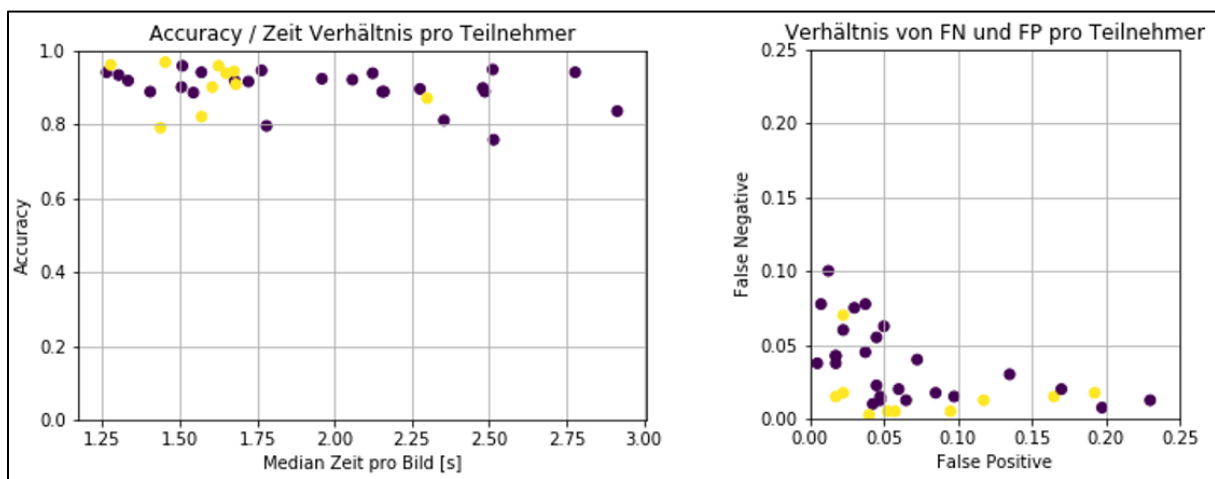


Abbildung 16: Qualität der Risserkennung von Experten (gelb) und Laien (violett) (nach BERTHOLD et al. 2023).

Zusammenfassend ist festzuhalten, dass allen Bestrebungen zum Trotz, einen objektiven und systematischen Erfassungs- und Bewertungsprozess auf Basis hochauflösender Kamera-

bilddaten zu etablieren, die Qualität der Zustandserfassung und -bewertung maßgeblich vom Auswertepersonal abhängt. Im selben Maße verdeutlicht sich hieran das Potential einer automatischen, messdatenbasierten Auswertung von Straßenoberflächenschäden.

2.2 Mobile Laserscanning

Mobile Mapping bezeichnet das von einer beweglichen Trägerplattform ausgehende Erfassen georeferenzierter Informationen. Während luftgestützte, digitale Systeme seit den 1990er-Jahren zur flächenhaften Erfassung der Topographie eingesetzt werden, sind kommerzielle Mobile-Mapping-Systeme seit Mitte der 2000er-Jahre verfügbar (FRITSCH 2015). Zu den Anwendungen, die auf Mobile-Mapping-Daten aufbauen, zählen unter anderem die Kartierung der Verkehrsinfrastruktur und ihres Zustands, die Vermessung und Kartierung von Straßeninventar und Versorgungseinrichtungen, die Vegetationskartierung und Baumkataster, das Building Information Modeling und das autonome Fahren. Mobile Mapping liefert demnach eine wichtige Datenquelle für die Planung, Simulation, Kartierung und Visualisierung dreidimensionaler digitaler Stadträume. Weiterführende Informationen zu potentiellen Anwendungen im Kontext der Straßenraumerfassung sind bei GUAN et al. (2016), WANG et al. (2019) und WANG et al. (2020) aufgeführt.

Mobile-Mapping-Systeme integrieren verschiedene Primär- und Sekundärsensoren auf einer in sich starren Plattform (PUENTE et al. 2013). Zu den Primärsensoren gehören Navigations- und Positionssensoren. Die direkte Georeferenzierung des Mobile-Mapping-Systems wird unter Nutzung globaler Navigationssatellitensysteme (GNSS) realisiert. Typischerweise werden zusätzlich leistungsfähige Inertialsysteme (INS) eingesetzt, sodass die Positionierung und Orientierung der Messplattform im übergeordneten Koordinatensystem noch genauer bestimmt werden kann (HOFMANN 2017). Die Sekundärsensoren können sowohl aktive als auch passive bildgebende Systeme umfassen, aus deren Daten Umgebungsinformationen extrahiert und kartiert werden können. Typischerweise werden Mobile-Mapping-Systeme zur Erfassung von Umgebungsdaten mit Kamera-, Laser- und Radarsensoren ausgerüstet. Um georeferenzierte Daten zu erhalten, müssen alle Sekundärsensoren mit den Primärsensoren zeitlich synchronisiert, sowie deren innere und äußere Kalibrierung bekannt sein (LUHMANN 2018). KUHLMANN & KLINGBEIL (2017) geben einen Überblick über Synchronisierungs- und Berechnungsprozesse innerhalb mobiler Multisensorsysteme, die für die räumliche Zuordnung aller Sensordaten und die Berechnung der Position und Orientierung der Messplattform im übergeordneten Koordinatensystem essentiell sind.



Abbildung 17: Beispiel für kompaktes und trägerfahrzeugunabhängiges Mobile-Mapping-System. Es umfasst neben einem inertialen Positionierungssystem zwei Laserscanner und sechs Kameras (BIENERT et al. 2018).

Gängige Kompaktsysteme bzw. transportable Komplettlösungen für die bodengestützte, mobile Datenerfassung sind beispielsweise das Trimble MX9 (TRIMBLE INC. 2021), Topcon IP-S3 (TOPCON CORPORATION 2021), Optech Lynx HS600 (TELEDYNE OPTECH INCORPORATED 2021), Leica Pegasus Two Ultimate (LEICA GEOSYSTEMS AG 2021) und Riegl VMX 250 (RIEGL LASER MEASUREMENT SYSTEMS GMBH 2012). Letzteres ist in Abbildung 17 dargestellt. Während die vorgenannten Lösungen hinsichtlich ihrer Sensorausstattung als Allround-Systeme für allgemeine Fragestellungen im Mobile-Mapping-Kontext angesehen werden können, existieren auch Speziallösungen mit anwendungsspezifischer Sensorik, wie beispielsweise MESAS (BAST 2018b), RoadSTAR (AIT 2019) oder S.T.I.E.R (LEHMANN+PARTNER GMBH 2022). In Abbildung 18 sind diese Messfahrzeuge dargestellt, die u. a. auch mit Oberflächenlaserscannern ausgerüstet sind.



Abbildung 18: Beispiele für Messfahrzeuge, die speziell für die Straßenzustandserfassung mit Laserscannern ausgerüstet sind: MESAS (I aus BAST 2018b), RoadSTAR (II aus AIT 2019) und S.T.I.E.R (III aus LEHMANN+PARTNER GMBH 2022).

Die grundlegende Datenaufnahmetechnologie der vorliegenden Arbeit ist das mobile Laserscanning (engl.: Mobile Laser Scanning, MLS). Dabei handelt es sich um einen Teilbereich des Mobile Mappings, bei dem der hauptsächliche Sekundärsensor ein Laserscanner ist (GLENNIE 2009). Allen Laserscannern gemein ist das grundlegende Funktionsprinzip, dass Objekte in ihrer Umgebung sequenziell punktförmig abgetastet werden. Dabei wird jeder Messpunkt durch Polarkoordinaten (Winkel und Entfernung) und Reflexionsintensitätswerte beschrieben. Zu den Kernkomponenten eines Laserscanners zählen das elektronische Distanzmessgerät (EDM) und ein optomechanisches Laserstrahlablenksystem (LIPKOWSKI & METTENLEITER 2019).

Seit die ersten MLS-Systeme zu Beginn der 2000er-Jahre vorgestellt wurden, hat sich diese Technologie in vielen Bereichen der Vermessung etabliert (GUAN et al. 2016). Nicht zuletzt deshalb, weil es durch sie erstmals möglich war, dreidimensionale Daten in hochdynamischen Umgebungen, wie Autobahnen oder innerstädtischen Straßen, hochauflösend, genau, flächendeckend und effizient zu erfassen, ohne Vermessungspersonal unnötigen Gefahren im Straßenverkehr auszusetzen (KUKKO et al. 2009).

LIPKOWSKI & METTENLEITER (2019) zufolge zählen Riegl (Österreich), Optech (Kanada), Faro (USA), Trimble (USA), Topcon (Japan), Surphaser (USA), Zoller+Fröhlich (Deutschland) und Leica (Schweiz) zu den bekanntesten Herstellern von Laserscannern im Bereich des Vermessungswesens. Auch das Fraunhofer Institut für Physikalische Messtechnik (Deutschland) entwickelt und vertreibt verschiedene Laserscannermodelle für Spezialanwendungen im Straßen- und Schienenbereich, wenn auch in deutlich geringeren Stückzahlen, als die vorgenannten Hersteller. Darüber hinaus werden für Mobile-Mapping-Anwendungen vermehrt Laserscanner verwendet, die ursprünglich aus dem Automotive-Bereich bzw. der Prozessautomations- und Logistikbranche stammen, wie beispielweise Modelle der Firmen Velodyne Lidar Inc. (USA) oder SICK AG (Deutschland).

In den folgenden Unterkapiteln werden die grundlegenden Technologien erläutert, die für Mobile-Laserscanning-Systeme wichtig sind: die Laserdistanzmessung und die satellitengestützte Positionsbestimmung.

2.2.1 Eigenschaften des Laserlichts

Das Akronym Laser steht für Light Amplification by Stimulated Emission of Radiation. Ein Laser erzeugt elektromagnetische Strahlung nach dem physikalischen Prinzip der stimulierten Emission von Photonen (CHAZETTE et al. 2016). Das dabei entstehende Laserlicht hat

Eigenschaften, die natürliches Licht nicht aufweist. „Je nach Art des eingesetzten Lasers ist das Laserlicht annähernd monochromatisch (exakt definierte Frequenz), stark gebündelt (kleiner Öffnungswinkel des Lichtstrahls), kohärent (Wellenpakete schwingen zueinander in Phase) und von hoher Intensität“ (WAGNER, ULLRICH & BRIESE 2003, S. 224).

Bei EICHLER & EICHLER (1991) sind die grundlegenden Eigenschaften von Laserlicht wie folgt beschrieben: Hinsichtlich des Wellenlängenspektrums besitzt Laserlicht ein ausgeprägtes Maximum in einem sehr engen Wellenlängenbereich, wohingegen für den Menschen sichtbares Licht deutlich breitbandiger ist (von 400 nm bis 700 nm). Laserlicht wird demnach als nahezu monochromatisch bezeichnet und weist eine hohe Leistungsdichte auf. Außerdem schwingen die elektromagnetischen Wellen mit einheitlicher Phasenlage. Damit ist Laserlicht zeitlich und räumlich kohärent, was für natürliche Lichtquellen meist nicht zutrifft. „Für die zeitliche Kohärenz gilt, dass eine feste Phasenbeziehung zwischen den Lichtwellen in Richtung der Strahlausbreitung gegeben ist. [...] Für die räumliche Kohärenz gilt, dass eine feste Phasenbeziehung zwischen den Lichtwellen transversal zur Strahlausbreitung gegeben ist“ (JUTZI 2007, S. 10). Die räumliche und zeitliche Kohärenz sind zentrale Eigenschaften von Laserlicht, aufgrund derer sehr kurze Pulse erzeugt und Signale zeitlich sehr hoch aufgelöst werden können. Die Kohärenz der Wellen ermöglicht es auch, dass ein Laserstrahl auf sehr geringe Strahldurchmesser fokussiert werden und demzufolge eine hohe Intensität erreichen kann.

Neben der mittleren Leistung, Wellenlänge oder Frequenz ist die Fokussierbarkeit von Lasern ein wichtiges Qualitätskriterium, um einen Laser für eine bestimmte Anwendung zu bewerten (EICHLER, DÜNKEL & EPPICH 2004). Sie hängt mit der sogenannten Strahldivergenz zusammen, die nachstehend erklärt wird. Im Allgemeinen fällt die Intensität eines Laserstrahls orthogonal zu dessen Ausbreitungsrichtung ab. Im linken Teil von Abbildung 19 ist die Intensitätsverteilung I_{r_0} eines in der Lasertechnik oft verwendeten Gaußstrahls bei seiner Aussendung in radialer Richtung ρ dargestellt. Die mittlere Teilabbildung zeigt, dass der Laserstrahl (rot) im Zuge seiner Ausbreitung von links nach rechts mit zunehmender Entfernung r einer gewissen Aufweitung unterliegt, die als Strahldivergenz θ bezeichnet wird. Die rechte Teilabbildung zeigt die Intensitätsverteilung I_{r_R} auf dem Zielobjekt in einer gewissen Entfernung r . Die Ausdehnung in radialer Richtung ρ ist größer geworden und die maximale Intensität ist dort niedriger als bei der Emission des Strahls an der Apertur. Die Strahldivergenz entsteht infolge der Beugung des Laserstrahls beim Austritt aus der Apertur und liegt typischerweise in einem Bereich zwischen 0,3 mrad und 2 mrad (JUTZI 2007).

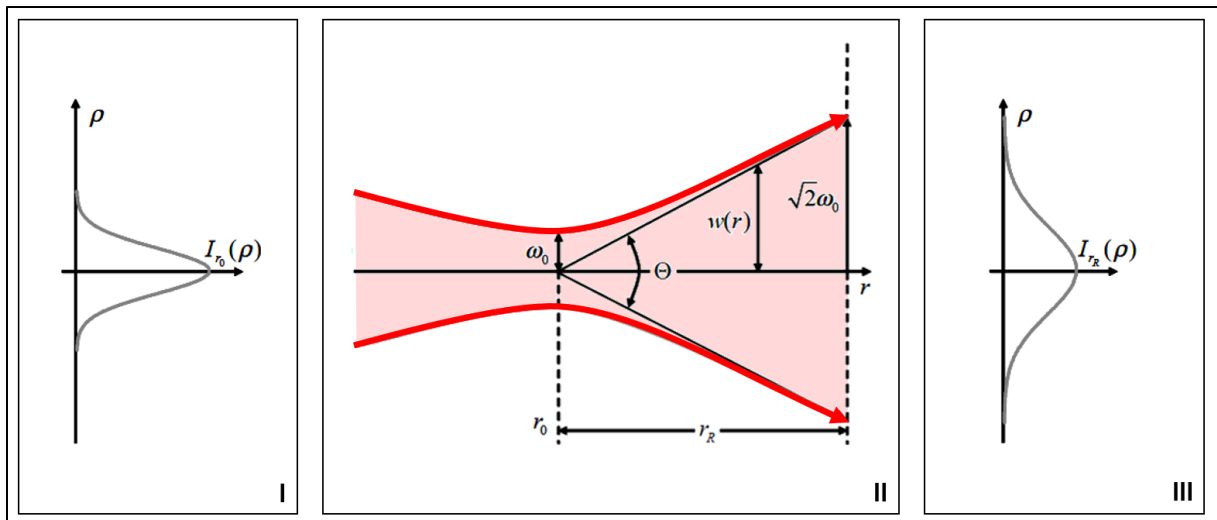


Abbildung 19: Ausbreitung eines gaußförmigen Laserstrahls mit entfernungsabhängiger Strahldivergenz (nach JUTZI 2007).

Wie in Abbildung 20 dargestellt, hat die Strahlaufweitung zur Folge, dass ein Laserstrahl beim Auftreffen auf das Zielobjekt eine gewisse Fläche in Form eines Kreises oder einer Ellipse einnimmt, die man auch als Footprint bezeichnet. Vor allem an Kanten können Signalformen entstehen, die unter Umständen mit eingeschränkter Genauigkeit von der Detektionseinheit weiterverarbeitet werden können.

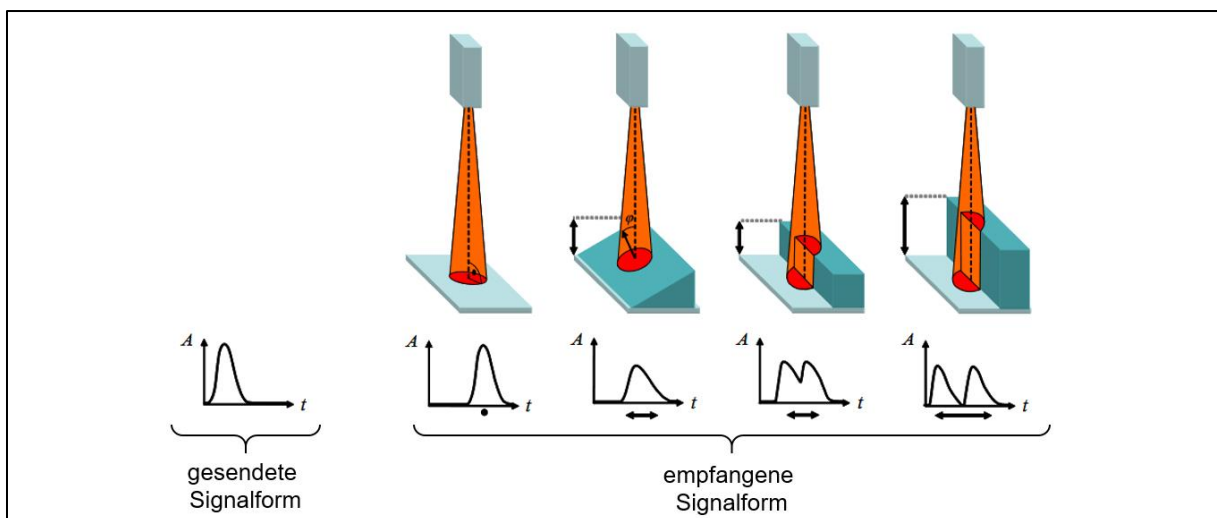


Abbildung 20: Einfluss unterschiedlicher Oberflächenformen auf die Signalform einer Lasermessung dargestellt in Diagrammen mit Amplitude A und Zeit t (nach JUTZI 2007).

Wie bereits dargestellt, hat die Geometrie des Zielobjektes im Zusammenspiel mit der Strahlaufweitung Einfluss auf die Signalform. Zusätzlich beeinflussen die Reflexionseigenschaften der Zieloberfläche, wie beispielsweise deren Rauigkeit, welche Anteile des emittierten Laserlichtes von einer Detektoreinheit empfangen und ausgewertet

werden können. Abbildung 21 zeigt vereinfachende Modellannahmen zur Reflexion auf Oberflächen. Die linke und rechte Teilabbildung zeigen die Reflexion (blau) eines einfallenden Laserstrahls (rot) auf einer ideal spiegelnden und einer ideal diffus reflektierenden Oberfläche. Die ideal diffus reflektierende Oberfläche ist gemäß dem Lambertschen Gesetz dadurch gekennzeichnet, dass einfallende Strahlung unabhängig von der Richtung, aus der sie kommt, in alle Richtungen gleich gestreut wird. Sie erscheint aus allen Betrachtungsrichtungen also gleich hell (JUTZI 2007). Bei einer spiegelnden Oberfläche sind Einfallswinkel α und Ausfallwinkel α' des Laserstrahls in Bezug zu Normale der Oberfläche identisch. In der Realität ist von Mischformen, wie beispielsweise einer gerichtet diffusen oder grob spiegelnden Reflexion, auszugehen. Diese sind in den beiden mittleren Teilabbildungen von Abbildung 21 dargestellt.

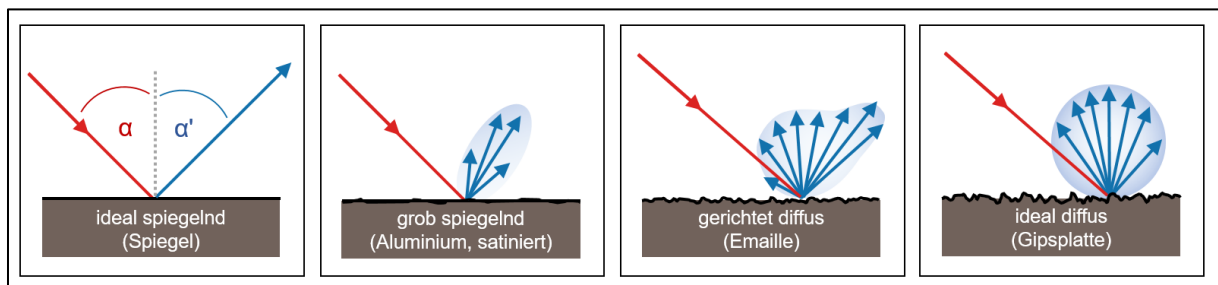


Abbildung 21: Reflexionen an Oberflächen unterschiedlicher Rauheit (nach JUTZI 2007; LANGE 2022).

Wie bereits erwähnt, ist eine wesentliche Eigenschaft von Laserlicht, dass sehr hohe Intensitäten erreicht werden können. In diesem Zusammenhang ist auch das Thema Sicherheit und Laserklassen von Relevanz. Die Laserklassen sind in einer internationalen Norm definiert und reichen von ungefährlichen, augensicheren Lasern der Klasse 1 bis zu den aufgrund ihrer hohen Leistung als gefährlich eingestuften Lasern der Klasse 4 (DIN EN 60825 2013). Werden Laser der Klasse 3 oder 4 im öffentlichen Raum eingesetzt, müssen bestimmte Vorkehrungen getroffen werden, da sonst die Gefahr besteht, dass Personen im öffentlichen Raum durch unbeabsichtigte Reflexionen oder direkte Einwirkung des Lasers Schaden an Augen und Haut nehmen können.

Die in diesem Kapitel beschriebenen Eigenschaften des Laserlichtes sind der Grund dafür, dass sich Laser seit ihrer Einführung in den 1960er-Jahren als essentieller Bestandteil von Informations- und Kommunikationstechnologien, medizinischen Anwendungen und Messtechnik etabliert haben. Für Messungen, die mit einem laserbasierten System durchgeführt werden, hat sich im Bereich der Geowissenschaften und der Vermessung der Begriff Light Detection and Ranging (LiDAR) durchgesetzt.

2.2.2 Elektrooptische Entfernungsmessung

Laser werden im Bereich der Geodäsie zur berührungslosen Entfernungsmessung zwischen dem Messgerät und dem Zielobjekt eingesetzt. Hierfür nutzen Laserscanner im allgemeinen ein Entfernungsmessgerät zur elektrooptischen Distanzmessung und ein Laserstrahlableitungs- und -abtastsystem zur flächenhaften Abtastung. Laserscannersysteme können anhand verschiedener technischer Kriterien typisiert werden (JUTZI 2007). Beispielsweise können sie auf Basis ihrer Reichweite und der damit verbundenen Art der Entfernungbestimmung (Messprinzip) in verschiedene Gruppen eingeteilt werden (BERALDIN 2004). Nahbereichsscanner für Messbereiche zwischen 0,1 m bis 5 m arbeiten häufig nach dem Triangulationsverfahren. Für Laserscanner, die in einem mittleren Entfernungsbereich von 1 m bis 100 m messen, ist das Phasenvergleichsverfahren typisch. Das Impulslaufzeitverfahren wird dagegen zumeist bei Laserscannern eingesetzt, die Reichweiten von mehreren hundert Metern erreichen sollen. Nachfolgend werden die drei genannten Messprinzipien kurz erläutert.

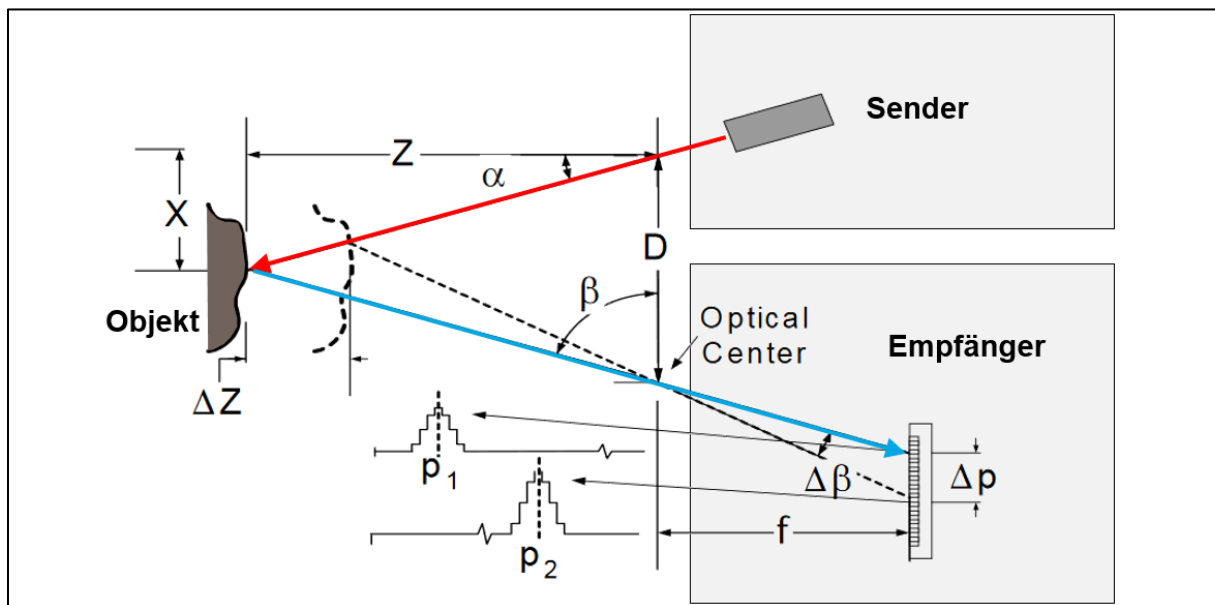


Abbildung 22: Prinzip der laserbasierten optischen Triangulation (nach BERALDIN 2004).

Das dem Triangulationsverfahren zugrundeliegende Prinzip realisiert die Abstandsmessung zur Objektoberfläche durch Winkelberechnungen (siehe Abbildung 22). Eine Laserquelle projiziert typischerweise einen Punkt oder eine Linie in einem bestimmten Winkel α auf die zu messende Objektoberfläche (rot). Dort wird das emittierte Laserlicht reflektiert (blau) und trifft in einem bestimmten Winkel β auf einem Detektorelement auf. In Abhängigkeit zur Entfernung Z des Messobjektes verändert sich der Auftreffwinkel $\Delta \beta$ und die Position Δp , an der das Laserlicht auf dem Detektorelement auftrifft. Mithilfe des Abstands D zwischen Sende- und Empfangs-

einheit sowie der Auftreffposition des Lasers auf dem Detektor kann der Abstand Z zum Messobjekt berechnet werden. Bei BERARDIN (2004) ist dieses Verfahren im Detail beschrieben.

Das Impulsmessverfahren, das auch als Time Of Flight (TOF) Verfahren bezeichnet wird, basiert auf der Messung der Laufzeit eines gesendeten Impulses bis zu seiner Detektion (WEHR & LOHR 1999). Wie in Abbildung 23 dargestellt, bestehen Laserscansysteme, die nach dem Impulsmessverfahren arbeiten, im Wesentlichen aus drei Komponenten: einem Sender, einem Empfänger und einem Zeitmesssystem. Eine von der Laserquelle kurzzeitig emittierte Trägerwelle mit aufmoduliertem Impuls A_S wird vom Zielobjekt reflektiert und von einer Empfangseinheit als Wellenpaket A_E detektiert. Ein Zeitmesssystem vergleicht die Laufzeit Δt zwischen ausgesendetem und empfangenem Impuls. Der Abstand Z zum Zielobjekt wird aus der Lichtgeschwindigkeit im Vakuum und dem jeweils vorliegenden Brechungsindex berechnet. Der Brechungsindex quantifiziert dabei die Ausbreitungsgeschwindigkeit von Licht in einem bestimmten Medium. Im Allgemeinen zählen zu den Vorteilen von Scannern, die nach dem Impulslaufzeitverfahren arbeiten, die großen Reichweiten und die Augensicherheit dieser Systeme. Zu den Nachteilen wird die im Vergleich zum Phasenvergleichsverfahren geringere Genauigkeit gezählt. Typischerweise werden TOF-Scanner im Airborne-Laserscanning-Bereich eingesetzt.

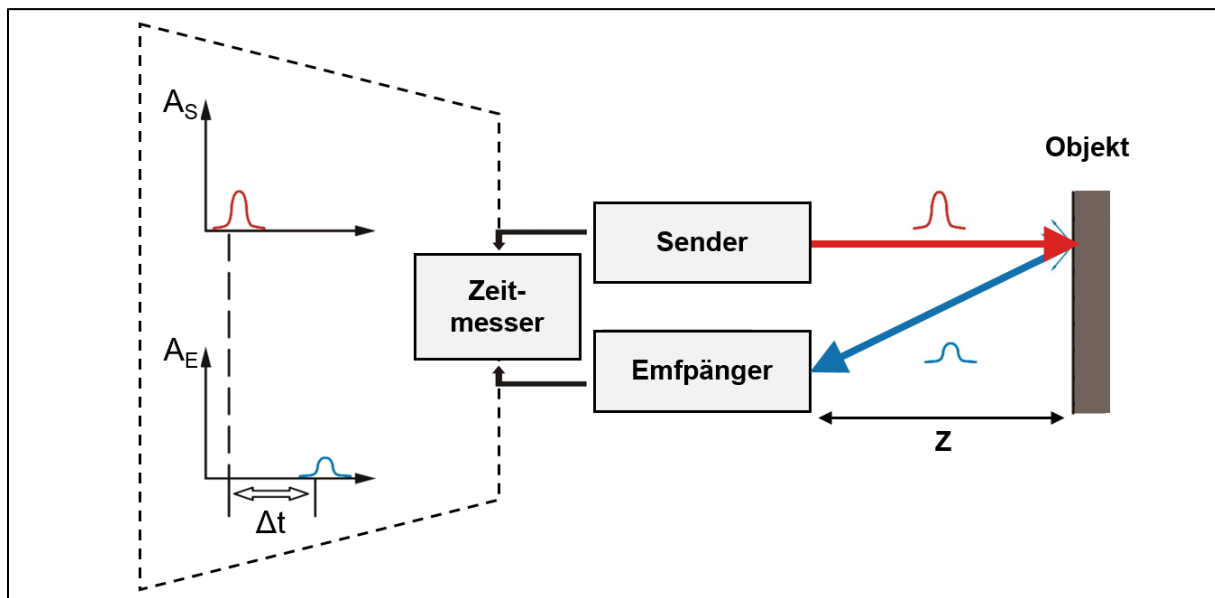


Abbildung 23: Prinzip der elektrooptischen Distanzmessung mithilfe des Impulslaufzeitverfahrens (nach WEHR & LOHR 1999; WAGNER, ULLRICH & BRIESE 2003).

Im Unterschied zum Impulsmessverfahren wird beim Phasenvergleichsverfahren, das auch als Continuous Wave (CW) Verfahren bezeichnet wird, ein kontinuierliches Signal emittiert (WEHR & LOHR 1999). Dabei handelt es sich um eine hochfrequente Trägerwelle, der eine Modulationswelle mit Wellenlänge λ aufmoduliert wird. In Abbildung 24 ist das Prinzip der Amplitudenmodulation veranschaulicht.

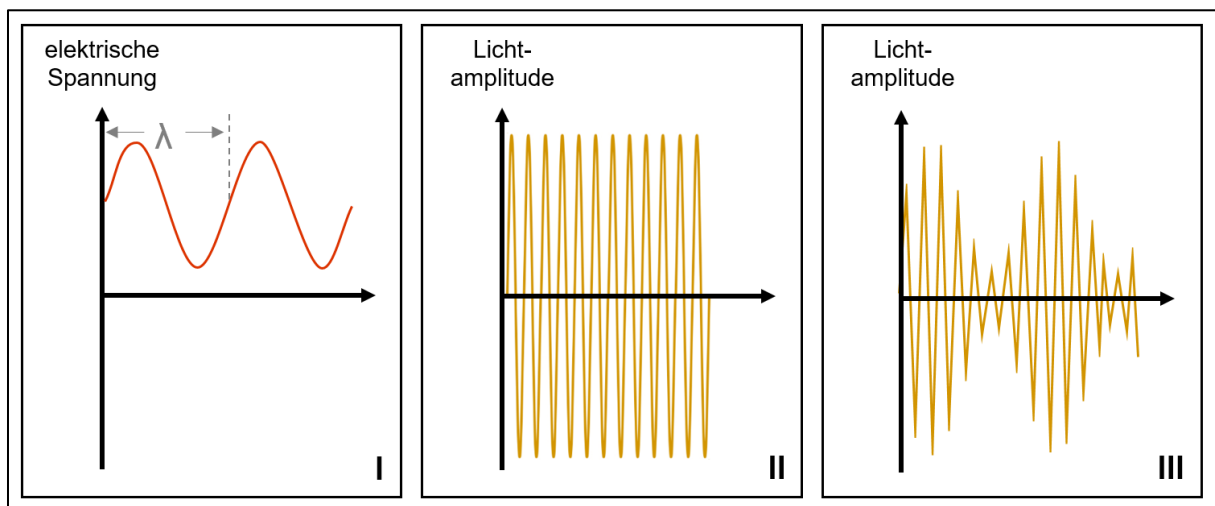


Abbildung 24: Prinzip der Amplitudenmodulation: Modulationswelle (I), Trägerwelle (II) und modulierte Trägerwelle (III). Dieses Prinzip ist die Basis für Laserscanner, die nach dem Phasenvergleichsverfahren arbeiten (nach WITTE, SPARLA & BLANKENBACH 2020).

Wie in Abbildung 25 schematisch dargestellt, wird am Sender die elektromagnetische Welle unter einem definierten Phasenwinkel φ_S emittiert und das Signal am Messobjekt zurück in Richtung Empfänger reflektiert. Der Phasenwinkel φ_E , der am Empfänger detektiert wird, ist neben der Ausbreitungsgeschwindigkeit der Welle auch von der zurückgelegten Distanz abhängig. Diesen Umstand nutzt das Phasenvergleichsverfahren, indem ein Phasenmesser die Phasendifferenz zwischen dem emittierten und reflektierten Signal misst. Aus dieser Phasenverschiebung $\Delta\lambda$ kann der Abstand zum Messobjekt ermittelt werden. Zur Abstandsmessung wird in einem ersten Schritt die Anzahl der vollständigen Wellendurchläufe über die Modulationswelle bestimmt. In einem zweiten Schritt wird die Länge des Reststücks $\Delta\lambda$ auf Basis der Trägerwelle ermittelt. Im Allgemeinen eignet sich dieses Messprinzip, wenn komplexe Geometrien in begrenzter Entfernung gemessen werden sollen und hohe Genauigkeiten bei hohen Messgeschwindigkeiten erforderlich sind.

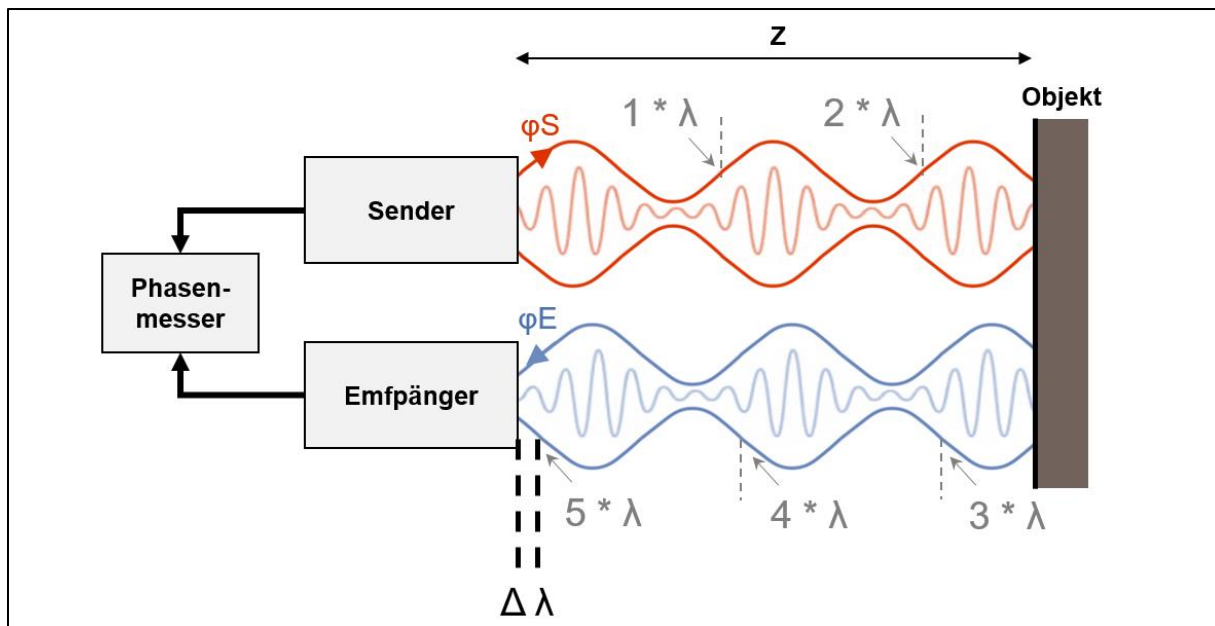


Abbildung 25: Prinzip des Phasenvergleichsverfahrens. Die äußere Wellenform symbolisiert die Modulationswelle die innere Wellenform die Trägerwelle (nach KUHLMANN & HOLST 2017).

2.2.3 Scansysteme

Um Oberflächen mit einem Laser schnell und flächenhaft zu erfassen – sie also zu scannen – werden typischerweise optomechanische Ablenkeinrichtungen eingesetzt. Diese lenken den Laserstrahl oder -impuls in einem bestimmten Muster gegen die Bewegungsrichtung der Messplattform in Richtung des Zielobjekts ab. Das resultierende Muster des Scanstreifens ist abhängig von Art und Bewegung der Ablenkeinrichtung, für die verschiedene konstruktive Lösungen existieren. Nachfolgend werden exemplarisch drei Scansystemtypen beschrieben. In Abbildung 26 sind diese Ablensysteme und die entsprechenden Abtastmuster schematisch dargestellt.

Bei Scansystemen mit oszillierendem oder schwingendem Spiegel befindet sich dieser typischerweise vor der waagrecht angebrachten EDM. Durch einen Motor angetrieben, schwingt der Spiegel zwischen zwei Extrempositionen. Mechanisch bedingt wird die Spiegelbewegung im Bereich ihrer Auslenkextrema abgebremst. Zwischen den Umkehrpunkten erfährt sie eine Beschleunigung. Dadurch entsteht ein Scanstreifen mit einem sinusförmigen Abtastmuster, bei dem die Punktdichte zu den Umkehrpunkten hin ansteigt (PETRIE 2011).

Wird der Spiegel durch die Ablenkmechanik sowohl gekippt als auch in Rotation versetzt, spricht man von einem System mit nutierendem Spiegel. Ein solches System produziert während der Vorwärtsbewegung der Plattform einen kontinuierlichen Streifen von sich überlappenden Ellipsen. Das resultierende Abtastmuster wird als Palmer-Scan bezeichnet. In der

Überlappung der elliptischen Scans liegt der große Vorteil dieses Systemtyps. Objektpunkte können dadurch wiederholt von verschiedenen Positionen aus aufgenommen werden (PETRIE 2011).

Die rechte Teilabbildung von Abbildung 26 zeigt einen dritten Systemtyp. Hier wird anstelle eines Schwingspiegels ein Polygonspiegel verwendet, der mit konstanter Geschwindigkeit rotiert. Daraus resultiert eine lineare Ablenkung des Laserstrahls zwischen den Streifenseiten. Auf der Zieloberfläche entstehen parallele Punktreihen mit gleichmäßiger Punktverteilung. Gegenüber Systemen mit Schwingspiegeln zeichnen sich solche mit Polygonspiegeln durch die Gleichmäßigkeit der Messpunktverteilung und die Stabilität der Scaneinrichtung aus. Ferner ermöglichen Polygonspiegelsysteme eine hohe Scanrate (KOHLSTOCK 2011).

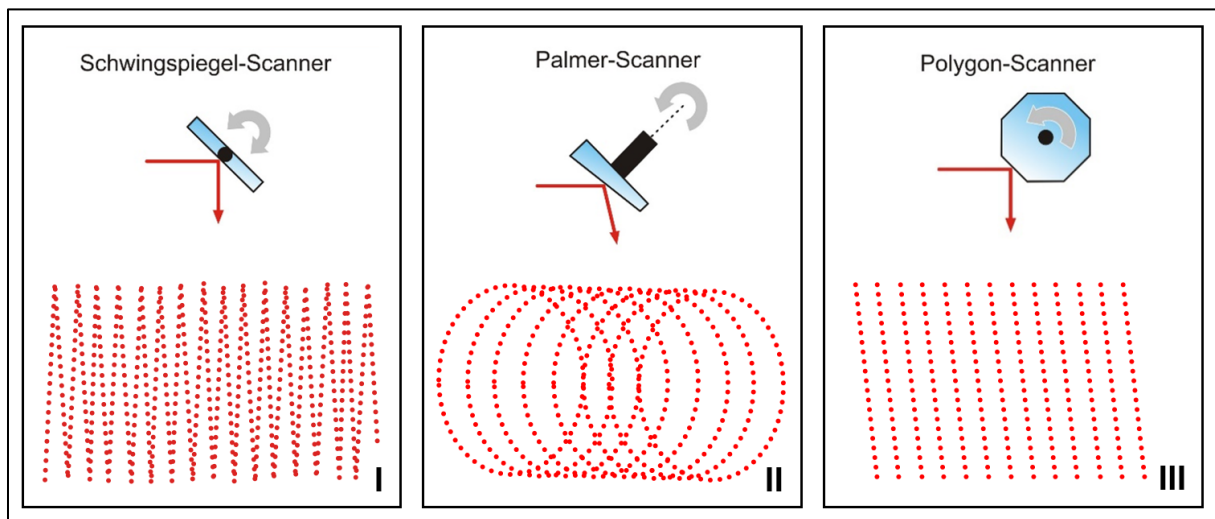


Abbildung 26: Beispiele für gängige Systemtypen und Abtastmuster von Laserscannern (nach KOHLSTOCK 2011; CHAZETTE et al. 2016).

CHAZETTE et al. (2016) bieten einen umfassenden Überblick über die LiDAR-Technologie im Bereich der Fernerkundung. Weiterführende Informationen zum Themenbereich Laserscanning mit dem Fokus auf Airborne- und Land-based-Anwendungen sind bei BERALDIN, BLAIS & LOHR (2010) und KUKKO (2013) zu finden.

2.2.4 Satellitengestützte Positionsbestimmung

Zur direkten Georeferenzierung der kinematischen Messplattform wird für das Mobile Mapping typischerweise eine Kombination aus GNSS und INS genutzt (GREJNER-BRZEZINSKA 1999; GROVES 2013). GNSS liefern geografische Positions- und Geschwindigkeitsdaten einer Empfängerantenne durch die Nutzung einer Konstellation von Satelliten in der Erdumlaufbahn

(BAUER 2017). Neben dem U.S.-amerikanischen Navigation Satellite Timing And Ranging Global Positioning System (NAVSTAR-GPS) stehen mit dem russischen GLONASS, dem chinesischen BeiDou (BDS) und dem europäischen Galileo verschiedene Satellitennavigationssysteme für die satellitengestützte Positionsbestimmung zur Verfügung. Eine ausführliche Beschreibung dieser ist bei BAUER (2017) zu finden. Am Beispiel des amerikanischen NAVSTAR-GPS (kurz GPS genannt) wird nachstehend das GNSS-Grundprinzip erklärt.

2.2.4.1 Global Positioning System

Das globale Positionierungssystem GPS besteht aus einem Weltraum-, Kontroll- und Nutzersegment. Das Weltraumsegment umfasst eine Konstellation von 24 Satelliten auf sechs Umlaufbahnen, wobei sich jeweils vier Satelliten auf einem gemeinsamen Orbit befinden. Das Konzept der GPS-Satellitenkonstellation sieht vor, dass es theoretisch von jedem Standpunkt auf der Erde aus möglich ist, vier Satelliten simultan zu beobachten.

Die Satelliten senden kontinuierlich definierte Funksignale aus, die mithilfe von Empfängern im Nutzersegment registriert werden können. Das Grundprinzip von GPS basiert darauf, dass die als Pseudostrecken bezeichneten Entfernungen zwischen einem Empfänger und mehreren Satelliten über Signallaufzeitmessungen bestimmt werden. Die Position der Satelliten muss dafür bekannt sein. Das Kontrollsegment umfasst hierzu ein weltweites Netz aus Bodenstationen, die den Zustand und Status der Satelliten überwachen. Über das Kontrollsegment wird sichergestellt, dass die Satellitenpositionen durch sogenannte Ephemeriden (Bahndaten) beschrieben werden können. Zu diesem Zweck sammelt eine Master Control Station die Pseudostreckenmessungen zwischen zahlreichen Kontrollstationen und Satelliten mit dem Ziel, unter Berücksichtigung von meteorologischen Daten präzise Ephemeriden und Uhrkorrekturen zu berechnen, vorherzusagen und zu veröffentlichen. Die sogenannten Broadcast Ephemeriden, die unmittelbar verfügbar sind, ermöglichen eine auf etwa 2 m bis 5 m genaue Bestimmung der Koordinate der aktuellen Position. Präzise Ephemeriden werden mit einiger Verzögerung vom International GNSS Service (IGS) bereitgestellt, ermöglichen dann aber eine wesentlich genauere Positionsbestimmung im Zentimeterbereich. Eine Zwischenlösung stellen sogenannte Rapid Ephemeriden dar, deren Berechnung und Veröffentlichung etwa einen Tag dauert.

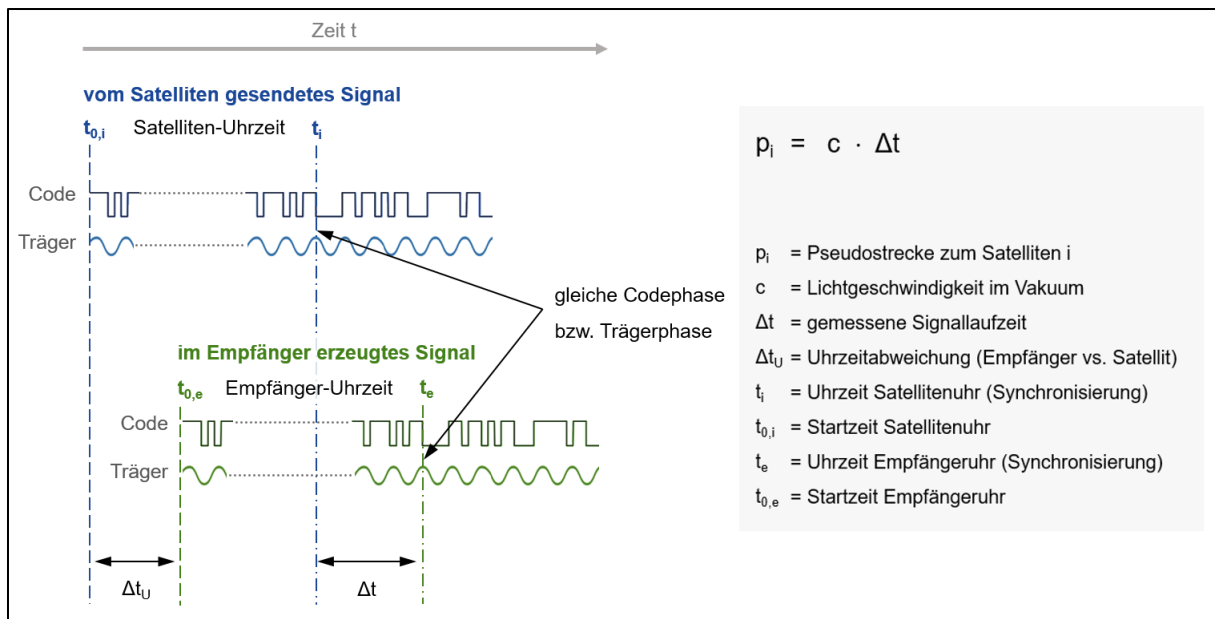


Abbildung 27: Prinzip der Pseudostreckenmessung (nach MANSFELD 2010).

Um mit einem Empfänger im Nutzersegment eine valide satellitengestützte Positionsbestimmung durchführen zu können, müssen mindestens vier Satelliten zur gleichen Zeit sichtbar sein. Sichtbarkeit meint hierbei, dass die Pseudostrecke zu den Satelliten bestimmbar sein muss. Die Pseudostrecke ergibt sich aus der wahren Strecke zwischen Satellit und Empfänger sowie einer Konstanten, die sich aus dem Uhrenfehler zwischen der Satelliten- und der Empfängeruhrzeit ableitet. Unter der Voraussetzung, dass die Positionen der Satelliten bekannt sind, ist es möglich, die Pseudostrecken unter Zuhilfenahme der Laufzeiten der von den Satelliten gesendeten Signale zu bestimmen. Abbildung 27 zeigt dies am Beispiel von nur einem einzigen Satelliten: Vom Satelliten werden zeitlich getaktete Codes auf die Trägerwellen L1 und L2 aufmoduliert und gesendet. Der C/A-Code ist dabei für die zivile, der P-Code für die militärische Nutzung vorgesehen. Die von den Satelliten gesendeten Codes lassen es zu, den sendenden Satelliten zu identifizieren. Der Empfänger kennt alle Satellitencodes, erzeugt Replikationen der entsprechenden Codes und sucht nach einem Maximum an Korrelation zwischen dem von ihm generierten und dem empfangenen Code. Sobald eine Übereinstimmung gefunden ist, kann die Ankunftszeit der Codesignale entsprechend der Empfängeruhrzeit sowie die Signallaufzeit bestimmt werden. Die Uhrzeiten von Sender und Empfänger laufen jedoch nicht synchron. Deshalb braucht es einen vierten Satelliten, mit dessen Hilfe der Uhrenfehler bestimmt und die initiale Laufzeitmessung korrigiert werden kann. Eine Positionsbestimmung ist prinzipiell auch möglich, wenn nur zu drei Satelliten Pseudostrecken bestimmt werden können, jedoch ist dies wegen der um den Uhrenfehler verfälschten Laufzeitmessung nur unter deutlichen Genauigkeitseinbußen möglich. Im Allgemeinen wirkt sich eine höhere Anzahl

sichtbarer Satelliten stabilisierend auf die Positionsbestimmung aus. Neben der beschriebenen Codemessung können auch Phasenmessungen auf den Trägerwellen der Signale durchgeführt werden. Damit sind im Vergleich zu Codemessungen deutlich höhere Genauigkeiten möglich. Eine umfassende und detaillierte Beschreibung von Grundlagen und Anwendungen der satellitengestützten Positionsbestimmung ist bei KAPLAN & HEGARTY (2017) zu finden.

2.2.4.2 Differentielle Korrekturverfahren

Es existieren verschiedene Korrekturmöglichkeiten, mithilfe derer die satellitengestützte Positionsbestimmung qualitativ verbessert werden kann. Nachfolgend werden die Konzepte Differential Global Positioning System (DGPS) und Real Time Kinematic (RTK) kurz vorgestellt (MANSFELD 2010). Beide Konzepte basieren darauf, dass GNSS-Referenzstationsdaten genutzt werden, um die Genauigkeit der Entfernungsmessung zwischen Satellit und Nutzerempfänger zu optimieren.

Dem DGPS bzw. DGNSS liegt die Annahme zugrunde, dass zwei räumlich nah beieinanderliegende Empfänger ähnlichen Messbedingungen und somit Fehlereinflüssen ausgesetzt sind. Sowohl eine ortsfeste Referenzstation mit bekannter Position als auch ein nahegelegener, mobiler Empfänger bestimmen ihre Position zeitgleich entsprechend der in Kapitel 2.2.4.1 beschriebenen Pseudostreckenmessung. Empfänger und Referenzstation sollten dabei nicht weiter als 200 km voneinander entfernt liegen. Für die Referenzstation können, aufgrund ihrer bekannten Position, Pseudostrecken zu Satelliten vorhergesagt und mit den gemessenen Pseudostrecken abgeglichen werden. Die festgestellten Differenzen werden dann als Korrektursignal von der Referenzstation an den nahegelegenen GNSS-Empfänger übermittelt (siehe Abbildung 28). Dieser wiederum kann die Korrekturdaten nutzen, um seine eigenen Pseudostreckenmessungen zu optimieren. Uhrenfehler, Ephemeridenfehler und atmosphärische Fehlereinflüsse können mithilfe von DGNSS reduziert werden, sodass eine Verbesserung der Positionierungsgenauigkeit für den Nutzerempfänger auf ca. 0,3 m bis 1 m erreicht werden kann. Die DGNSS-Korrektur kann sowohl in Echtzeit als auch im Post-processing und sowohl auf Basis von Code- als auch Phasenmessungen der Trägerwellen vorgenommen werden.

Geschieht eine DGNSS basierte Korrektur in Echtzeit unter Nutzung von Phasendaten, bezeichnet man dies als Real Time Kinematic (RTK) Messung (BAUER 2017). Die Phasenmehrwertigkeiten der Trägerwellen, die im Zuge der RTK-Optimierungen aufgelöst werden müssen, erfordern es, dass Referenzstation und Empfänger nur wenige Kilometer voneinander entfernt

sind. Mit RTK ist es auch möglich, den Fehlereinfluss von Multipath-Effekten und Messrauschen zu minimieren und so eine Positionierungsgenauigkeit im Zentimeterbereich zu realisieren.

Typischerweise erlauben Softwareprodukte kommerzieller Anbieter von GNSS-Instrumenten eine nachträgliche DGNSS- bzw. RTK-Korrektur der Empfängermessungen im Postprocessing. Über entsprechende Dienste, wie beispielsweise den Satellitenpositionierungsdienst der deutschen Landesvermessung (SAPOS), der ca. 260 Referenzstationen in ganz Deutschland betreibt, können hierzu archivierte Beobachtungsdaten von Referenzstationen bezogen werden. Abbildung 29 zeigt eine SAPOS-Station, die in Bamberg GNSS-Beobachtungen aufzeichnet.

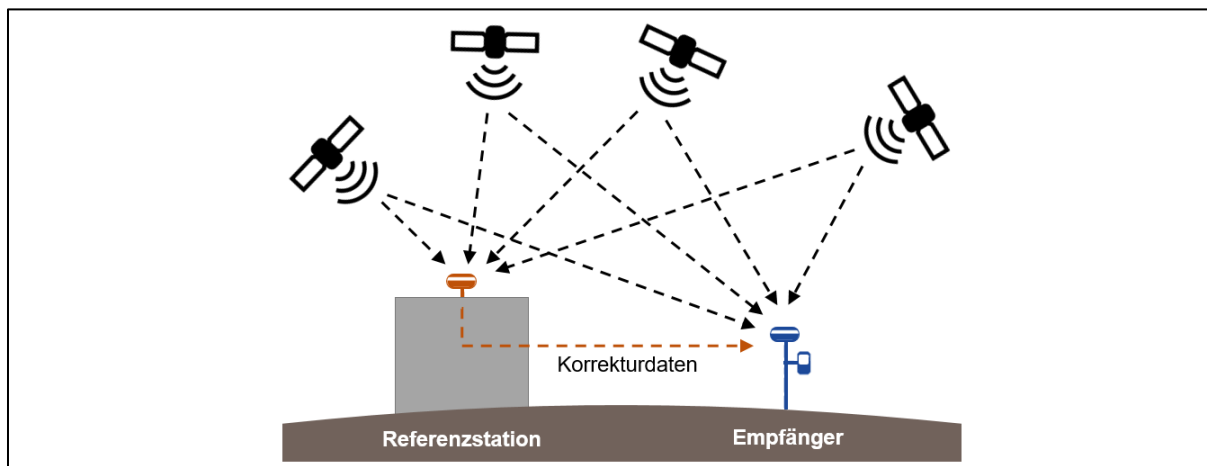


Abbildung 28: Prinzip des differentiellen GNSS (DGNSS).

Für die Übertragung von DGPS-Korrekturdaten gibt es verschiedene Verfahren, die auf den Standards RTCM und RTCA basieren. Für den Mobile-Mapping-Bereich ist in Deutschland vor allem NTRIP (Networked Transport of RTCM via Internet Protocol) relevant (WEBER, DETTMERING & GEBHARD 2005). Dabei handelt es sich um ein Verfahren zur Bereitstellung von GNSS-Korrekturdatenströmen über eine Internetverbindung. Die Verbindung wird typischerweise über Mobilfunknetze aufgebaut und der paketorientierte Funkdienst GPRS genutzt.

Mithilfe der vorgestellten Konzepte ist es prinzipiell möglich, die Genauigkeit der satellitengestützten Positionsbestimmung einer mobilen Messplattform so zu optimieren, dass deren absolute Position im Weltkoordinatensystem auf wenige Zentimeter genau bestimmt werden kann.

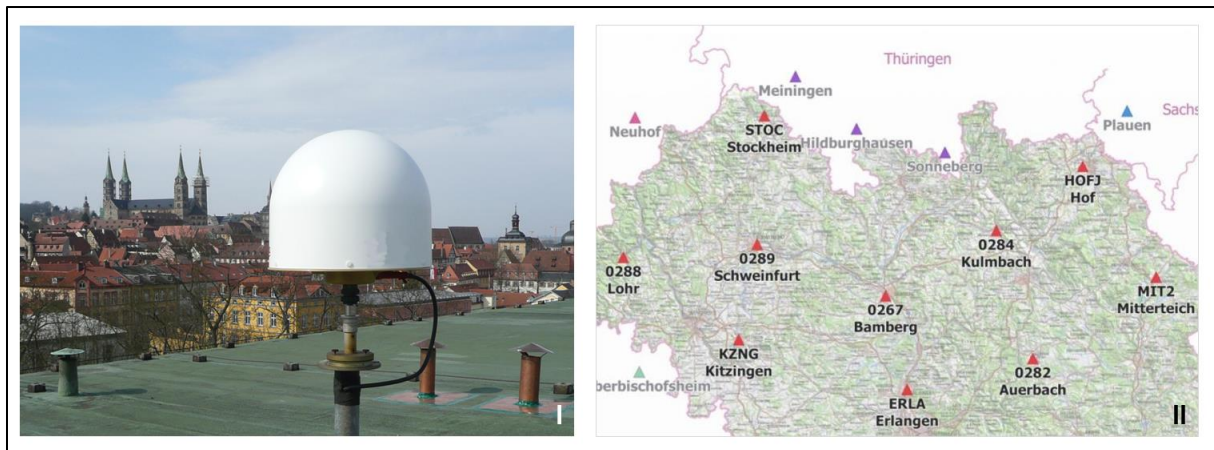


Abbildung 29: SAPOS-Referenzstationsantenne auf einem Dach in Bamberg (I aus LANDESAMT FÜR DIGITALISIERUNG, BREITBAND UND VERMESSUNG 2022). Ausschnitt aus einer Übersichtskarte von SAPOS-Stationen in Bayern (II aus LANDESAMT FÜR DIGITALISIERUNG, BREITBAND UND VERMESSUNG 2022).

2.2.4.3 Positionierungssysteme mit Inertialsystem

Wie vorstehend erläutert, müssen für einen GNSS-Empfänger die Signale von mindestens vier Satelliten sichtbar sein, um eine valide Positionsbestimmung zu realisieren. Kommt es auf Grund von Abschattungen zum Abriss eines oder mehrerer Signale, kann eine Positionierung nur mit eingeschränkter Genauigkeit oder gar nicht erfolgen. Außerdem sind GPS-Daten nur mit einer relativ niedrigen Frequenz von 10 Hz verfügbar, weshalb Mobile-Mapping-Systeme die Positionsbestimmung nicht auf Basis von GNSS allein realisieren (KUKKO 2013). Um im Allgemeinen in einer höheren Datenrate Informationen zur Ausrichtung der Messplattform zu erhalten und auch bei GNSS-Ausfall eine adäquate Positionsbestimmung gewährleisten zu können, sind Mobile-Mapping-Systeme typischerweise zusätzlich zu GNSS mit einem INS in Form eines sogenannten Strapdown-Systems ausgestattet. Ein INS berechnet die relative Position eines Objekts sowie dessen Positionsänderung über die Zeit mit Hilfe von Rotations- und Beschleunigungsmessungen einer inertialen Messeinheit (IMU). Eine IMU besteht aus Beschleunigungsmesssensoren (Akzelerometer) und Drehratensensoren (Gyroskop) für jede der drei Raumachsen (siehe Abbildung 30).

Ein Gyroskop misst die Drehgeschwindigkeit für die jeweilige Achse des INS. Für die in Gyroskopen eingesetzten Sensoren gibt es optische und mechanische Lösungen (HINÜBER 2002). Ein zentrales, bauartbedingtes Problem von Gyroskopen ist die sogenannte Kreisel drift, die vor allem bei mechanischen Lösungen auftritt. Die Drehachse des Kreiselinstruments wandert dabei langsam aus seiner ursprünglichen Lage heraus. Dieses Verhalten ist u. a. auf ein Zusammenwirken mechanischer Imperfektionen, lagerbedingter

Reibung und Auswirkungen der Erdrotation zurückzuführen. Weiterhin spielen Temperaturänderungen, elektrostatische Effekte und Vibrationen eine Rolle. „Die optischen Kreisel zeichnen sich gegenüber den mechanischen [...] Kreiseln dadurch aus, dass das sensierende Element keiner mechanischen Beanspruchung unterliegt und gegenüber Vibration und Beschleunigung weitestgehend unempfindlich ist“ (HINÜBER 2002, S. 588). Für Mobile-Mapping-Systeme kommen typischerweise Ringlasergyroskope (RLG), faseroptische Kreisel (FOG) und mikroelektronisch-mechanische Systeme (MEMS) zum Einsatz (HOFMANN 2017). MEMS sind i. d. R. preisgünstiger, liefern aber vergleichsweise geringere Genauigkeiten (DEPPE et al. 2017). Für Anwendungen mit hohen Ansprüchen an Genauigkeit und zeitliche Stabilität werden daher eher FOG und RLG genutzt, wobei FOG basierte Systeme im Bereich Vermessung derzeit dominieren.

Auch für Beschleunigungssensoren existieren verschiedene Bauarten. Es können grundlegend mechanische oder halbleiterbasierte Akzelerometer unterschieden werden (WOODMAN 2007). Ein mechanischer Beschleunigungsmesser besteht aus einer an Federn aufgehängten Masse. Die Verschiebung der Masse wird mit einem Wegaufnehmer gemessen, der ein Signal liefert, das proportional zu der Kraft ist, die auf die Masse in Richtung der Eingangsachse wirkt. Mithilfe des zweiten Newtonschen Gesetzes kann daraus die wirkende Beschleunigung berechnet werden. Zu den halbleiterbasierten Beschleunigungsmessern zählen unter anderem piezoelektrische Akzelerometer oder MEMS-Sensoren. Eine detaillierte Beschreibung des Funktionsprinzips von MEMS-Beschleunigungssensoren und MEMS-Drehratensensoren ist bei WILD-PFEIFFER (2015) zu finden.

Der Begriff Strapdown-System beschreibt, dass die Trägheitssensoren der IMU fest auf einer sich bewegenden Plattform befestigt sind. Gyroskope dienen in einem Strapdown-System nicht der Stabilisierung einer Sensorplattform, wie beispielsweise bei stabilisierten Trägheitssystemen, sondern aus ihnen wird direkt durch Integration der Drehraten die Orientierung abgeleitet (siehe Abbildung 30). Zur Positionsbestimmung werden die Signale der Akzelerometer zusammen mit den integrierten Signalen der Gyroskope um den Breitengradabhängigen Wert der Erdbeschleunigung korrigiert und abschließend zweifach integriert (WOODMAN 2007).

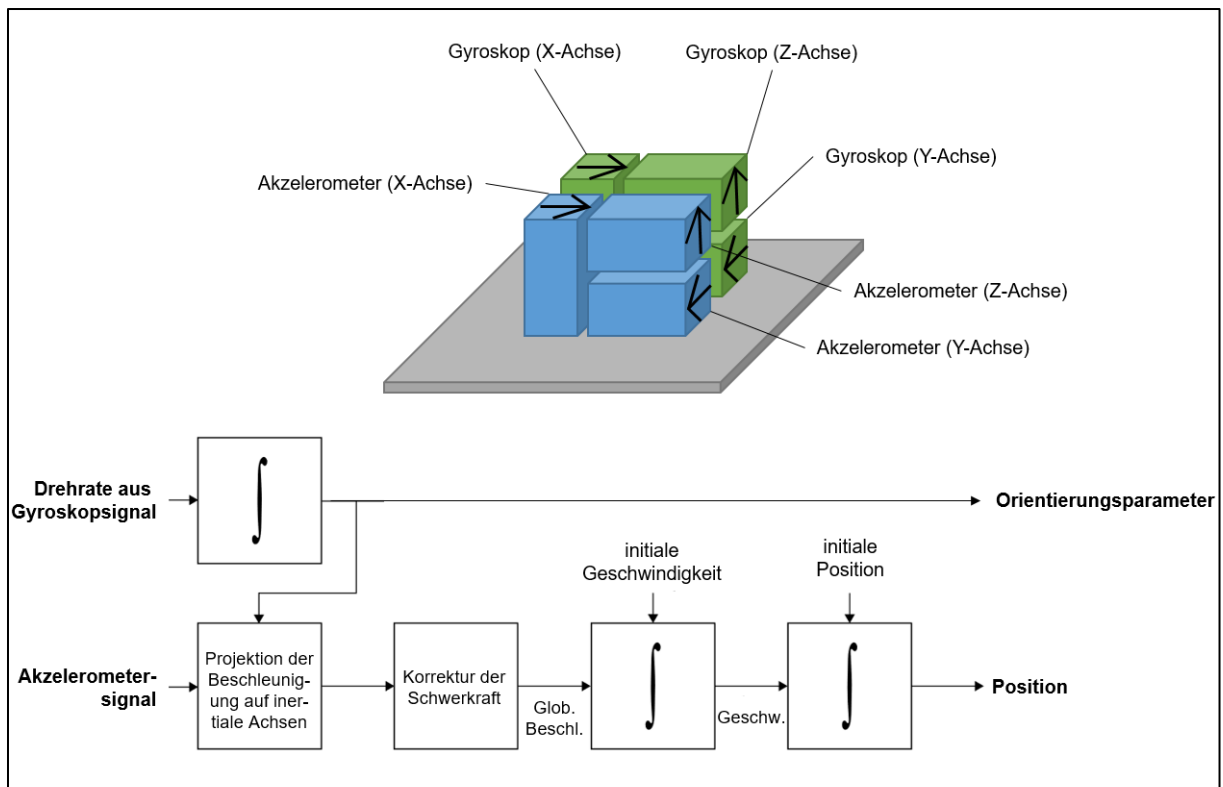


Abbildung 30: Aufbau und inertialer Navigationsalgorithmus von Strapdown-Systemen (nach WOODMAN 2007).

Betrachtet man eine kurze Zeitspanne, dann liefert eine für Mobile-Mapping-Anwendungen typische IMU mit faseroptischem Kreiselsystem und einer Datenrate von 200 Hz sehr genaue Orientierungs-, Beschleunigungs- und Positionsinformationen (KUKKO 2013). Über eine längere Zeitspanne jedoch wirkt sich der Einfluss der sensorspezifischen Drift negativ auf die Orientierungs- und Positionsgenauigkeit aus. Dieser Driftfehler kann, wie in Abbildung 31 schematisch dargestellt, mit Hilfe der GNSS-Daten geschätzt und korrigiert werden.

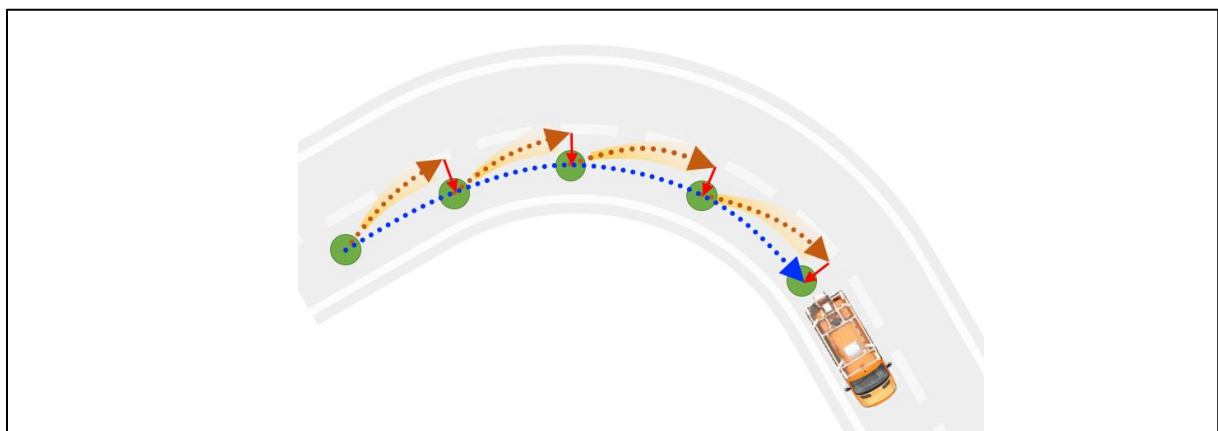


Abbildung 31: Prinzipdarstellung der Korrektur (rot) von driftenden IMU-Daten (orange) durch GNSS-Messungen (grün) zu einer integrierten Trajektorie (blau) in der Draufsicht.

Die Schätzung der Fehler des INS wird typischerweise mit einem rekursiven Bayes'schen Schätzverfahren, dem Extended Kalman Filter Verfahren, realisiert (GROVES 2013). Dabei werden optimale Schätzungen des Systemzustands sowie Genauigkeiten dieser Schätzungen (Schätzfehler) berechnet (siehe Abbildung 32). Beim Kalman Filter werden die Fehler der Beschleunigungs- und Drehratensensoren mittels einer Reihe von linearisierten und diskretisierten Differentialgleichungen modelliert (KLINGBEIL & ROMANOVAS 2012). Die Differenzen zwischen der Messungen des INS und der Messungen des GNSS werden im Kalman Filter verarbeitet, um den langsam wachsenden Positionsfehler des INS zu schätzen.

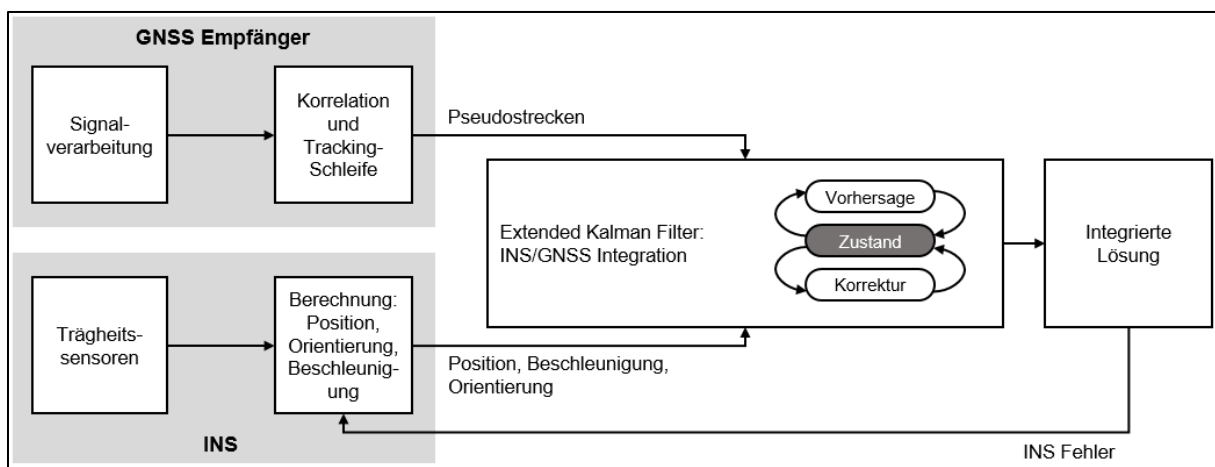


Abbildung 32: Aufbau eines integrierten INS/GNSS-Positionierungssystems mit Closed-Loop-Integrationsstrategie (nach ZHOU et al. 2017).

Der Algorithmus des Kalman Filters besteht dabei im Wesentlichen aus den Schritten Vorhersage und Korrektur, die für jede neue Beobachtung des Systemzustands durchgeführt werden. Der Vorhersageschritt beinhaltet die Schätzung eines zukünftigen Systemzustands mit einem vermuteten Schätzfehler. Im Korrekturschritt wird die Zustandschätzung mit den Beobachtungen korrigiert und auch die Fehlerkovarianz für die nachfolgende Iteration aktualisiert. Die sogenannte Closed-Loop-Integrationsstrategie sieht vor, dass die vom Kalman Filter geschätzten Fehler an das INS zurückgeführt und für die nächste Kalman Filter Iteration als Korrektur angebracht werden. Eine detaillierte Herleitung des Kalman Filters im Kontext der Positionsbestimmung ist bei THRUN, BURGARD & FOX (2005) zu finden.

Als Ergebnis liefert ein inertiales Positionierungssystem die sogenannte Trajektorie des Mobile-Mapping-Systems. Sie ist beschrieben durch die Positionen X, Y, und Z im übergeordneten Koordinatensystem sowie die Orientierung der Messplattform mithilfe der achsenbezogenen Neigungswinkel Roll, Pitch und Heading zu definierten Zeitpunkten der Messung (Zeitstempel).

2.2.5 Georeferenzierung von Sekundärsensordaten

Die Sekundärsensoren eines Mobile-Mapping-Systems, zu denen u. a. auch Laserscanner zählen, arbeiten unabhängig voneinander in ihren eigenen, sensorspezifischen Koordinatensystemen (KUKKO 2013).

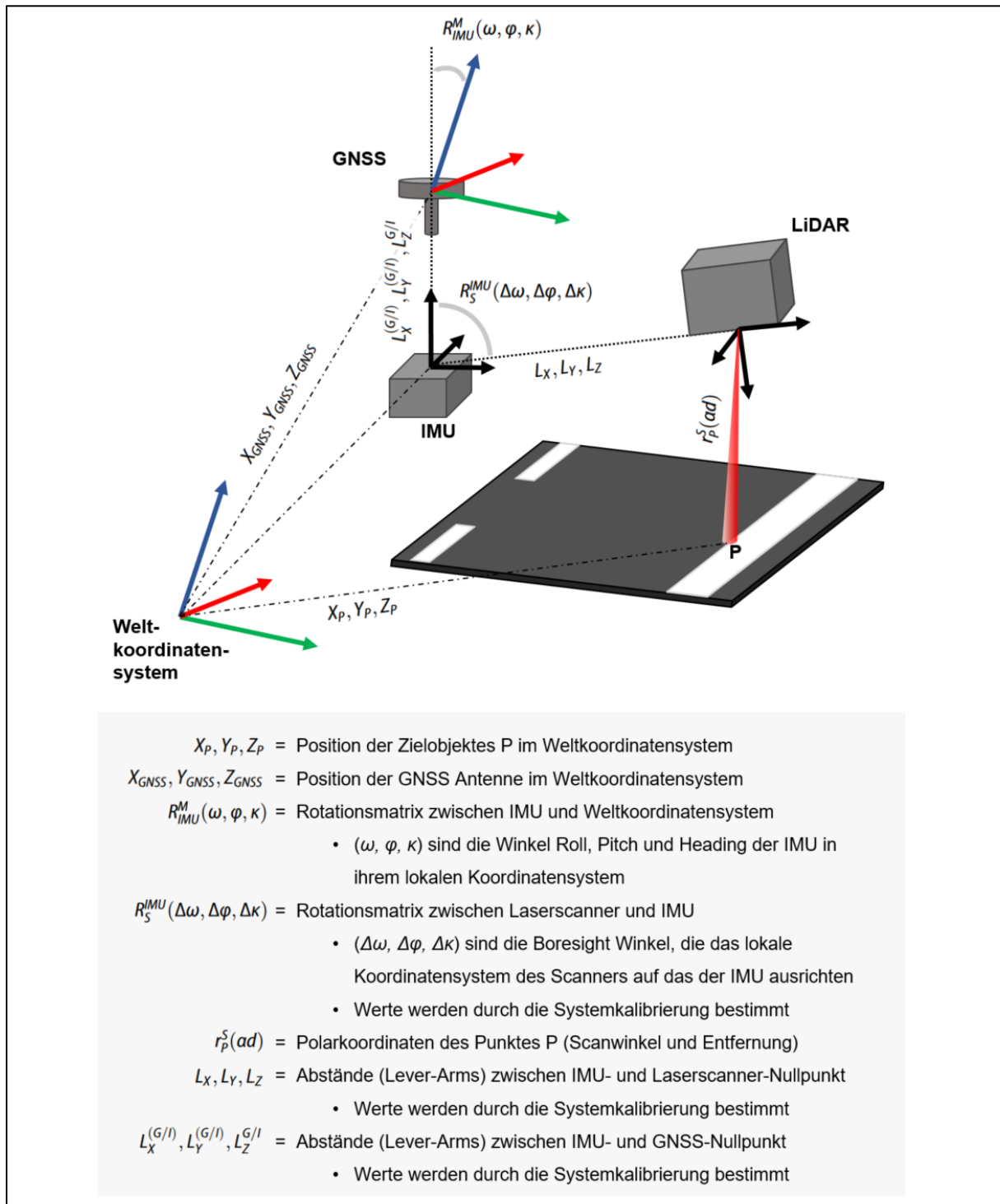


Abbildung 33: Skizze der Transformationen zwischen den an der direkten Georeferenzierung beteiligten Koordinatensystemen (nach GUAN et al. 2016).

Jede Messung wird innerhalb dieses Sensorkoordinatensystems definiert und ist zunächst nicht direkt mit den Weltkoordinaten verbunden. Alle Sekundärsensoren sind an das Koordinatensystem der Messplattform gebunden. Dieser Koordinatenrahmen wird dadurch definiert, wie die IMU angebracht ist (Strapdown Konzept). Es gibt mehrere lokale Koordinatensysteme, die über Transformationen zueinander in Beziehung stehen (siehe Abbildung 33). Die Bestimmung dieser Transformationen ist Aufgabe eines als Boresight Alignment bezeichneten Kalibrierungsprozesses (RIEGER et al. 2010).

Typischerweise wird die räumliche Ähnlichkeitstransformation bzw. die 3D-Helmerttransformation genutzt, um Koordinatensysteme aufeinander abzubilden (KAHMEN 2011). Die sieben Parameter der 3D-Helmerttransformation umfassen drei Translationen (X_0, Y_0, Z_0) zwischen den Koordinatensystemursprüngen, drei Rotationen (ω, φ, κ) um die Achsen des Zielkoordinatensystems sowie einen Massstabsfaktor.

Wie bei GUAN et al. (2016) in Bezug auf die in Abbildung 33 dargestellte Skizze beschrieben, können die 3D-Koordinaten eines Objektpunkts P mithilfe folgender Gleichung berechnet werden:

$$\begin{bmatrix} X_P \\ Y_P \\ Z_P \end{bmatrix}^M = \begin{bmatrix} X_{GNSS} \\ Y_{GNSS} \\ Z_{GNSS} \end{bmatrix}^M + R_{IMU}^M(\omega, \varphi, \kappa) \cdot \left(R_S^{IMU}(\Delta\omega, \Delta\varphi, \Delta\kappa) \cdot r_P^S(ad) + \begin{bmatrix} L_X \\ L_Y \\ L_Z \end{bmatrix}_S^{IMU} - \begin{bmatrix} G \\ L_X^T \\ G \\ L_Y^T \\ G \\ L_Z^T \end{bmatrix}_{GNSS}^{IMU} \right) \quad (1)$$

Zur Georeferenzierung der Laserscannermessungen muss also zum einen die Position und Orientierung des Laserscanners relativ zur IMU und zum GNSS-Empfänger bekannt sein (Boresight Alignment) und zum anderen müssen die beteiligten Koordinatensysteme ineinander überführt werden (3D-Helmerttransformation). Außerdem müssen die Daten des Positionierungssystems für die Sensorintegration des Laserscanners in das Gesamtsystem und letztlich für die Bestimmung der genauen Weltkoordinaten der Messpunkte mit einem genauen Zeitstempel versehen werden. Über die Zeit und entsprechend der Bewegung der Messplattform entsteht sukzessive aus einzelnen 3D-Punkten das primäre Datenprodukt einer MLS-Aufnahme: ein georeferenzierter Laserscan in Form einer 3D-Punktwolke.

2.2.6 Fehlereinflüsse

In den vorstehenden Kapiteln wurde dargestellt, dass beim mobilen Laserscanning komplexe Systeme zum Einsatz kommen und eine Vielzahl unterschiedlicher Sensoren in den jeweiligen Subkomponenten beteiligt sind. Jede Messung ist dabei in gewisser Weise fehlerbehaftet und kann die Qualität der 3D-Punktwolke beeinflussen, die als Ergebnis der Synchronisation von Laserscanner und inertialem Positionierungssystem entsteht. Verschiedene Fehlerquellen tragen also unterschiedlich stark zum Gesamtfehlerhaushalt des Laserscans bei. Die Messfehler auf Seiten des GNSS haben den quantitativ größten Einfluss und betreffen vor allem die absolute Positionsgenauigkeit eines Laserscans. Die Fehlereinflüsse aus IMU und Scanner beeinflussen dagegen eher die Genauigkeit relativer Messungen innerhalb einer 3D-Punktwolke.

Im Folgenden sind die wichtigsten Fehlereinflüsse der am Laserscanning beteiligten Systemkomponenten überblicksartig beschrieben. Eine detaillierte Analyse von Fehlereinflüssen luft- und bodengestützter Laserscanning-Systeme ist bei GLENNIE (2007), GORDON (2008) und CHAZETTE et al. (2016) zu finden.

2.2.6.1 Laserscanner

Bei den Unsicherheiten von LiDAR-Messungen handelt es sich zum einen um innere Einflussfaktoren auf der Ebene der Erfassungskette und zum anderen um äußere Fehlereinflüsse, die aus der Messumgebung resultieren (CHAZETTE et al. 2016). Die systematischen Fehlerquellen der ersten Kategorie werden meist herstellerseitig durch eine entsprechende Sensorkalibrierung minimiert. Bei GORDON (2008) sind die verschiedenen intrinsischen Faktoren ausführlich beschrieben. So können beispielsweise durch die Optik des Ablenksystems sogenannte Totwege des Messstrahls erzeugt werden, die dann die Distanzmessung entsprechend verfälschen. Beim Phasenvergleichsverfahren kann es außerdem zu zyklischen Phasenfehlern kommen, die u. a. aus einer zu kleinen baulichen Distanz zwischen Sender und Empfänger resultieren. Dies kann dazu führen, dass „elektromagnetische Strahlung direkt vom Sender zum Empfänger gelangen und sich mit der vom Messobjekt reflektierten Strahlung überlagern [kann]“ (GORDON 2008, S. 28).

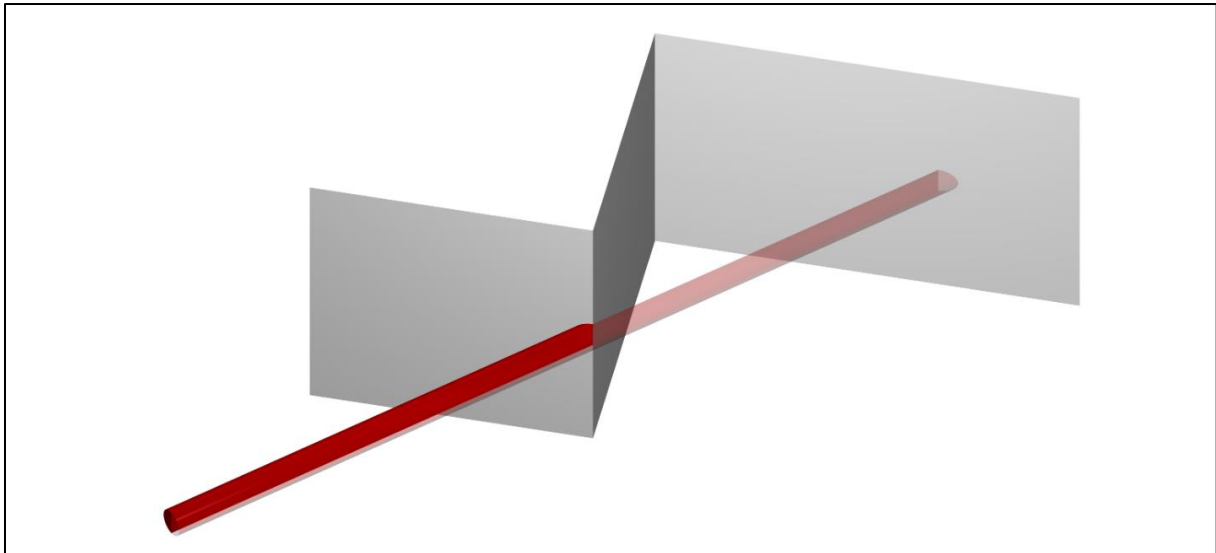


Abbildung 34: Problem der eingeschränkten Erfassbarkeit von Kanten (GORDON 2008).

Zur Kategorie der äußeren Fehlerquellen zählen beispielsweise Einflüsse der Atmosphäre, Reflexionseigenschaften der Messobjektoberfläche oder der Auftreffwinkel des Laserstrahls. Anders als beim Airborne Laserscanning, spielt bei im Nahbereich stattfindenden Messungen die atmosphärische Brechung keine nennenswerte Rolle. „Für übliche Messaufgaben liegt der Einfluss [meteorologischer Faktoren] meist unter der Genauigkeit des Messgerätes“ (GORDON 2008, S. 30). Dagegen beeinflussen meteorologisch bedingte Eigenschaften einer Oberfläche, wie beispielsweise deren Nässe infolge von Regen, die Lasermessung erheblich. Problematisch ist Oberflächenfeuchtigkeit vor allem bei Lasern, die im nahen Infrarotbereich arbeiten, da Infrarotlicht von Wasser absorbiert wird. Auch die oberflächenrauigkeitsbedingten Reflexionseigenschaften können die Lasermessung verfälschen. So werden beispielsweise in manchen Fahrbahnmarkierungen kleine Glaspartikel verwendet, in denen es zu mehrfachen Brechungen und Reflexionen kommen kann. Dies führt zu sogenannten Mehrwegeeffekten, die eine tendenziell zu lange Distanzmessung nach sich ziehen. Überwiegt die spiegelnde Reflexion, kann es im Extremfall vorkommen, dass gar keine Signalanteile zum Detektor reflektiert werden und somit Datenlücken entstehen. Auch die Form des Laserstrahls beim Auftreffen auf das Zielobjekt hat einen Einfluss auf das Messergebnis. Der Laserstrahl nimmt infolge der Strahldivergenz eine gewisse Fläche auf dem Zielobjekt ein, die als Footprint bezeichnet wird. Wie in Abbildung 34 dargestellt, kann es bei bestimmten Konstellationen dazu kommen, dass Teile des Footprints auf einem Objekt im Vordergrund und Teile auf einem Objekt im Hintergrund liegen. In der Folge entsteht ein Mischsignal, das eine Entfernung zwischen den beiden Objekten repräsentiert. Dadurch ergibt sich ein ausrundender Effekt an Kanten.

Laser system	
Laser class	1 (according to EN60825-1 / ANSI Z136.1)
Beam divergence	< 0.5 mrad (half angle, 1/e ²)
Beam diameter	Approx. 1.9 mm (at 0.1 m distance)
Ambiguity distance	182 m (above, range reading restarts at zero)
Minimum distance	0.3 m
Range resolution	0.1 mm
Data acquisition rate	Max. 1.094 million pixel/sec.
Accuracy / Precision	
Linearity error	≤ 1 mm
Range drift (full -10° C ... +45° C)	< 0.3 mm (with internal reference), < 2 mm (without ref.)
Target Distance	White (80%) ¹ Grey (37%) ¹ Black (14%) ¹
1 Sigma Range Noise, 1 m	0.2 mm 0.3 mm 0.4 mm
1 Sigma Range Noise, 2 m	0.2 mm 0.2 mm 0.3 mm
1 Sigma Range Noise, 5 m	0.2 mm 0.2 mm 0.3 mm
1 Sigma Range Noise, 10 m	0.2 mm 0.2 mm 0.3 mm
1 Sigma Range Noise, 25 m	0.2 mm 0.3 mm 0.5 mm
1 Sigma Range Noise, 50 m	0.4 mm 0.6 mm 1.2 mm

Abbildung 35: Auszug aus dem Datenblatt eines Laserscanners, der zum Einsatz auf MLS-Systemen konzipiert ist (Z+F 2021).

Auch ein flacher Auftreffwinkel ist ein Einflussfaktor. Trifft ein Laserstrahl auf eine schräge Fläche oder trifft der Laserstrahl in einem sehr flachen Winkel auf eine Oberfläche, dann dominieren die räumlich näher gelegenen Anteile des reflektierten Signals, was zu einer tendenziell zu kurzen Distanzmessung führt (KERN 2003).

Typischerweise geben Hersteller von Laserscannern zur Fehlerabschätzung die Genauigkeit der Entfernungsmessung in ihren Spezifikationen an. Meistens werden hierbei Root Mean Square (RMS) Werte für bestimmte Reflektivitätsniveaus bei bestimmten Entfernungen zum Zielobjekt aufgeführt. Beispielhaft ist in Abbildung 35 ein Auszug aus dem Datenblatt des Z+F Profilers 9020 dargestellt. Die meisten derzeit in MLS-Systemen eingesetzten Laserscanner bestimmen die Entfernung auf 1 mm bis 10 mm genau. Submillimetergenaue LiDAR-Systeme sind bislang noch die Ausnahme.

2.2.6.2 Satellitengestützte Positionsbestimmung

Die Fehlereinflüsse, die auf die satellitengestützte Positionsbestimmung wirken, lassen sich in drei Gruppen untergliedern: Satellitenfehler, Fehler in der Signalausbreitung und Empfängerfehler (HOFMANN 2017). In Tabelle 5 sind acht Hauptfehlerursachen und ihre Größenordnung aufgeführt.

Tabelle 5: Fehlereinflüsse der satellitengestützten Positionsbestimmung im C/A-Code (nach BAUER 2017).

Kategorie	Fehlerursache	Größenordnung
Satellitenfehler	Bahnfehler	ca. 5 m
	Uhrenfehler	ca. 1 m
Fehler in der Signalausbreitung	Ionosphäre	Zentimeter bis Dezimeter
	Troposphäre	Dezimeter
	Multipath-Effekte	ca. 5 m
Empfängerfehler	Rauschen	ca. 1 bis 10 m
	Phasenzentrumsvariationen	Millimeter bis Meter
	Gerätebedingte Latenzen	Meter

Von besonderer Bedeutung für das Mobile Mapping sind neben Bahn- und Uhrzeitfehler die Laufzeitfehler durch Mehrwegeausbreitung (Multipath-Effekte) und Abschattungen. Abschattungen entstehen, wenn Satellitensignale an einem bestimmten Ort vom Empfänger nicht registriert werden können, weil sich Objekte, wie beispielsweise Häuser, im Signalweg befinden (siehe Abbildung 36). Multipath-Effekte dagegen entstehen, wenn ein Satellitensignal mehrfach vom Empfänger registriert wird, weil es beispielsweise an Häuserfassaden reflektiert wird. Multipath-Effekte treten vor allem in dicht und hoch bebautem Gebiet auf, insbesondere wenn das Fassadenmaterial spiegelnd ist.

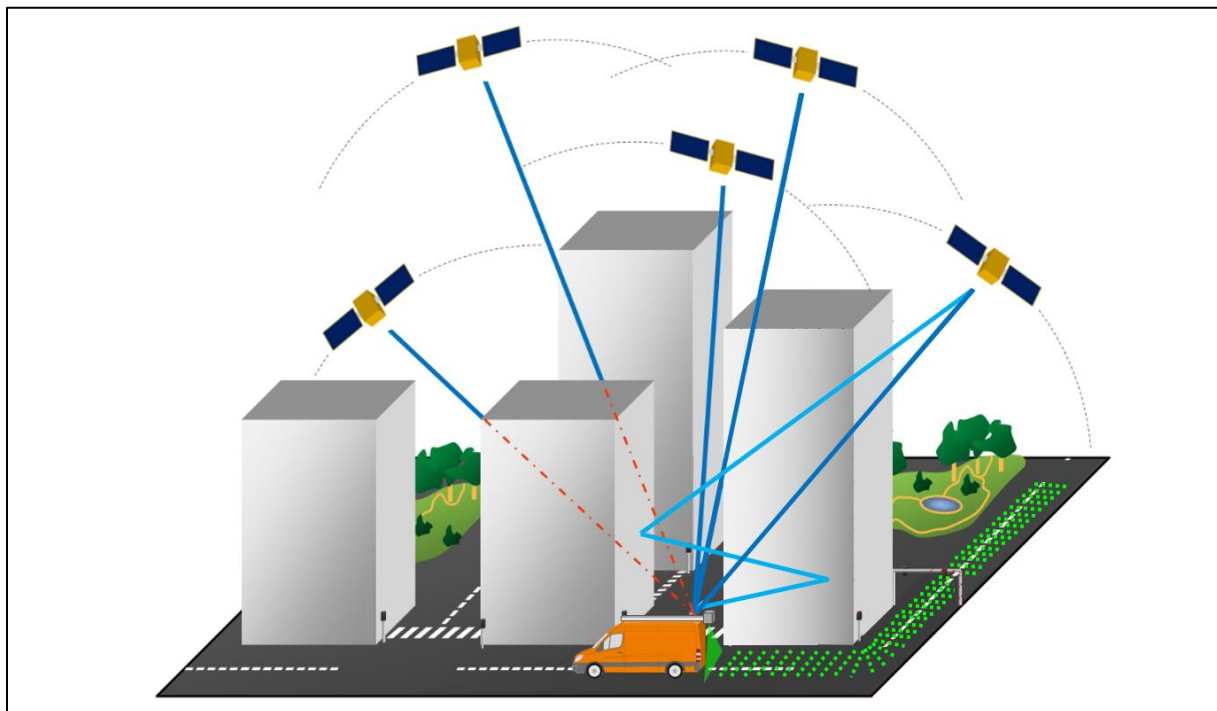


Abbildung 36: Prinzipdarstellung von GNSS-Fehlerquellen. Dargestellt sind Abschattungen durch Häuser im Signalweg (rot) und die Mehrwegeausbreitung (hellblau) aufgrund von Reflexionen des Satellitensignals an Häuserfassaden (nach KUKKO et al. 2012).

Die Auswirkung der genannten Faktoren auf die Positionsbestimmung im Kontext von Mobile-Mapping-Anwendungen liegt im Bereich von wenigen Millimetern bis zu mehreren Metern. Das Ausmaß des Gesamtfehlers ist dabei abhängig von der jeweils vorliegenden geometrischen Konstellation von Satelliten und Empfänger. „Die Fehlereinflüsse auf die Beobachtungen können mittels [...] DGNSS unterschiedlich gut reduziert werden: Satellitenuhrfehler werden vollständig eliminiert, Satellitenbahnfehler sowie ionosphärische und troposphärische Einflüsse können verringert werden“ (HOFMANN 2017, S. 17).

Multipath-Effekte können nur schwer mithilfe von DGNSS kompensiert werden. Unter Nutzung einer RTK-Korrektur ist es aber möglich, diesen Fehlereinfluss zu reduzieren.

2.2.6.3 Inertialsystem

Die drei wesentlichen Fehlerquellen der IMU sind der Biasfehler, der Skalenfaktorfehler und das Messrauschen (GROVES 2013). Der Biasfehler, gemessen in Grad pro Stunde, ist ein systematischer Fehler und bezeichnet die Drehrate bzw. die Beschleunigung, die Gyroskope bzw. Akzelerometer in Ruhe ausgeben. Temperaturänderungen und Stoßeinwirkungen haben hierbei besonderen Einfluss. Auch der Skalenfaktorfehler gehört in die Kategorie der systematischen Fehler und quantifiziert den Fehler des Verhältnisses zwischen Eingangs- und Ausgangssignal. Er wird typischerweise in Parts Per Million (ppm) oder Prozent angegeben. Das Messrauschen bezeichnet im Allgemeinen jede zufällige Schwankung eines gemessenen Outputs, wenn ein Sensor einem konstanten Input bei konstanten Bedingungen ausgesetzt ist. Es wird i. d. R. in $^{\circ}/\sqrt{h}$ angegeben. Wird ein verrauschtes Signal integriert, beispielsweise um aus einem Drehratensignal einen Winkel abzuleiten, kommt es über die Zeit zu einem zufälligen Anwachsen von Fehlern. Diese Drift wird als Random Walk bezeichnet. Die Größenordnung der IMU-spezifischen Fehler hängt maßgeblich vom Qualitätsniveau der genutzten Komponenten ab. Dieses wiederum richtet sich vor allem nach dem Kalibrierungsaufwand, der herstellerseitig zur Kompensation der Fehlerquellen geleistet wird. Während beispielsweise Biasfehler von Gyroskopen aus der Kategorie Strategic Grade im Bereich $0,0001^{\circ}/h$ liegen, sind bei Tactical-Grade-Systemen Biasfehler von mehr als $0,1^{\circ}/h$ möglich (JEKELI 2001). In Tabelle 6 sind typische Werte IMU-spezifischer Fehler aufgeführt. Als Ergebnis ihrer Studie geben ZHANG, NIU & SHI (2020) den Einfluss von IMU-Fehlern auf die Positionsbestimmung im Mobile-Mapping-Bereich mit ca. 1 cm an.

Tabelle 6: Typische Größenordnung IMU-spezifischer Fehler im Mobile-Mapping-Bereich (nach JEKELI 2001; ZHANG, NIU & SHI 2020).

Fehler	Größenordnung bei Navigation Grade Sensorik
Biasfehler	0,005 °/h
Skalenfaktorfehler	10 ppm
Messrauschen (White Noise)	0,002 °/√h

2.3 Digitale Oberflächenmodellanalyse in Geo-Informationssystemen

Im Bereich der Geowissenschaften spielen Geo-Informationssysteme (GIS) eine zentrale Rolle bei der quantitativen Analyse und Modellierung von räumlichen Phänomenen (BILL 2016). Sie stellen eine Art von Informations- und Kommunikationssystemen dar und können wie folgt definiert werden können: „Ein Geo-Informationssystem [...] ist ein rechnergestütztes System, das aus Hardware, Software und Daten besteht und mit dem sich raumbezogene Problemstellungen in unterschiedlichsten Anwendungsgebieten modellieren und bearbeiten lassen. Die dafür benötigten raumbezogenen Daten/Informationen können modelliert, digital erfasst und redigiert, verwaltet, reorganisiert und verteilt, analysiert sowie alphanumerisch und graphisch präsentiert/visualisiert werden“ (BILL 2016, S. 8). Dementsprechend kann unter dem Begriff GIS eine Software zur Erfassung, Verwaltung, Analyse und Ausgabe von Geodaten verstanden werden, welche sich grundlegend in Vektor- und Rasterdaten unterscheiden lassen.

Auch wenn moderne GIS-Softwareprodukte Möglichkeiten zur direkten Integration von dreidimensionalen Daten bieten, werden Laserscandaten in GIS-Anwendungen klassischerweise in Form von vektorbasierten oder gerasterten Oberflächenmodellen verwendet. Aus diesem Grund werden in den folgenden Unterkapiteln verschiedene Arten von Oberflächenmodellen, räumliche Interpolationsverfahren zur Generierung von Oberflächenmodellen sowie die Wissenschaftsdisziplin vorgestellt, die sich mit der digitalen Analyse von Oberflächenmodellen befasst – die Geomorphometrie.

2.3.1 Oberflächenmodellierung

Das primäre Datenprodukt einer MLS-Kampagne ist eine 3D-Punktwolke des Messobjektes. Obwohl moderne Laserscanner in der Lage sind, sehr hohe Genauigkeiten, Messraten und Punktdichten zu erreichen, so ist doch die unorganisierte räumliche Struktur dieser Daten ein Nachteil. Für viele Endanwendungen und Analysen ist ein solches Datenformat nicht geeignet. So werden beispielsweise im Zuge der AsBuilt-Modellierung von Gebäuden Fassadenelemente

auf Basis von Laserscans semi-automatisch nachmodelliert – die Punktwolken selbst finden oft keine weitere Verwendung (FRITSCH 2015; BLANKENBACH 2017). Auch fahrdynamische Simulationen aus dem Automotive-Bereich nutzen Laserscans von Straßen nur indirekt über spezielle Modelle, wie beispielsweise das OpenCRG Format (RAUH & MÖSSNER-BEIGEL 2008; POTÓ, CSEPINSZKY & BARSÍ 2018).

Unter dem Begriff Oberflächenmodellierung wird im Kontext dieser Arbeit das Überführen von unorganisierten 3D-Punktwolken in ein strukturiertes 2,5D-Datenformat verstanden. Diese Art der geometrischen Modellierung bildet die Schnittstelle zwischen 3D und 2D, da im GIS die Höhe meist als thematisches Attribut einer Lageposition geführt wird (BILL 2016). Die Bezeichnung 2,5D bringt dabei zum Ausdruck, dass hier keine vollständige Integration aller räumlichen Dimensionen realisiert ist, sondern jeder Lageposition nur exakt eine Attributausprägung zugeordnet werden kann. Ein klassisches Beispiel für 2,5D-Daten im Bereich der Geowissenschaften sind Digitale Höhenmodelle. Diese können laut BILL (2016, S. 376) wie folgt definiert werden: „Als Digitales Höhenmodell (DHM) bezeichnet man die Menge der digital gespeicherten Höhenwerte, die als Funktion der Lage der Punkte die Höhenstruktur des Objekts hinreichend repräsentieren.“ In der englischsprachigen Literatur sind die Begriffe digital elevation model (DEM) und digital terrain model (DTM) gebräuchlich und entsprechen den im deutschsprachigen Raum oft synonym verwendeten Begriffen DHM und Digitales Geländemodell (DGM).

Über die reine 3D-Information hinaus kann eine MLS-Punktwolke auch zusätzliche Daten enthalten, wie beispielsweise Informationen zur Reflexionsintensität, zum Scanwinkel, zur Entfernungsmessung oder zum Messzeitpunkt. Auch der im Rahmen der vorliegenden Arbeit verwendete Laserscanner (siehe Kapitel 3.1.1.2) ermöglicht die Nutzung von zwei Informationskanälen: Neben der Höhe als geometrische Information steht die Reflexionsintensität als radiometrische Information zur Verfügung. Da für beide Datenkanäle organisierte 2,5D-Daten erzeugt werden sollen, wird im Folgenden der allgemeinere Begriff des Oberflächenmodells dem Geländemodell Vorzug gegeben.

In den folgenden beiden Unterpunkten werden zunächst die grundlegenden Arten von Oberflächenmodellen beschrieben und im Anschluss daran Algorithmen, die es ermöglichen, aus diskreten Punkten Oberflächen zu generieren.

2.3.1.1 Arten von Oberflächenmodellen

Im Bereich der Geowissenschaften werden zwei Arten von Oberflächenmodellen unterschieden: Modelle auf Gitterbasis und Modelle auf Dreiecksbasis (HEYWOOD, CORNELIUS & CARVER 2006; HENGL & EVANS 2009; BILL 2016). Vereinzelt werden Isolinien als dritte Art aufgeführt (WATSON 1995), jedoch handelt es sich dabei eher um einer Art der Geländevisualisierung, auch wenn die Modellierung von Oberflächen auf Basis von Isolinien grundsätzlich möglich ist (WISE 1998). Im Folgenden werden Dreiecks- und Gittermodelle überblicksartig vorgestellt. Auch wenn dabei i. d. R. Höhenwerte als zu modellierende Information beschrieben werden, sind die Aussagen zu den Oberflächenmodellen prinzipiell auch auf andere Phänomene, wie beispielsweise Temperatur- oder Druckverteilungen, übertragbar (BILL 2016). Werden räumlich verteilte Messpunkte mittels Triangulation zu einer Oberfläche vermascht, die aus einem Netz unregelmäßiger Dreiecke besteht, spricht man von einem Triangulated Irregular Network (TIN) (VAN KREVELD 1997). Die gemessenen Höhen bilden dabei Knotenpunkte, die über Kanten miteinander verbunden werden. Es ist irrelevant, in welchem Muster die Messpunkte vorliegen oder ob die Messpunktdichte Schwankungen unterliegt. Sind die Knoten unregelmäßig verteilt, entstehen unterschiedlich große Dreiecksflächen. Für die Bereiche, die nicht durch einen Messpunkt beschrieben sind, wird die Höheninformation linear interpoliert (TSAI 1993). Eine Möglichkeit, Punktdaten zu einem Netz aus unregelmäßigen, zusammenhängenden und sich nicht überlappenden Dreiecksflächen zu vermaschen, ist die Delaunay-Triangulation. Sie wird in Kapitel 4.2.1 näher beschrieben. In Abbildung 37 ist auf der linken Seite ein TIN dargestellt. Zu den Vorteilen eines TIN zählen die Erhaltung der Höheninformation an den Messpunkten, der vergleichsweise einfache und überschaubare Berechnungsalgorithmus, die Möglichkeit der unmittelbaren Berücksichtigung von Bruchlinien durch die Eingabe von Zwangsseiten bei der Dreiecksbildung und die einfache Durchführbarkeit von Volumenberechnungen (KOHLSOCK 2011). Nachteilig ist, dass ein TIN mit ebenen Dreiecksfacetten per Definition diskontinuierlich ist, sich die Werte zwischen den Dreiecken also abrupt ändern können (HENGL & EVANS 2009). Das dem TIN zugrundeliegende Konzept der Dreiecksvermaschung wird nicht nur in den Geowissenschaften im Bereich der Oberflächenmodellierung, sondern auch in anderen Anwendungsdomänen, wie beispielsweise der Videokompression, genutzt (DINAS & BAÑÓN 2014).

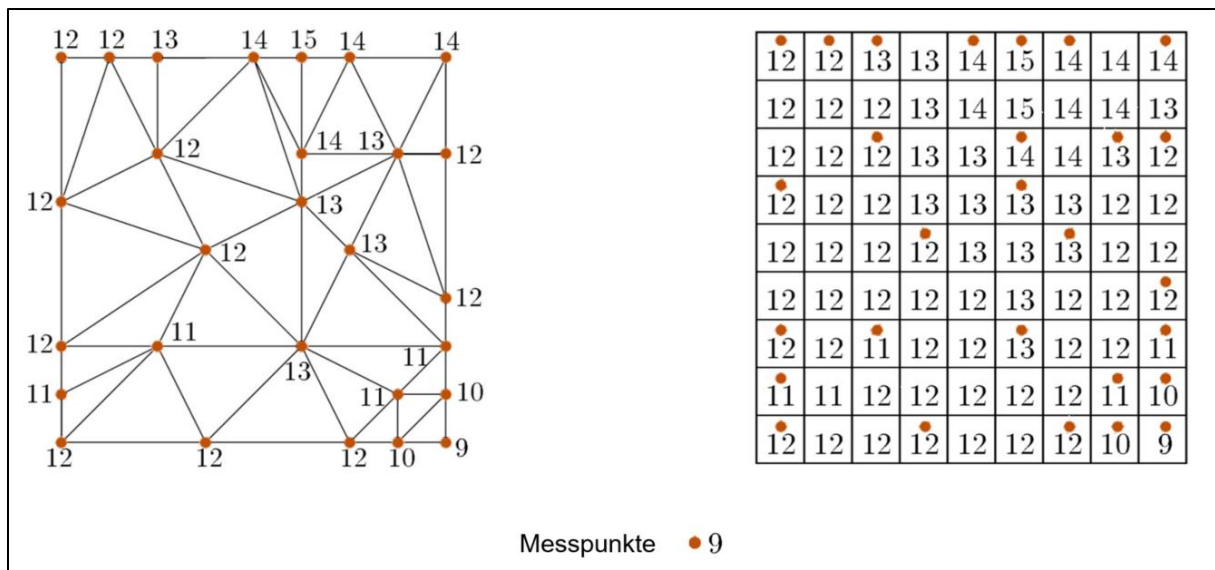


Abbildung 37: Schematische Darstellung des Grundrisses von Dreiecks- und Gittermodell im Vergleich (nach VAN KREVELD 1997).

Im Unterschied zu Dreiecksmodellen speichern Oberflächenmodelle auf Gitterbasis Höhenwerte in Form einer Matrix. Im Allgemeinen werden Gitter verwendet, deren Gitterlinien rechtwinklig zueinander verlaufen und viereckige, meist quadratische Gitterzellen bilden (siehe Abbildung 37 rechts). Abweichende Gitterarten, wie beispielsweise regelmäßige Gitter in Dreiecks- oder Hexagonalform sind ebenfalls möglich (BILL 2016). Dort, wo sich die Gitterlinien schneiden, werden die Messwerte eingesetzt. Für Gitterlinienschnittpunkte, die nicht durch einen gemessenen Höhenwert beschrieben sind, werden Höhenwerte mithilfe von Interpolationsverfahren berechnet (RASE 2021). Es gibt jedoch auch andere Interpretationen des Gitters. So können die Lagekoordinaten den Schnittpunkt der Gitterlinien definieren oder die Mittelpunkte von Gitterzellen. Softwareprodukte zur Erzeugung von Gittermodellen bieten i. d. R. eine Möglichkeit, dies zu spezifizieren. Der zentrale Vorteil eines Gittermodells ist, dass Höhenwerte kontinuierlich an jeder Position im Raster bei einer festgelegten Auflösung abgegriffen werden können (PIKE, EVANS & HENGL 2009). Ferner begünstigt die Matrixstruktur eine einfache Umsetzung von geomorphometrischen Algorithmen und nicht zuletzt die Adaption von Methoden der digitalen Bildanalyse und Mustererkennung (HENGL & EVANS 2009). Auch aus rechentechnischer Perspektive sind Modelle auf Gitterbasis vorteilhaft, da in ihnen Interpolationen einfach zu realisieren und Höhenwerte einfach zu speichern sind (BILL 2016). Die kontinuierliche Grundstruktur von Gittermodellen kann je nach Zielanwendung auch als Nachteil verstanden werden, da sie ungeachtet der Morphologie der betrachteten Oberfläche starr ist (PIKE, EVANS & HENGL 2009). Dazu gehört auch, dass die Wahl der Zellgröße eines Gittermodells die Möglichkeiten späterer Anwendungen determiniert. Bei der

Parametrisierung ist die Kenntnis hilfreich, wofür das Oberflächenmodell später verwendet werden soll. Ist die Rasterzellengröße so groß gewählt, dass mehrere Messpunkte in einer Zelle liegen, ist eine Interpolationsmethode erforderlich und es können dabei im schlechtesten Fall Informationen verloren gehen (VAN KREVELD 1997). Als weiterer Nachteil der digitalen Geländedarstellung in Gitterform sei das vergleichsweise große Datenvolumen erwähnt. Die Erzeugung und Weiterverarbeitung sehr großer und gleichzeitig sehr hochauflöster Gitteroberflächenmodelle stellt erhöhte Anforderungen an die Rechentechnik.

Eine Sonderform gerasterter Oberflächenmodelle stellt das sogenannte Curved Regular Grid (CRG) dar, welches vorrangig zur Straßenoberflächenmodellierung eingesetzt wird (POTÓ, CSEPINSZKY & BARSÍ 2018; SU et al. 2019). Das CRG-Format wurde von der Daimler AG eingeführt und intern für Simulationsanwendungen wie Reifen-, Schwingungs- und Fahrsimulation verwendet (RAUH & MÖSSNER-BEIGEL 2008). Die Besonderheit des CRG-Formats ist, dass das Höhenwertgitter an einer Straßenachse oder Fahrlinie aufgespannt wird – es ist also an der Straßengeometrie orientiert. Die Achse ist durch einen Start- und einen Endpunkt definiert, zwischen denen Achspunkte in einem festen Abstand liegen. An jedem Achspunkt wird orthogonal ein Querprofil konstruiert, auf dem wiederum die eigentlichen CRG-Knotenpunkte äquidistant erzeugt werden. Jeder CRG-Knotenpunkt ist durch seine Lage in Bezug zur Achse eindeutig definiert (siehe Abbildung 38).

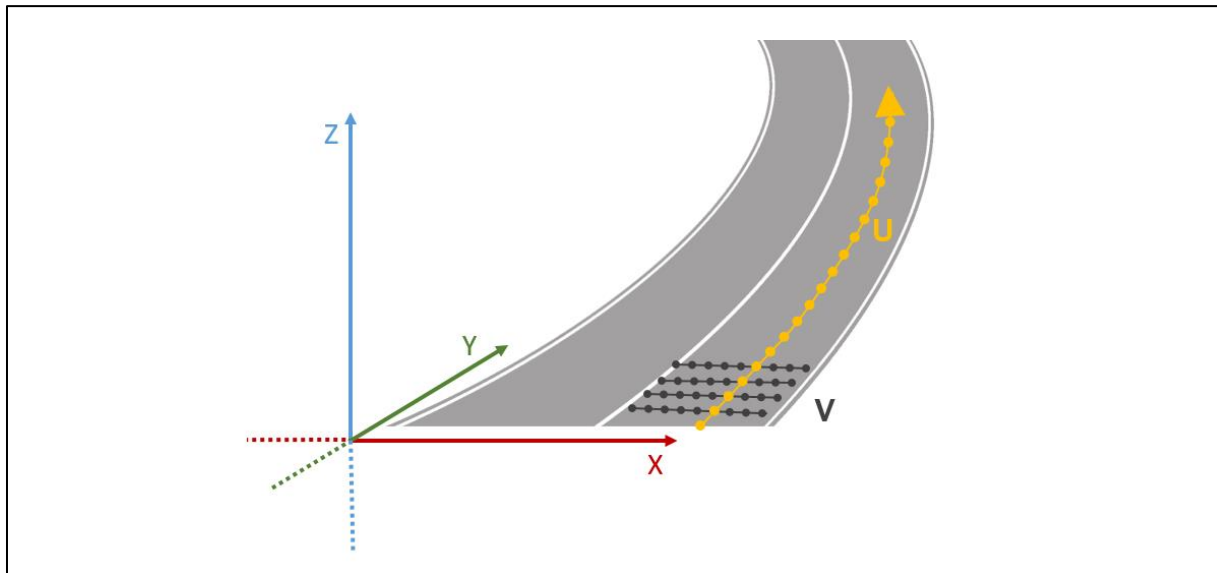


Abbildung 38: Schematischer Aufbau eines CRG aus einer Achse bzw. Referenzlinie U und daran rechtwinklig orientierten Querprofilen V.

2.3.1.2 Räumliche Interpolationsverfahren

Interpolationsmethoden werden in den Geowissenschaften eingesetzt, wenn es darum geht, aus räumlich verteilten, diskreten Messungen eine kontinuierliche Verteilung zu approximieren, um auch an nicht beprobten Stellen Aussagen über das jeweilige räumliche Phänomen machen zu können. BILL (2016, S. 521) definiert den Begriff Interpolation wie folgt: „Interpolation weist gegebenen diskreten Daten [...] eine kontinuierliche Funktion zu, die diese Daten abbildet. Die Interpolation ist stets auch eine Art von Approximation: Die betrachtete Funktion wird durch die Interpolationsfunktion in den Stützstellen oftmals exakt wiedergegeben und in den restlichen Punkten immerhin näherungsweise. Die Approximationsgüte hängt dabei vom gewählten mathematischen Ansatz wie auch von der Dichte und Verteilung der Messwerte ab.“ Räumliche Interpolationsverfahren können nach verschiedenen Kriterien typisiert werden (LAM 1983; LI & HEAP 2014). Oft werden sie in lokale oder globale, exakte oder approximierende und deterministische oder geostatistische Verfahren kategorisiert. Dieses Kapitel orientiert sich an der Taxonomie, die LI & HEAP (2014) verwenden. Dementsprechend werden im folgenden die Grundlagen zur deterministischen und geostatistischen räumlichen Interpolation beschrieben.

Verfahren der Geostatistik zeichnen sich dadurch aus, dass sie raumbezogene Daten unter Nutzung von Stochastik analysieren. Geostatistische Verfahren interpretieren sowohl die beobachteten Werte als auch die zu interpolierenden Werte als Realisationen von Zufallsvariablen, die gewisse Voraussetzungen, wie Stationarität, räumliche Autokorrelation, Normalverteilung oder Isotropie erfüllen müssen (BILL 2016). Der Geostatistik liegt dabei die Annahme zugrunde, „dass sich die räumliche Variabilität jeder Zufallsvariable Z durch die Summe von drei Komponenten erklären lässt: Einer strukturellen Komponente $m(x)$, die als Trendfläche bezeichnet werden kann, einer zufälligen Komponente $\epsilon'(x)$ und dem zufälligen Rauschen $\epsilon''(x)$ mit Mittelwert 0 und Varianz σ^2 .“ (BILL 2016, S. 536)

$$Z(x) = m(x) + \epsilon'(x) + \epsilon''(x) \quad (2)$$

Sofern die Beobachtungen einer Trendfunktion unterliegen, muss diese im Vorfeld der Interpolation geschätzt werden. Hierzu wird ein sogenanntes Variogramm bestimmt. Dabei handelt es sich um eine Darstellung der Semivarianzen $\gamma(h)$ in Bezug zum Abstand h der Autokorrelation. Abbildung 39 zeigt einen typischen Variogrammverlauf. Wie bei BILL (2016) beschrieben, markiert der Punkt c_0 den Wert, der für das in den Daten enthaltene, zufällige Rauschen angenommen wird. Mithilfe der Reichweite a wird definiert bis zu welchem Abstand

mit vorherrschender Autokorrelation der Datenwerte gerechnet wird. Ab einer gewissen Distanz kann davon ausgegangen werden, dass ein Datenpunkt keinen Einfluss mehr auf das Interpolationsergebnis hat. Dieser Punkt wird im Variogramm durch den Schwellenwert c_1 markiert. In der englischsprachigen Literatur werden für c_0 , c_1 und a die Begriffe nugget, sill und range verwendet.

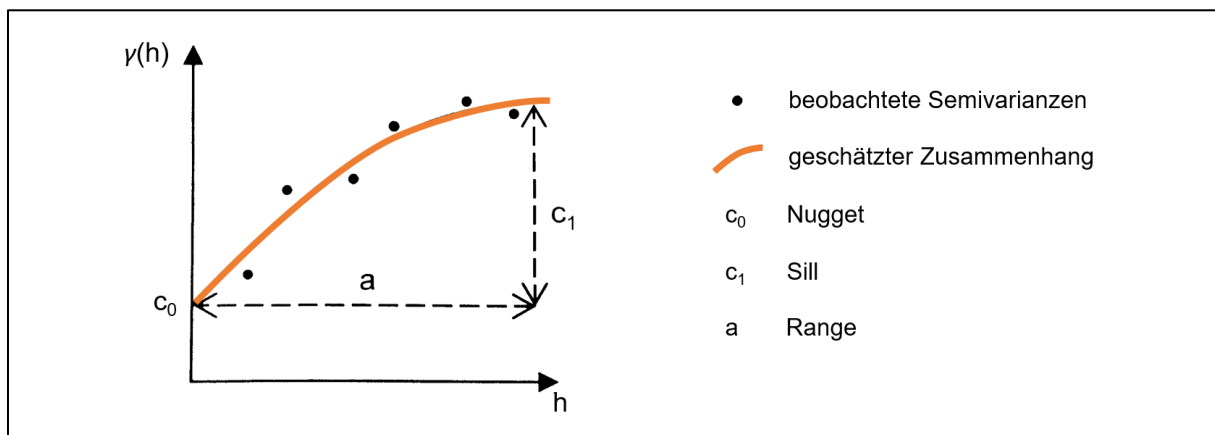


Abbildung 39: Prinzipdarstellung eines Variogramms (nach BILL 2016).

Im Rahmen einer geostatistischen Interpolation ist es also erforderlich, den Zusammenhang zwischen den beobachteten Semivarianzen durch ein Modell zu approximieren. Denkbar sind hierfür unter anderem die lineare Regression sowie gauß'sche und exponentielle Modelle. Ein bekanntes geostatistisches Interpolationsverfahren ist das Kriging. Zu den Vorteilen dieser Methode zählt, dass nicht nur eine Schätzung des Attributwertes erfolgt, sondern auch die damit verbundene Unsicherheit mittels einer Standardabweichung quantifiziert wird (TEMME et al. 2009).

In Abgrenzung zu den geostatistischen Methoden subsummieren LI & HEAP (2014) unter dem Begriff deterministische Interpolation all jene Verfahren, die das Konzept des Zufalls nicht mit einbeziehen, da sie nur Schätzungen erstellen aber keine Unsicherheiten oder Fehler im Zusammenhang mit den vorhergesagten Werten liefern.

Deterministische Verfahren können entweder Werte auf Basis der Gesamtheit der verfügbaren Datenpunkte (global) prädictieren oder bereichsweise für Punktnachbarschaften innerhalb dieser Gesamtheit (lokal). Auch können sie exakt oder approximierend sein. Dies bezeichnet die Eigenschaft, dass die aus der Interpolation resultierende Oberfläche an den Datenpunkten entweder durch diese Punkte verlaufen muss oder dass die Oberfläche von den zur Interpolation verwendeten Messungen abweichen darf. Die Nutzung approximierender Methoden ist vor allem dann sinnvoll, wenn davon ausgegangen werden kann, dass die Messpunkte verstärkt von

Unsicherheiten beeinflusst sind, die Erhaltung ihrer Werte durch den Interpolator daher nicht von vordringlichem Interesse ist.

LI & HEAP (2014) geben einen umfassenden und detaillierten Überblick über Interpolationsverfahren, die im Bereich der Geowissenschaften Einsatz finden. Im Rahmen der Untersuchungen der vorliegenden Arbeit werden ausschließlich deterministische räumliche Interpolationsverfahren verwendet (siehe Kapitel 4.2).

2.3.2 Geomorphometrie

PIKE, EVANS & HENGL (2009) definieren die Geomorphometrie als die Wissenschaft der quantitativen Landoberflächenanalyse (land-surface analysis) und als einen analytisch-kartografischen Ansatz zur Darstellung der Erdoberflächentopografie mithilfe computer-gestützter Verarbeitung von Geländehöheninformationen. Die Autoren stellen heraus, dass der Begriff Gelände (terrain) ihrem Verständnis nach unpräzise sei und schlagen daher in Abgrenzung zu WILSON & GALLANT (2000) den Begriff Landoberfläche (land-surface) vor.

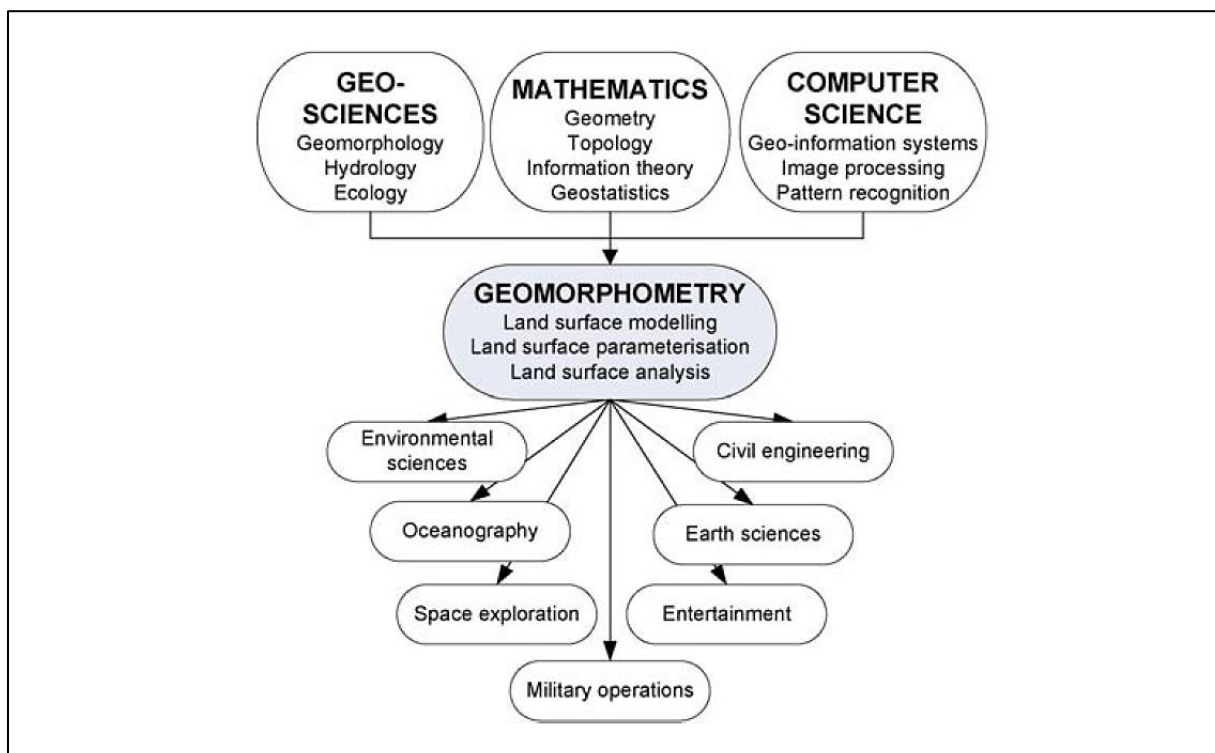


Abbildung 40: Beziehungen der Geomorphometrie zu ihren Quelldisziplinen und Anwendungsdomänen (PIKE, EVANS & HENGL 2009).

Aus der vorstehenden Definition wird deutlich, dass es sich bei der Geomorphometrie um ein interdisziplinäres Forschungsfeld handelt, das seine Wurzeln in den Geowissenschaften, der Mathematik und der Informatik hat (siehe Abbildung 40). Vor allem aus dem Bereich der digitalen Bildanalyse und Mustererkennung wurden Konzepte und Methoden adaptiert. Im folgenden werden die grundlegenden Konzepte der Geomorphometrie vorgestellt. Eine umfassende und zugleich detaillierte Auseinandersetzung mit diesem Wissenschaftszweig ist bei HENGL & REUTER (2009) zu finden.

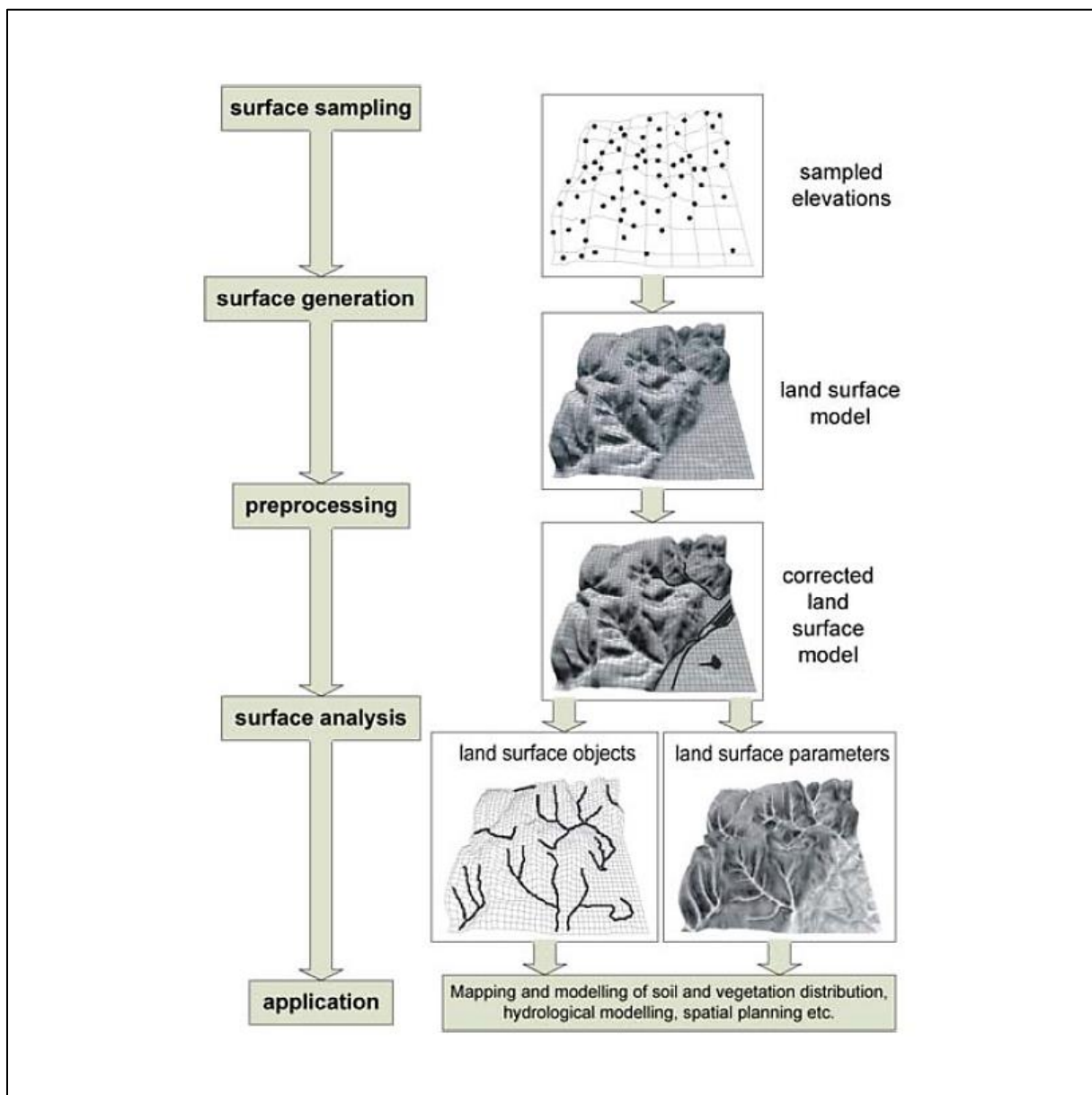


Abbildung 41: Ablaufschema einer geomorphometrischen Analyse (PIKE, EVANS & HENGL 2009).

Die Basisoperation in der Geomorphometrie ist die Extraktion von Parametern und Objekten aus digitalen Topographien. In Abbildung 41 sind die typischen Schritte einer geomorphometrischen Analyse dargestellt. Typischerweise arbeiten geomorphometrische Algorithmen auf digitalen Rasterhöhenmodellen. Da diese immer gewissen Fehlern unterliegen können, ist eine Vorverarbeitung Teil der Prozesskette. So müssen beispielsweise Höhenmodelle zunächst von Artefakten, wie beispielsweise abflusslosen Senken, bereinigt werden, bevor Abflussmodelle berechnet werden können. Eine grundlegende Restriktion der Geomorphometrie ist, dass eine Oberfläche stetig differenzierbar sein muss. Das bedeutet, dass kein Punkt in der XY-Ebene mehr als eine Höhe haben kann (2,5D-Oberflächenmodell).

Die Kernbegriffe der Geomorphometrie sind in Tabelle 7 aufgeführt. Während unter einem Landoberflächenparameter ein beschreibendes Maß für eine Oberflächenform verstanden wird (z. B. Neigung, Ausrichtung, Krümmung), so ist ein Landoberflächenobjekt ein diskretes räumliches Phänomen (z. B. Doline), das aus einem DEM extrahiert und idealerweise in eine Vektorform (Punkte, Linien oder Polygone) überführt wird. Hinsichtlich der Parameter werden primäre und sekundäre Parameter unterschieden, je nachdem, ob sie direkt aus dem Höhenmodell abgeleitet werden können oder ob sie zusätzliche Verarbeitungsschritte und Eingabeinformationen benötigen.

Tabelle 7: Wichtige Begriffe der Geomorphometrie (nach WILSON & GALLANT 2000; PIKE, EVANS & HENGL 2009).

Begriff	Bedeutung	Beispiel
Parameter	(lokales) beschreibendes Maß für eine Oberfläche - Indizes für Erosion oder Massenbewegungen	Slope
		Flow Accumulation Topographic Wetness Index
Objekt	ein diskretes räumliches Phänomen	Doline
		Schwemmfächer Drainage-Network
Element	- Sonderform von Objekt - durch topografische Diskontinuitäten begrenzter Teil der Landoberfläche, mit (relativ) einheitlicher Morphometrie	Drumlin Düne
		Landformen, wie z. B. Dünenfelder, sind aus Landformelementen aufgebaut

Morphometrische Algorithmen basieren mehrheitlich auf Nachbarschaftsoperationen (lokale Filter). Dabei wird eine kleine, regelmäßige Zell-Matrix (auch: Filter-Fenster oder Kernel) schrittweise über das Eingaberaster von der oberen linken bis zur unteren rechten Ecke bewegt.

An jeder Position auf dem Eingaberaster wird eine mathematische Operation angewendet, die durch die Werte des Filter-Kernels definiert ist und eine thematische Karte erzeugt. Dieses Verfahren wird als Moving Window bezeichnet und ist eines der grundlegenden Konzepte der digitalen Bildanalyse.

2.4 Künstliche Intelligenz, maschinelles Lernen und Deep Learning

Der Begriff Künstliche Intelligenz (KI) wird in der öffentlichen Diskussion für ein breites Spektrum an Technologien und Algorithmen verwendet. Häufig geht es dabei um Themen wie Digitalisierung, virtuelle Assistenten, die Automatisierung von Prozessen oder das Erschließen von Wissen aus großen Datenbeständen. Die teils unscharfe Verwendung des Begriffes KI ist nicht zuletzt darauf zurückzuführen, dass lange Zeit eine offizielle Definition fehlte (MONETT, LEWIS & THÓRISSON 2020). Es herrscht aber weitestgehend Konsens darüber, dass sich ein intelligentes System dadurch auszeichnet, dass es Probleme lösen kann, die nicht durch einfache Regelwerke beschreibbar sind und dass es in der Lage ist, aus Erfahrungen zu lernen (WANG 2008).

Ein Teilgebiet von KI ist das maschinelle Lernen (engl.: Machine Learning, ML). CHOLLET (2021) beschreibt das maschinelle Lernen als neues Programmierparadigma, da ML-Modelle nicht mehr explizit programmiert, sondern vielmehr trainiert werden (siehe Abbildung 42). In Abhängigkeit der verwendeten Methoden wird das maschinelle Lernen in unüberwachtes, überwachtes und bestärkendes Lernen unterteilt (ALLOGHANI et al. 2020).

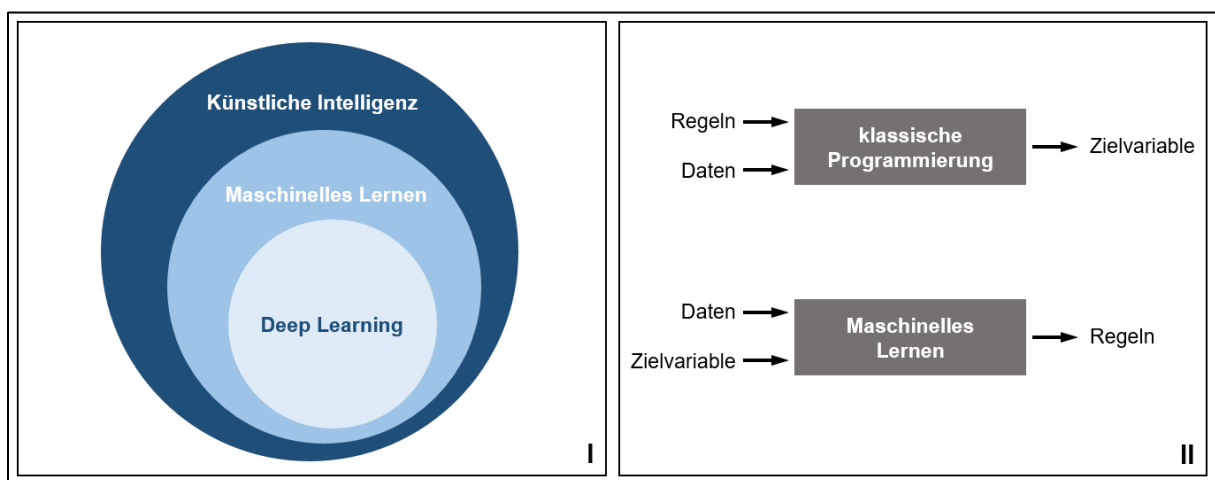


Abbildung 42: Schematische Darstellung der Einbettung von Deep Learning in den Themenkomplex Künstliche Intelligenz (I). Rechts ist der Paradigmenwechsel von klassischer Programmierung zu maschinellem Lernen dargestellt (II nach CHOLLET 2021).

Das Ziel von unüberwachten Lernmethoden ist es, in unstrukturierten Daten Zusammenhänge, Strukturen oder Muster zu erkennen und sie somit für Menschen leichter interpretierbar zu machen (WROBEL, JOACHIMS & MORIK 2012). Beispiele hierfür sind der K-Means-Algorithmus zum Clustering von Daten oder die Hauptkomponentenanalyse zur Dimensionsreduktion multidimensionaler Information. Als Eingabe für diese unüberwachten Algorithmen werden nur die Daten selbst benötigt.

Überwachte Lernverfahren untergliedern sich, je nachdem, ob die zu prognostizierende Zielvariable eine reelle Zahl oder ein diskreter Wert ist, in Regressions- und Klassifikationsansätze. Um ein Klassifikationsmodell überwacht trainieren zu können, müssen den Daten zuvor Beschreibungen zugewiesen werden. Die in der englischsprachigen Fachliteratur als Targets, Labels oder Annotations bezeichneten Beschreibungen kodieren im Fall von Klassifikationsproblemen die jeweilige Klassenzugehörigkeit eines Datenpunktes mithilfe eines diskreten Zahlenwertes. Eingabedatenpunkt und Zielvariable zusammen bilden beim überwachten maschinellen Lernen ein sogenanntes Trainingsdatenpaar. „In der Trainingsphase wird aus den Trainingsdatenspaaren [...] ein Modell gelernt, dessen Kern eine Vorhersagefunktion ist. Die Parameter der Vorhersagefunktion werden durch Optimierung geschätzt, sodass die Funktion den Input auf den Referenzoutput abbildet“ (ROSCHER & DREES 2022, S. 17). Eine Sonderform des überwachten Lernens ist das sogenannte bestärkende Lernen. Verfahren dieser Kategorie werden typischerweise in der Robotik eingesetzt, wenn für eine gegebene Problemstellung noch kein Endergebnis definiert ist, aber jede Entscheidung des Modells dahingehend bewertet werden kann, ob sie der Zielerreichung dienlich ist oder nicht (DÖBEL et al. 2018).

Zu den Lernaufgaben, bei denen überwachtes Lernen im Kontext visueller Erkennungsaufgaben typischerweise eingesetzt wird, zählen die Klassifikation, Objekterkennung und die semantische Segmentierung, die bei ROSCHER & DREES (2022, S. 18 f.) wie folgt definiert sind: „Als Klassifikation werden alle Aufgaben bezeichnet, bei denen die Inputdaten zu einer vorher definierten Klasse zugeordnet werden. Der Erfolg der Klassifikation wird an der Übereinstimmung von Schätzung und Referenz gemessen. [...] [Die Objekterkennung, die] meist im Bereich Bildinterpretation [genutzt wird,] hat zum Ziel, Objektinstanzen zu erkennen, denen eine bestimmte Semantik zugeordnet werden kann. Die Instanzen werden in den meisten Anwendungen durch achsparallele Bounding Boxen repräsentiert. [...] Bei der semantischen Segmentierung wird jedem Datenpunkt (z. B. jedem Pixel [...]) eine vorher definierte Klasse zugeordnet, wobei benachbarte Datenpunkte gleicher Klasse ein semantisches Segment bilden.“

Deep Learning stellt einen vergleichsweise jungen Teilbereich des überwachten Lernens dar und verwendet tiefe neuronale Netze für das Training von Modellen (ROSCHER & DREES 2022). Obwohl künstliche neuronale Netze bereits seit den 1950er-Jahren in Wellen aus Stagnation und Innovation wissenschaftlich untersucht wurden, tauchten Deep Neural Networks und Deep Learning als Begriffe erst seit dem Jahr 2010 auf (CHOLLET 2021). Das “Deep“ in Deep Learning bezieht sich dabei auf die hohe Anzahl tief gestaffelter Schichten in den verwendeten neuronalen Netzwerken. Im Unterschied zu Deep-Learning-Ansätzen können konventionelle ML-Ansätze, wie beispielsweise Support Vector Machines, Random Forest oder flache neuronale Netze, in der Regel nicht direkt auf die jeweiligen Eingabedaten angewendet werden. Sie benötigen zunächst eine problemspezifische Datenvorverarbeitung und Merkmalsextraktion, deren Konzeption Expertenwissen erfordert. Verfahren aus der Gruppe des Deep Learnings hingegen benötigen diese Merkmalsextraktion nicht. Durch die Tiefe und Struktur der genutzten neuronalen Netzwerke kann auch die Merkmalsextraktion aus den Daten erlernt und damit optimal auf die Problemstellung angepasst werden. Aufgrund dieser Eigenschaft wurden mit Deep-Learning-Verfahren zahlreiche Durchbrüche in Anwendungsgebieten wie Bildanalyse, Spracherkennung oder dem autonomen Fahren erzielt, deren Resultate die der vorherigen Verfahren automatischer Mustererkennung bei weitem übertrafen (CHOLLET 2021). Innerhalb von weniger als einer Dekade hat Deep Learning eine regelrechte Revolution mit bemerkenswerten Ergebnissen für Fragestellungen maschineller Wahrnehmung ausgelöst, bei denen es um Fähigkeiten geht, die für Menschen natürlich und intuitiv erscheinen, für Maschinen aber lange Zeit schwer zu erfassen waren (CHOLLET 2021).

In den folgenden Unterkapiteln werden grundlegende Konzepte künstlicher neuronaler Netze beschrieben. Dabei liegt der Fokus darauf, die wesentlichen Architekturmerkmale und die Funktionsweise von Convolutional Neural Networks im Themenfeld der semantischen Bildsegmentierung und Deep Learning einzuführen und zu erklären. Abschließend erfolgt ein Exkurs zum Thema Geospatial Artificial Intelligence (GeoAI) um zu zeigen, wie Deep Learning als Methodik das Entstehen einer neuen Teildisziplin der Geowissenschaften beeinflusst hat.

2.4.1 Künstliche neuronale Netze

DÖRN (2018, S. 89) definiert Neuronale Netze im Allgemeinen als „informationsverarbeitende Systeme, die aus einer großen Anzahl von Neuronen (Nervenzellen) bestehen“. Natürliche oder biologische neuronale Netze bilden die Grundlage zur Verarbeitung von Reizen und steuern

damit u. a. die Wahrnehmung und Lernfähigkeit von Organismen. Abbildung 43 zeigt schematisch die Informationsverarbeitung in einem Neuron. Impulse werden an den Dendriten empfangen und zum Zellkern geleitet. Von dort wird ein Signal über das Axon zu den Synapsen transportiert, von wo es an das nächste Neuron weitergegeben wird. Vereinfacht ausgedrückt, findet Lernen in einem Nervensystem (z. B. Gehirn) durch Veränderungen in der Stärke der Verbindungen zwischen einzelnen Nervenzellen statt.

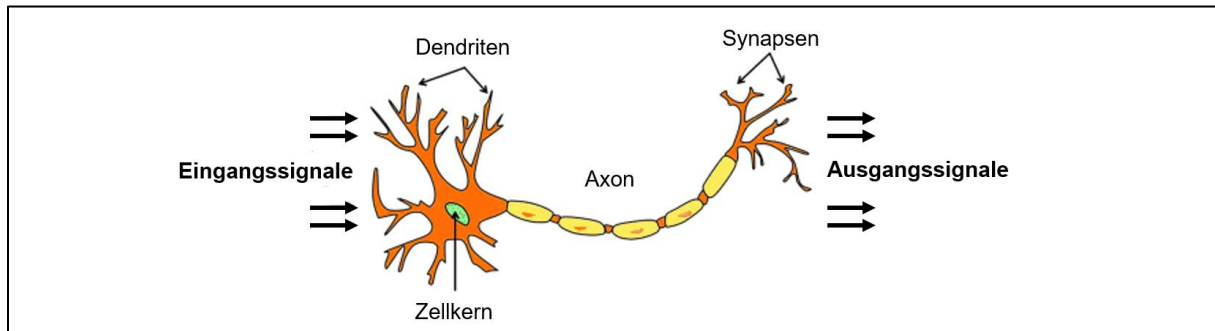


Abbildung 43: Prinzip der Übertragung von Reizen in Nervenzellen (nach DÖRN 2018).

In Anlehnung an das biologische Prinzip der Signalübertragung zwischen Nervenzellen, wird in künstlichen neuronalen Netzen eine numerische Eingabe in Form eines Signals über typischerweise mehrere Schichten in eine Ausgabe umgewandelt: Das Wissen eines künstlichen neuronalen Netzes ist in der Stärke der Verbindung zwischen einzelnen Neuronen kodiert. Diese Verbindungsstärke zwischen Neuronen aufeinanderfolgender Schichten wird in einem Trainingsprozess gelernt. Dabei werden die Fehler zwischen dem vom Netz ausgegebenen Ergebnis und der hinterlegten Label-Information auf die an der Entscheidung beteiligten Verbindungen zurückgerechnet und die Verbindungsstärke iterativ angepasst, bis die Prädiktion des Modells in Bezug auf das hinterlegte Ergebnis korrekt ist (DÖBEL et al. 2018). In den folgenden Unterkapiteln werden die grundlegenden Konzepte künstlicher neuronaler Netze erklärt und aufgezeigt, wie der Lernvorgang in einem künstlichen neuronalen Netz realisiert und beeinflusst werden kann.

2.4.1.1 Aufbau und Funktionsweise

KRUSE et al. (2015) definieren ein KNN als einen gerichteten Graph im Sinne der Graphentheorie, dessen Knoten als Neuronen und dessen Kanten als Verbindungen bezeichnet werden. Es werden im Allgemeinen drei Arten von Neuronen unterschieden: Eingabe-, Ausgabe- und versteckte Neuronen. Den Verbindungen zwischen den einzelnen Neuronen ist jeweils ein

sogenanntes Gewicht zugeordnet und jedes Neuron ist durch die Netzeingabe, Aktivierung und Ausgabe charakterisiert. In Abbildung 44 ist die Grundstruktur eines KNN mit vier Schichten dargestellt. An die Neuronen der Eingabeschicht wird eine Umgebungsinformation, beispielsweise ein Bild, angelegt. Jedes Pixel entspricht dabei einem Neuron. Die Neuronen der Ausgabeschicht wiederum geben eine Information an die Umgebung ab. Dies könnte beispielsweise eine Klasseninformation sein. Um bei dem Beispiel zu bleiben, könnten die beiden Ausgabeneuronen dafür stehen, ob das Bild ein Gesicht zeigt oder nicht. Da die übrigen Neuronen nicht direkt mit der Umgebung in Verbindung stehen, werden die entsprechenden Schichten als versteckte Schichten (engl.: Hidden Layers) bezeichnet. Sind nur wenige versteckte Schichten vorhanden, so wird das Netz als flach bezeichnet. Umfasst das Netz dagegen viele versteckte Schichten, spricht man von tiefen und ab 16 Schichten von sehr tiefen Netzen (HE et al. 2016).

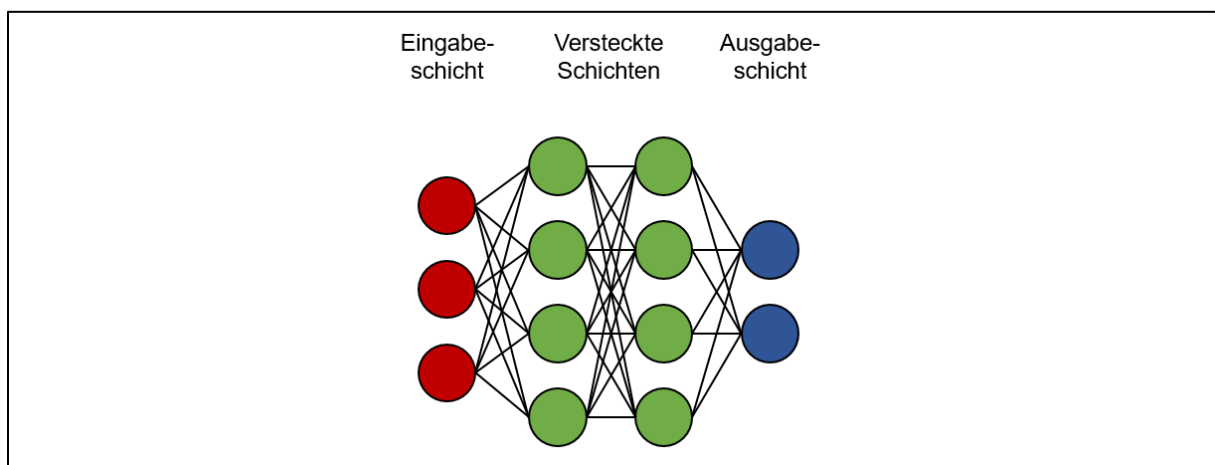


Abbildung 44: Schematische Darstellung eines künstlichen neuronalen Netzes (nach ROSCHER & DREES 2022).

In Abhängigkeit von der Art der Neuronenverbindung lassen sich zwei grundsätzliche Kategorien von künstlichen neuronalen Netzen unterscheiden (CHOLLET 2021): Feedforward Neural Network und Recurrent Neural Network. Während der interneuronale Informationsfluss in Feedforward Netzen nur in eine Richtung verläuft, sieht die Struktur von Recurrent Netzen auch eine Verbindung von Neuronen vorangegangener Schichten vor, weshalb sie auch als rückgekoppelte neuronale Netze bezeichnet werden. In der vorliegenden Arbeit werden nur Feedforward Netze betrachtet.

Abbildung 45 zeigt die innere Struktur eines einzelnen Neurons einer Schicht. Jedes Neuron einer Schicht besteht aus einer Übertragungs- und einer Aktivierungsfunktion. Zunächst werden die Eingabewerte x aus der Vorgängerschicht mit Gewichten w multipliziert und alle Produkte zu einem Ausgabewert z aufsummiert und der Biaswert hinzuaddiert. Der Bias- oder Offset-

Wert beträgt immer 1. Diese mit Koeffizienten gewichteten Eingaben gehen in die Aktivierungsfunktion f ein, wo sie integriert werden und die Ausgabe eines Neurons in Form der sogenannten Aktivierung erzeugt wird. Die Aktivierungsfunktion innerhalb eines Neurons definiert also ob und in welcher Stärke ein Signal weitergegeben wird (DÖBEL et al. 2018).

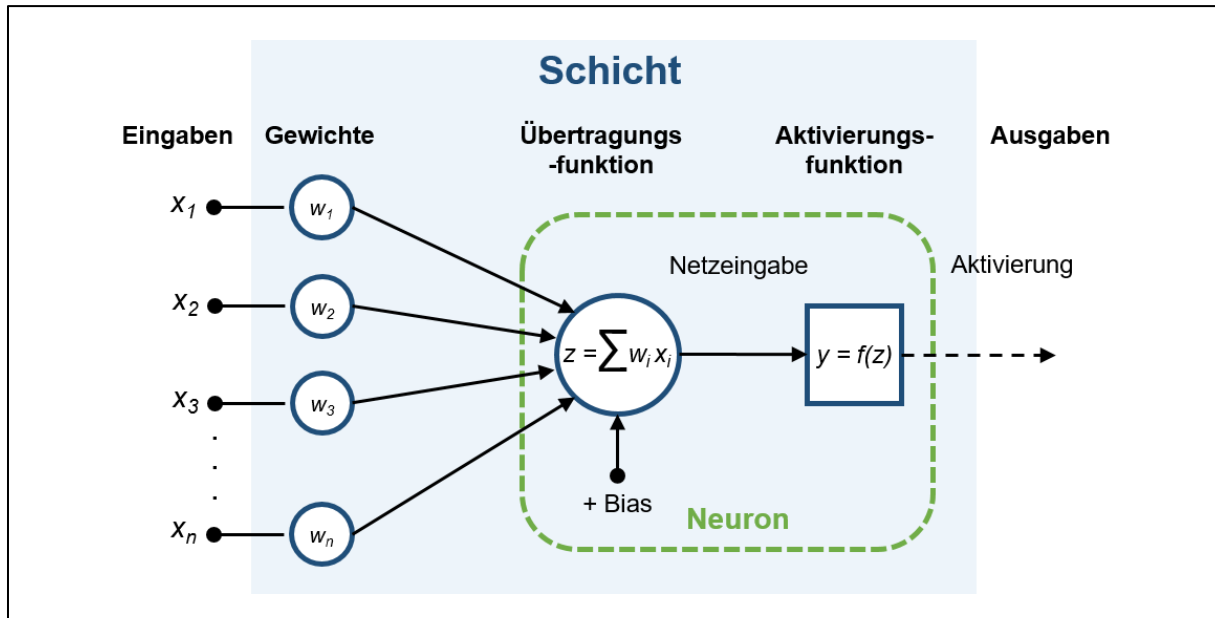


Abbildung 45: Innere Struktur eines Neurons (nach KIMURA et al. 2020).

Ursprünglich wurden vor allem nichtlineare Funktionen, wie die Sigmoid- oder die hyperbolische Tangensfunktion als Aktivierungsfunktion verwendet, weil sie das Verhalten eines biologischen Neurons mathematisch nachbilden (YAMASHITA et al. 2018). Heute wird typischerweise die sogenannte Rectified Linear Unit (ReLU) als nichtlineare Aktivierungsfunktion eingesetzt. Die ReLU-Aktivierungsfunktion ähnelt prinzipiell einer binären Schwellenwertfunktion, jedoch aktiviert ReLU nicht sprunghaft beim Erreichen eines Schwellenwertes, sondern die Aktivierung nimmt beim Überschreiten des Schwellenwertes kontinuierlich linear zu. Dies führt dazu, dass die Summe der Eingangssignale das Neuron erst ab einer gewissen Signalstärke überhaupt aktivieren und wenn eine Aktivierung erfolgt, die Ausprägung des Signals auch entsprechend stark oder schwach an das nachfolgende Neuron weitergeleitet wird. Eine weitere wichtige Aktivierungsfunktion im Kontext von Klassifikationsproblemen ist die sogenannte Softmax Aktivierungsfunktion. Sie kommt typischerweise in den Neuronen der letzten Schicht eines Netzes zum Einsatz und transformiert die in sie eingehenden Werte in eine kategoriale Wahrscheinlichkeitsverteilung (CHOLLET 2021). Die Summe der Ausgaben ergibt dabei immer 1. Der Wert, der für ein Ausgabeneuron mithilfe der Softmax-Aktivierungsfunktion erzeugt wird, repräsentiert die Wahrscheinlichkeit,

dass der betrachtete Input zu dieser bestimmten Klasse gehört. In Abbildung 46 sind typische Aktivierungsfunktionen künstlicher neuronaler Netze dargestellt.

Es ist festzuhalten, dass sowohl bei flachen als auch bei tiefen künstlichen neuronalen Netzen Signale über Schichten von künstlichen Neuronen verarbeitet werden. Die Stärken der Signalübertragungen können in einem Trainingsprozess überwacht angelernet und in einem Modell gespeichert werden, mit dem Ziel, dass das Modell in der Lage ist, auch auf unbekannten Eingabedaten korrekte Vorhersagen zu erzeugen. Man unterscheidet demnach die Trainingsphase, in der die Neuronenverbindungen so angepasst werden müssen, dass die Signalverarbeitung von Eingabe zu Ausgabe möglichst geringe Fehler aufweist und Prädiktionsphase (engl.: Inference), in der ein trainiertes Netzmodell lediglich angewendet wird und auf Basis seines Wissens Ausgaben für Eingaben erzeugt.

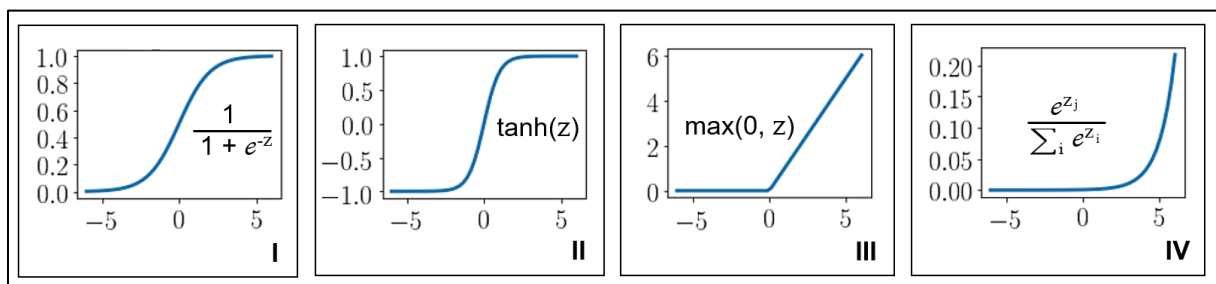


Abbildung 46: Übliche nichtlineare Aktivierungsfunktionen in künstlichen Neuronen Netzen (nach JOHNSON et al. 2020): Sigmoid (I), Tangens Hyperbolicus (II), ReLU (III) und Softmax (IV).

2.4.1.2 Lernen eines neuronalen Netzwerkmodells

In den vorherigen Kapiteln wurde erläutert, dass es beim überwachten maschinellen Lernen darum geht, unter Verwendung einer Vielzahl an Beispieldaten eine Eingabe (z. B. Bild) auf ein Ziel (z. B. Klasseninformation) abzubilden. Künstliche neuronale Netze realisieren diese Zuordnung von Eingaben zu Zielen über eine mehr oder weniger tiefe Abfolge von einfachen Datentransformationen in schichtweise arrangierten Neuronen (CHOLLET 2021). Wie genau sich die Signalumwandlung von Eingaben zu Ausgaben auf Neuronenebene vollzieht, ist in den Gewichten der Neuronenverbindungen gespeichert. Die Gewichte parametrisieren also diese Transformationsprozesse und werden deshalb auch als Parameter eines Netzwerkmodells bezeichnet. Der Begriff des Lernens umschreibt in diesem Zusammenhang, dass die Werte aller Parameter aller Neuronenverbindungen eines Netzes gefunden werden müssen, sodass ein künstliches neuronales Netz die bereitgestellten Trainingsdatenpaare korrekt zuordnet. Dabei ist zu erwähnen, dass vor allem tiefe neuronale Netze mehrere Millionen Neuronenver-

bindungen aufweisen können, die sich darüber hinaus gegenseitig beeinflussen. Das Finden optimaler Werte für alle Parameter ist also unter Umständen sehr rechenaufwändig, weshalb beim Training eines Deep-Neural-Network-Modells wesentlich höhere Hardware-Anforderungen gestellt werden, als zur Anwendung desselben Modells später nötig sind.

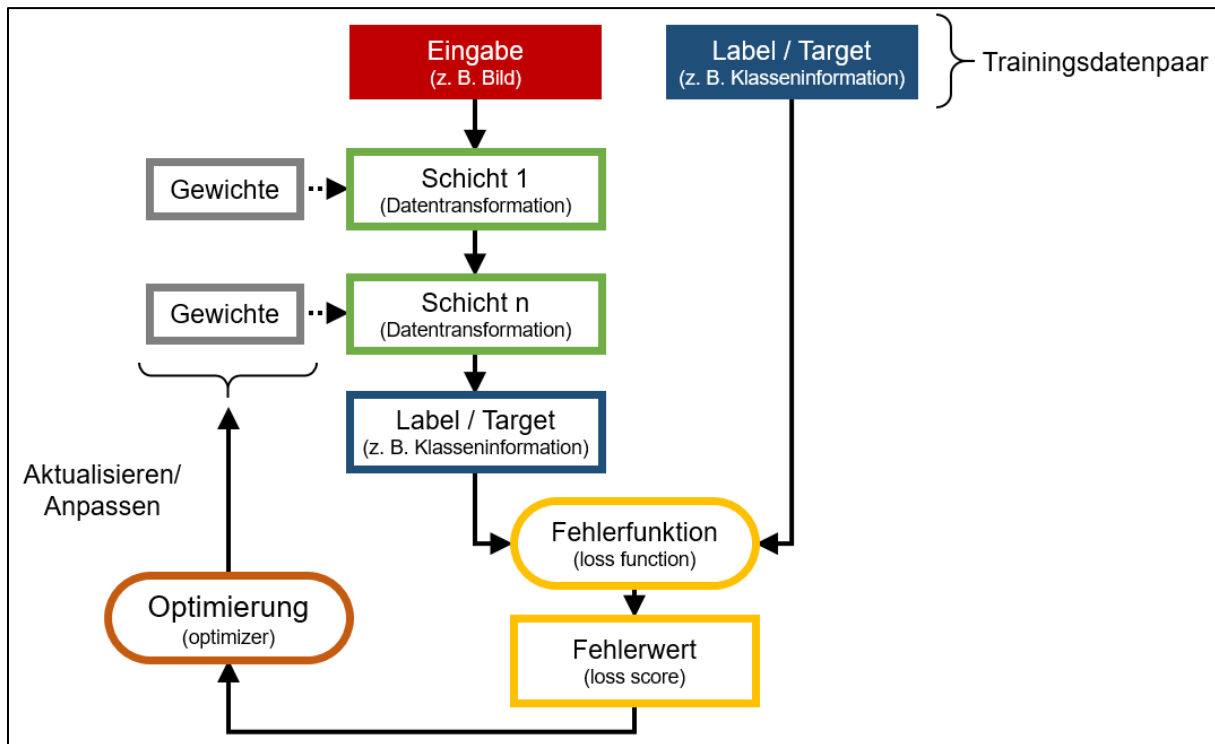


Abbildung 47: Ablaufschema zum Lernen eines neuronalen Netzwerkmodells. Ein neuronales Netz wird durch seine Gewichte parametrisiert. Es gilt, iterativ die richtigen Werte für diese Gewichte zu finden. Eine Fehlerfunktion misst die Qualität der Netzausgabe und der Fehlerwert wird als Feedback-Signal verwendet, um die Gewichte anzupassen (nach CHOLLET 2021).

Um die Ausgabe eines neuronalen Netzes zu kontrollieren, muss man zunächst messen können, wie weit die Netzwerkausgabe vom erwarteten Ergebnis abweicht. Dies ist die Aufgabe der Fehlerfunktion oder Kostenfunktion (engl.: Loss Function) eines Netzes, für die es eine Vielzahl von verschiedenen Algorithmen gibt. Typische Loss Funktionen sind der Mean Squared Error (MSE) oder der Mean Absolute Error (MAE) für Regressionsprobleme und die Binary Cross-Entropy oder die Categorical Cross-Entropy für Klassifikationsprobleme. Letztere wird besonders häufig verwendet, wenn ein Multiklassenproblem zu lösen ist und die letzte Schicht eines Netzwerks eine Softmax Aktivierung beinhaltet. Für alle Fehlerfunktionen gilt, dass sie die Prädiktionsgüte des Netzes für ein Trainingsdatenpaar bewerten. Kern der Funktionsweise von künstlichen neuronalen Netzen ist, dass die Bewertung der Fehlerfunktion als Feedback-Signal genutzt wird, um die Stärke der Neuronenverbindungen schrittweise anzupassen. Diese Anpassung erfolgt in die Richtung, die den Loss Score für das aktuelle

Trainingsdatenpaar senkt, und ist die Aufgabe des sogenannten Optimierers. Dieser wiederum implementiert den Backpropagation-Algorithmus, der für das Lernen in neuronalen Netzwerkmodellen zentral ist. Zu Beginn des Trainings werden die Gewichte eines Netzes zufällig initialisiert, sodass die Wertabbildung von Eingabe auf Ausgabe zufällig geschieht und die Prädiktion der Netzausgabe mit hoher Wahrscheinlichkeit stark vom Zielwert abweicht. Der Loss Score ist dementsprechend initial hoch. Mit jeder weiteren Iteration, in der dieses und auch andere Trainingsdatenpaare das Netz durchlaufen, werden die Gewichte ein wenig angepasst, sodass der Loss Score schrittweise sinkt. Ziel dieser Trainingsschleife ist es, die Gewichte im Netz so anzupassen, dass die Netzausgabe und die Zielwerte den kleinstmöglichen Abstand aufweisen – der Loss Score also minimal wird. In Abbildung 47 ist grafisch dargestellt, wie das Lernen in neuronalen Netzen abläuft.

Wie vorstehend aufgezeigt und auch bei ROSCHER & DREES (2022) beschrieben, lässt sich ein Trainingsprozess also in drei Schritte unterteilen: den Vorwärtsschritt (engl.: Forward Pass), die Berechnung der Verlustfunktion und den Rückwärtsschritt (engl.: Backward Pass). Zentral für den Rückwärtsschritt ist der angesprochene Backpropagation-Algorithmus, der im Folgenden vereinfacht beschrieben wird. Eine detaillierte mathematische Herleitung ist bei RUMELHART et al. (1996) zu finden.

Das Training eines Netzwerks kann als die Optimierung einer Kostenfunktion (auch: Fehlerfunktion oder Verlustfunktion) beschrieben werden. Ziel der Optimierung ist es, ein globales Minimum des Loss Scores zu finden. Das Problem ist, dass nicht bekannt ist, ob ein globales Minimum überhaupt existiert oder wie viele lokale Minima es gibt. Dieses Problem adressiert der Backpropagation-Algorithmus. Dabei handelt es sich um einen Algorithmus zur iterativen Berechnung von Gradienten für die Gewichte in der Graphenstruktur eines neuronalen Netzes. Der Gradient ist aus mathematischer Sicht eine partielle Ableitung der Verlustfunktion und gibt die Richtung an, in der die Funktion die steilste Anstiegsrate hat. Jedes Gewicht wird in der entgegengesetzten Richtung des Gradienten mit einer bestimmten Schrittweite aktualisiert, die auf der Grundlage der sogenannten Lernrate bestimmt wird (YAMASHITA et al. 2018). Im Zuge des Backward Passes werden die mittels Kostenfunktion ermittelten Fehler unter Verwendung dieses als Gradientenabstieg bezeichneten Verfahrens rückwärts-gerichtet an die Gewichte angebracht, die an den Fehlern beteiligt waren. Dabei nutzt der Backpropagation-Algorithmus die Kettenregel der Differentialrechnung: Gradienten aus früheren Schichten werden multipliziert um Gradienten der folgenden Schichten des Netzes abzuleiten. Das beschriebene Verfahren hat den Nachteil, dass die Optimierung unter Umständen nicht zu einem globalen Minimum hin konvergiert, sondern gegen ein

suboptimales, lokales Minimum strebt (siehe Abbildung 48). Außerdem kann das Training verlangsamt werden, wenn flache Fehlerplateaus auftreten, die sehr viele Iterationen der Gradientenberechnung benötigen, um sie zu überwinden.

Um diese Probleme zu lösen, wurden verschiedene Verbesserungen und Modifikationen des ursprünglichen Gradientenabstiegsverfahrens entwickelt, wie beispielsweise der stochastische Gradientenabstieg (SGD), SGD mit Momentum, RMSprop oder Adam. Die verschiedenen Optimierer werden in dieser Arbeit nicht weiter behandelt. Ein Überblick über Optimierer, die im Anwendungsbereich Bildklassifizierung verwendet werden, ist bei DOGO, AFOLABI & TWALA (2022) zu finden.

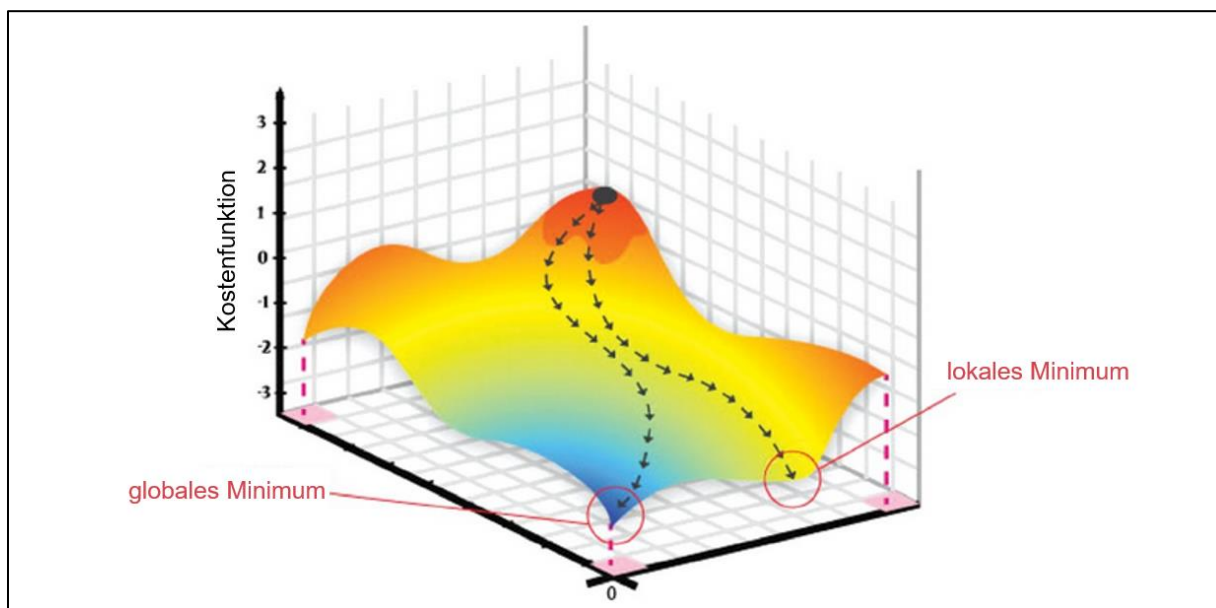


Abbildung 48: Prinzipdarstellung der Berechnung des Gradientenabstiegs auf der Fehleroberfläche einer Kostenfunktion (nach KRITTANAWONG et al. 2019).

Die Begriffe Epoche, Batch-Größe, Iteration und Lernrate wurden teilweise schon genannt, werden an dieser Stelle aber noch einmal explizit beschrieben, da sie für das Verständnis späterer Ausführungen wichtig sind. Der Begriff Epoche beschreibt, dass die Gesamtheit der Trainingsdatenpaare einmal vorwärts und rückwärts durch das Netz propagiert wurde (AGHDAM & HERAVI 2017). In der Regel ist die Datenmenge zu groß, um sie in einer einzigen Epoche in ein Netz einzuspeisen, weshalb die Gesamtheit der Trainingsdatenpaare in kleinere Einheiten, die sogenannten Batches, aufgeteilt wird. Die Batch-Größe gibt die Anzahl der Trainingsdatenpaare innerhalb eines Batches an, die im Zuge der Backpropagation für das Anpassen der Gewichte verwendet werden. Eine Batch-Größe > 1 führt dazu, dass die Anpassung der Gewichte nicht für jedes Trainingsbeispiel einzeln erfolgt, sondern die Gewichte mit dem Durchschnitt der Gradienten im Batch aktualisiert werden. Damit können potentiell

wesentlichere Informationen robuster gelernt werden. Da die Optimierung ein iterativer Prozess ist, reicht es nicht aus, den Trainingsdatensatz nur ein einziges Mal vorwärts und rückwärts durch das Netz zu propagieren. Je nach dem, wie komplex das zu lösende Problem ist, sind mehrere dutzend oder mehrere hundert Epochen nötig, bevor das Netz konvergiert und das Training als abgeschlossen betrachtet werden kann. Der Begriff Iterationen meint die Anzahl der Batches, die benötigt werden, um eine Epoche abzuschließen. Liegen beispielsweise 5000 Trainingsdatenpaare vor, dann werden bei einer Batch-Größe von 50 insgesamt 100 Batches erzeugt. Um eine Epoche abzuschließen, müssen also insgesamt 100 Iterationen vollendet werden. Die Lernrate ist ein Hyperparameter, der als Teil der Optimierer-Konfiguration für den Trainingserfolg entscheidend ist. YAMASHITA et al. (2018) beschreiben die Lernrate sogar als einen der wichtigsten Hyperparameter, die vor Beginn des Trainings festgelegt werden müssen. Der zu verwendende Wert liegt typischerweise zwischen 0 und 1 und hängt stark von der Problemstellung, der Netzwerkarchitektur und den Trainingsdaten ab. Wird ein zu geringer Wert verwendet, steigt der Zeit- und Rechenaufwand für das Training. Ein zu großer Wert kann dazu führen, dass das Training sprunghaft und unkontrolliert konvergiert. Eine gebräuchliche Strategie ist es, das Training mit einer relativ großen Lernrate zu beginnen und nach einer gewissen Anzahl von Iterationen den Wert stufenweise zu reduzieren (siehe Abbildung 49 links).

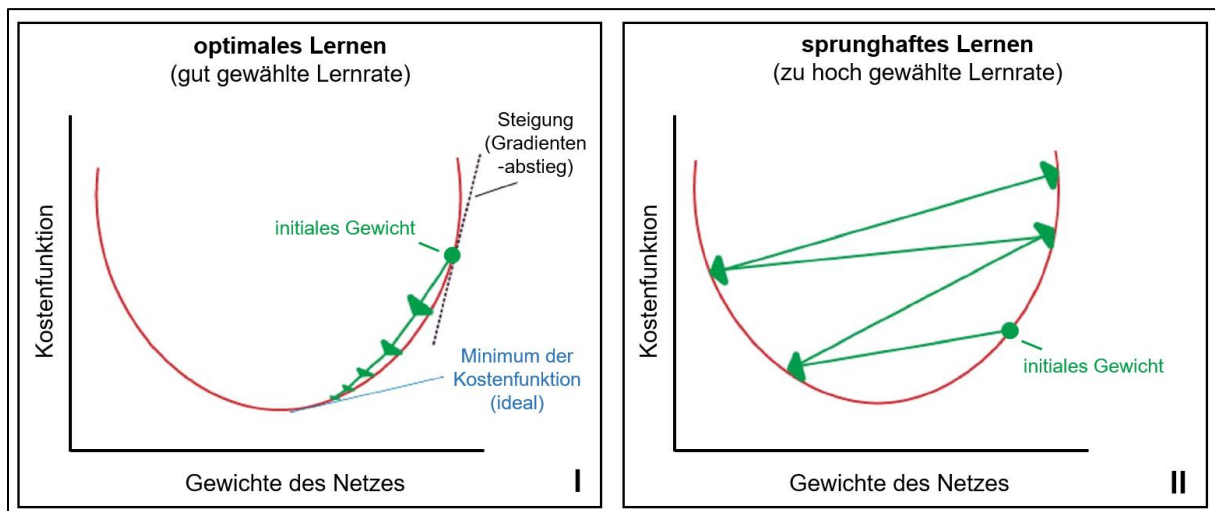


Abbildung 49: Einfluss der Lernrate auf das Training eines Modells (nach KRITTANAWONG et al. 2019). Eine schrittweise Reduktion der Lernrate kann dazu beitragen, das Minimum der Kostenfunktion effizient zu identifizieren (I). Ist die Lernrate zu hoch eingestellt, kann das Minimum gegebenenfalls gar nicht gefunden werden (II).

Das Training eines neuronalen Netzmodells kann als beendet angesehen werden, wenn sich der Loss Score auf einem Validierungsdatensatz nicht mehr weiter minimieren lässt (CHOLLET 2021). Nach wie vielen Epochen dieser Punkt erreicht ist, kann beim Start eines Trainings noch

nicht sicher abgeschätzt werden. Da ein zu langes Training zu einer Überanpassung des Modells auf die Trainingsdaten führt, ist das Identifizieren eines optimalen Zeitpunkts für das Trainingsende von entscheidender Bedeutung für die letztendliche Leistungsfähigkeit eines trainierten Modells.

2.4.1.3 Generalisierungsfehler und Regularisierungsstrategien

Im Bereich des maschinellen Lernens werden Fehler, die ein trainiertes Modell auf unbekannten Daten macht, Generalisierungsfehler (engl.: Out-Of-Sample Error) genannt. Eine genaue Berechnung des wahren Fehlers, den das Modell bei der Inferenz fortlaufend auf unbekannten Daten erzeugen wird, ist nicht möglich – wohl aber eine Approximation dieses wahren Fehlers, beispielsweise über den empirischen Fehler. Zu diesem Zweck wird das sogenannte Hold-Out Splitting angewendet und dabei der insgesamt verfügbare Datensatz bestehend aus Daten und Zielwerten in mindestens zwei Teile aufgeteilt (CHOLLET 2021): einen Trainings- und einen Testdatensatz. Der Trainingsdatensatz enthält dabei Daten-Zielwert Paare, die für das Training eines Modells verwendet werden, wobei der Loss im Forward Pass berechnet und die lernbaren Parameter im Backward Pass angepasst werden. Der Testdatensatz dient lediglich der Bestimmung der Leistungsfähigkeit des trainierten Modells auf Daten, die während des Trainings nicht verwendet wurden und damit letztlich der Schätzung des Generalisierungsfehlers. Die beiden Teildatensätze sollten unabhängig voneinander sein, sich aber hinsichtlich der Besetzungshäufigkeit der Klassen möglichst ähneln. Sollen beispielsweise Verkehrszeichen in Bildern klassifiziert werden, sollten die Beispiele im Trainings- und Testdatensatz möglichst nicht entlang derselben Straße aufgenommen worden sein. Für eine realistische Schätzung des Generalisierungsfehlers wäre es besser, die Beispiele stammten aus verschiedenen Städten. Wenn das Hold-Out Splitting optimal umgesetzt wird, dann kann der empirisch bestimmte Fehler auf den Testdaten den wahren Fehler in der späteren Anwendungspraxis durchaus gut annähern. Eine weitaus komplexere, aber auch genauere Schätzung des Generalisierungsfehlers kann u. a. über eine frequentistische oder Bayes'sche Statistik erreicht werden, was in der vorliegenden Arbeit aber nicht weiter betrachtet wird. Es sei hier erwähnt, dass auch noch ein zusätzlicher dritter Teildatensatz separiert werden kann – der Validierungsdatensatz. Validierungsdaten dürfen eine gewisse Ähnlichkeit zu den Trainingsdaten aufweisen, da sie lediglich dazu dienen, das Modell mit dem besten Hyperparameter-Setting auszuwählen, welches dann auf dem Testdatensatz evaluiert wird. Bei einer geringen Datenverfügbarkeit kann die Modellauswahl aber auch auf dem Testdatensatz erfolgen.

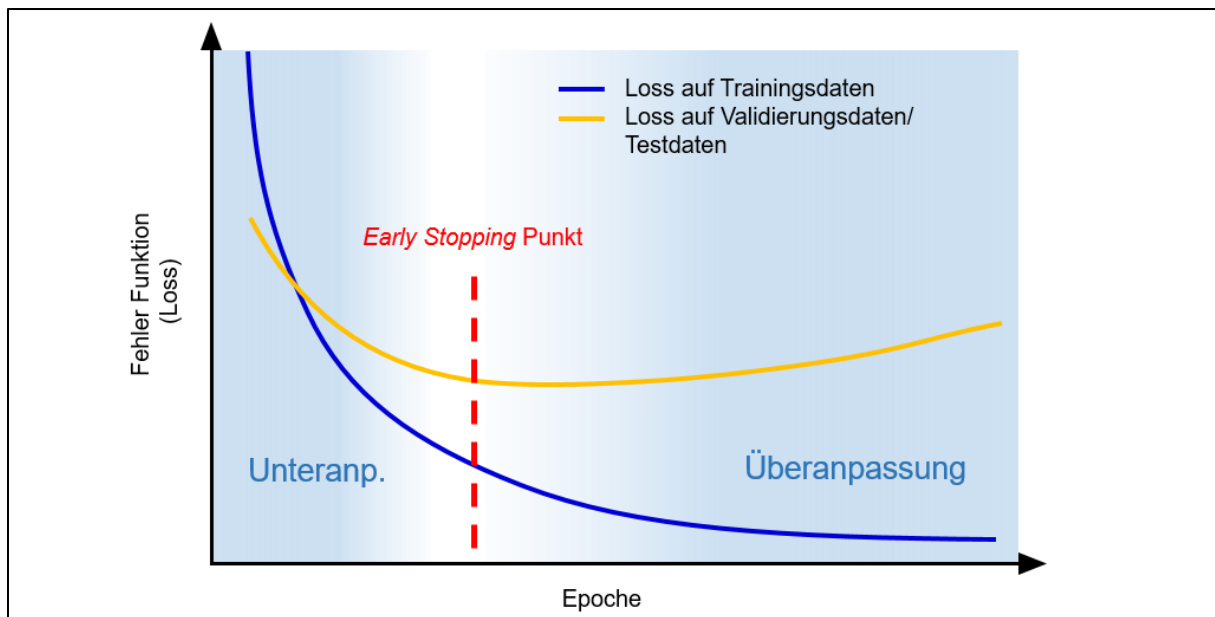


Abbildung 50: Verlauf der Loss Kurven auf Trainings- und Validierungsdaten zur Identifikation des Early Stopping Zeitpunktes, der den Punkt markiert, ab dem das trainierte Modell beginnt sich auf die Trainingsdaten überanzupassen (nach CHOLLET 2021).

Generalisierungsfehler können aus einer Unter- oder einer Überanpassung eines trainierten Modells auf seine Trainingsdaten resultieren (siehe Abbildung 50). Während der ersten Trainingsiterationen sind die Loss Scores sowohl auf den Trainings- als auch auf den Testdaten noch vergleichsweise hoch, denn das Netz hat noch nicht alle relevanten Muster in den Trainingsdaten modelliert. Diese Phase des Trainings wird als Unteranpassung bezeichnet. Mit zunehmender Iterationszahl verbessert sich die Generalisierungsfähigkeit des Modells immer mehr und die Loss-Kurven sinken. Ab einem gewissen Punkt stagniert diese Entwicklung jedoch und der Loss Score kann sich unter Umständen verschlechtern. Diese Phase des Trainings wird als Überanpassung bezeichnet und beschreibt eine Situation, in der ein Modell für die Trainingsbeispiele spezifische Muster oder irrelevante Anomalien erlernt, die für unbekannte Daten jedoch eine untergeordnete oder gar keine Rolle spielen (YAMASHITA et al. 2018). Ein solches Modell kann zwar auf den Daten, auf denen es trainiert ist, einen sehr niedrigen Loss Score erreichen, bei der Anwendung auf neue Daten jedoch in der Praxis unbrauchbare Ergebnisse erzielen. Ein niedriger Loss-Wert eines Modells hat keine Aussagekraft im Hinblick auf den Generalisierungsfehler.

In der Fachliteratur werden Techniken, die das Ziel haben, den Generalisierungsfehler zu reduzieren, als Regularisierungstechniken oder -strategien beschrieben (SMIRNOV, TIMOSHENKO & ANDRIANOV 2014). Eine Maßnahme gegen Unteranpassung wäre die Verwendung eines tieferen bzw. komplexeren Netzwerkmodells. Um eine Überanpassung des Modells im Zuge des Trainings zu vermeiden, sollte der Trainingsprozess zu dem Zeitpunkt beendet

werden, an dem sich der Loss Score auf dem Validierungs- bzw. Testdatensatz nicht weiter minimieren lässt (CHOLLET 2021). Dies wird als Early Stopping bezeichnet (siehe Abbildung 50).

Darüber hinaus beschreiben YAMASHITA et al. (2018) folgende Methoden, um das Risiko einer Modellüberanpassung zu minimieren: Das wirksamste Mittel gegen Überanpassung ist das Erweitern des Trainingsdatensatzes um mehr und diversere Beispieldaten. Zusätzlich kann Data Augmentation unterstützend eingesetzt werden (CHOLLET 2021). Dabei handelt es sich um einen Ansatz, zur künstlichen Verbreiterung der Beispieldatenbasis, in dem zufällige Variationen der bestehenden Daten erzeugt werden. Bei Bilddaten werden beispielsweise skalierte, gedrehte, gespiegelte oder hellkeitsveränderte Versionen der eigentlichen Eingabedaten erzeugt.

Die L2-Regularisierung (engl.: Weight Decay) ist eine weit verbreitete Regularisierungstechnik und setzt in der Kostenfunktion des Netzes an. Durch Hinzufügen eines zusätzlichen Strafterms zum eigentlichen Gradienten-Term bei der Berechnung des Loss Scores wird erreicht, dass bestimmte Gewichte im Netz nur sehr kleine Werte annehmen. Neuronenverbindungen mit sehr niedrigen Gewichten führen dazu, dass über sie kein wesentlicher interneuronaler Signalfluss mehr stattfindet, das Modell dadurch weniger komplex und letztlich robuster gegen Überanpassung wird. Um effektiv gegen Überanpassung wirken zu können, sollte der Weight-Decay-Koeffizient der L2-Regularisierung für jede Schicht einzeln ermittelt werden, sodass das Verhältnis zwischen Gradienten-Term und Weight-Decay-Term über das gesamte Netz hinweg gleich bleibt. Details hierzu sind bei ISHII & SATO (2018) beschrieben.

Eine andere Regularisierungsmethode ist das Dropout. Dabei werden zufällig ausgewählte Aktivierungen während des Trainings auf 0 gesetzt, was dazu führt, dass diese Neuronen quasi temporär aus dem Netz entfernt sind. Wie bei der L2-Regularisierung führt dies zu einer Abnahme der Modellkomplexität, in deren Folge das Modell weniger empfindlich auf bestimmte Verbindungen im Netz reagiert (SRIVASTAVA et al. 2014). In Abbildung 51 ist dieses Prinzip dargestellt.

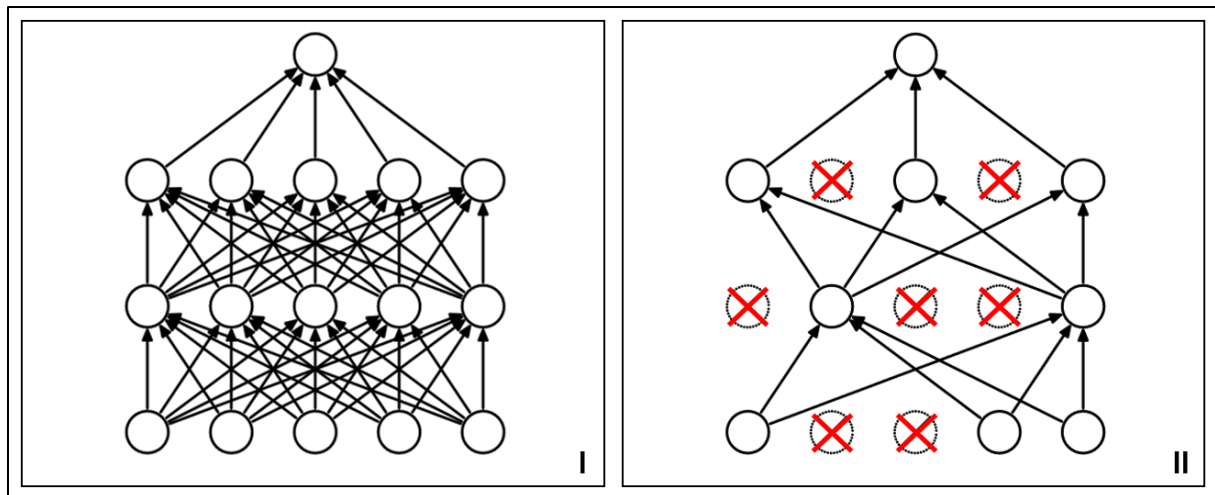


Abbildung 51: Schematische Darstellung eines künstlichen neuronalen Netzes mit zwei versteckten Schichten (I). Rechts ist dasselbe Netz abgebildet, bei dem jedoch einzelne Neuronen mithilfe der Regularisierungstechnik Drop-out entfernt wurden (II nach SMIRNOV, TIMOSHENKO & ANDRIANOV 2014).

Ein weiteres Regularisierungskonzept ist die Batch-Normalisierung (LOFFE & SZEGEDY 2015). Dieses Konzept sieht vor, die Neuronenverbindungen in den versteckten Schichten eines Modells zu normalisieren, sodass das Aktivierungssignal jedes Neurons einer Standard-Normalverteilung entspricht. Dazu werden für die Trainingsdaten des aktuellen Batches Mittelwert und Standardabweichung der Aktivierungswerte berechnet und das Aktivierungssignal damit normalisiert. Dieser Normalisierungsschritt wird für den aktuellen Batch typischerweise direkt vor der Aktivierungsfunktion durchgeführt und kann als separate Schicht interpretiert werden, die an gewünschten Stellen innerhalb des Netzgraphen eingesetzt werden kann. Abbildung 52 zeigt Histogramme über die Gradienten einer bestimmten versteckten Schicht mit und ohne vorherige Batch Normalization. Für die nicht normalisierte Schicht sind die Gradienten mit ausgeprägten Flanken verteilt, während die Gradienten für die normalisierte Schicht um den Mittelwert konzentriert sind.

In vielen Fällen trägt Batch Normalization dazu bei, dass robustere Modelle schneller trainiert werden können. Zu den Gründen dafür, warum diese Regularisierungstechnik so wirkungsvoll ist, existieren verschiedene Hypothesen (LOFFE & SZEGEDY 2015; SANTURKAR et al. 2018). Mutmaßlich verringert Batch Normalization die wechselseitigen Abhängigkeiten zwischen versteckten Schichten während des Trainings. Ein weiterer wichtiger Mechanismus für die Wirksamkeit scheint zu sein, dass die Normalisierungstransformation eine glattere Fehleroberfläche erzeugt, was die Optimierung im Zuge der Backpropagation begünstigt. Auch YAMASHITA et al. (2018) weisen darauf hin, dass mit Batch Normalization ein optimierter Gradientenfluss im Netz herrscht, was wiederum ermöglicht, höhere Lernraten beim Training zu nutzen.

Schließlich kann auch die Verwendung einer weniger komplexen Netzarchitektur dazu beitragen, Überanpassung zu vermeiden, da auch dies die Anzahl der lernbaren Parameter insgesamt reduziert.

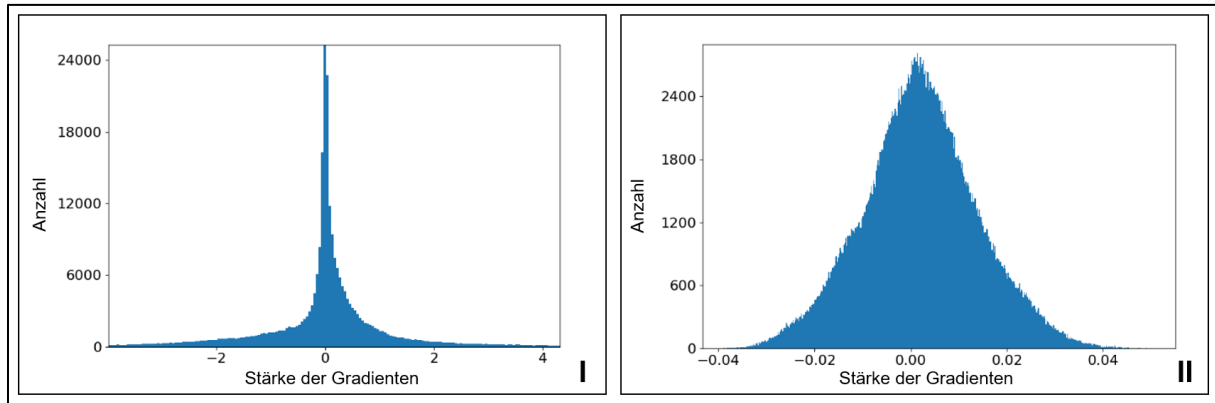


Abbildung 52: Idealisierte Histogramme der Gradientenstärke einer neuronalen Netzwerkschicht mit (I) und ohne Batch-Normalization (II) (nach BJORCK et al. 2018).

2.4.2 Semantische Bildsegmentierung mithilfe von Deep Learning

Während es bei der Bildklassifizierung mithilfe künstlicher neuronaler Faltungsnetze im Kern darum geht, einem gesamten Bild eine bestimmte Klasse zuzuweisen, zielt die semantische Bildsegmentierung darauf ab, Objekte verschiedener Klassen in einem Bild zu lokalisieren und voneinander abzugrenzen. Mithilfe semantischer Bildsegmentierung versucht man sich der Frage zu nähern, welche Inhalte in einem Bild vorhanden sind und wo sie sich im Bild genau befinden.

Semantische Bildsegmentierung ist dabei keine ausschließlich dem Deep Learning vorbehaltene Aufgabenstellung – obschon tiefe neuronale Faltungsnetze dieses Anwendungsfeld der digitalen Bildanalyse zweifelsohne revolutionierten (CHOLLET 2021). Auch konventionelle Algorithmen, wie beispielsweise die schwellenwertbasierte Bildbinarisierung, die Wasserscheiden-Transformation oder Cluster-Analysen, können ein Eingabebild in homogene Bereiche mit ähnlichen Eigenschaften unterteilen (JÄHNE 2012). Jährlich stattfindende, weltweite Wettbewerbe, wie beispielsweise die ImageNet Large-Scale Visual Recognition Challenge (ILSVRC), bei denen verschiedene Methoden in den Anwendungsfeldern Bildklassifizierung, Objekterkennung und Bildsegmentierung konkurrieren (DENG et al. 2009), wurden vor dem Aufkommen von Deep-Learning-Ansätzen meist von klassischen merkmalsbasierten Machine-Learning-Ansätzen, wie Boosting oder Random Forest gewonnen (CHOLLET 2021). Im Bereich der Analyse von Fernerkundungsdaten dominierten lange Zeit die

Object-based Image Analysis (OBIA) (BLASCHKE & STROBL 2001), die stark mit den kommerziellen Softwarelösungen eCognition und ERDAS Imagine assoziiert sind (BENZ et al. 2004). Das Grundkonzept von OBIA beruht darauf, dass in einem Segmentierungsschritt eine Clusterung der Eingabedaten auf einer oder mehreren Skalenebenen durchgeführt wird. Anschließend werden in einem Klassifikationsschritt Eigenschaften wie Textur-, Topologie- und Formparameter für die gefundenen Cluster ermittelt und den Objektklustern mithilfe von schwellenwertbasierten Entscheidungsregeln und Fuzzy-Logikansätzen semantische Information zugewiesen (NOACK 2019).

Seit dem Jahr 2012 setzte eine Entwicklung ein, bei der tiefe neuronale Faltungsnetze (engl.: Convolutional Neural Networks) alle bisherigen Verfahren aus Fachartikeln, Konferenzbeiträgen und Praxisanwendungen nahezu verdrängten, sodass es auf den einflussreichen Konferenzen zu Computer Vision, wie der CVPR, in den Jahren 2015 und 2016 fast unmöglich war, Präsentationen zu finden, die nicht in irgendeiner Form Faltungsnetze enthielten (CHOLLET 2021).

Durch die Entwicklung des sogenannten Fully Convolutional Networks (LONG, SHELHAMER & DARRELL 2015) wurde es möglich, ein gesamtes Eingabebild in einem einzigen Durchgang semantisch zu segmentieren. Im Vergleich zu vorher verbreiteten Bildausschnittbasierten Ansätzen, bei dem nach dem Sliding-Window-Prinzip Bildausschnitte aus einem Eingabebild extrahiert und eine Segmentierung auf Basis der Patch-weisen Bildklassifikation erzeugt wird (HOU et al. 2016), sind vollbildbasierte Ansätze besser in der Lage, Kontextinformationen zu verarbeiten und sie sind effizienter (OLIVEIRA 2019). In Folgejahren wurde das in der Fully Convolutional Network Architektur noch eher einfach implementierte Encoder/Decoder-Konzept in verschiedenen Architekturen weiterentwickelt. Dabei wurden verschiedene Ansätze entwickelt, die aus den dichten Merkmalsrepräsentationen am Ende des Convolutional Neural Networks im Encoder-Teil eine detailgenaue semantische Segmentierung mit den Dimensionen der Eingabedaten im Decoder-Teil erzeugen.

Die folgenden Kapitel bieten einen Überblick über Charakteristika von Convolutional Neural Networks und ausgewählte Encoder-Decoder-Architekturen, welche wichtige Konzepte und Entwicklungen beinhalten, die in dem Deep-Learning-System, das im Rahmen der vorliegenden Arbeit genutzt wird, aufgegriffen und optimiert wurden.

2.4.2.1 Convolutional Neural Networks

Convolutional Neural Networks (CNN) bilden eine spezielle Gruppe von tiefen neuronalen Netzen für die Verarbeitung von Daten, die eine Gitterstruktur aufweisen. Inspiriert vom Aufbau und von der Funktionsweise des visuellen Kortex von Menschen und Tieren, sind CNN in der Lage automatisch und adaptiv räumliche Hierarchien von Merkmalen zu lernen – von einfachen, universellen bis hin zu komplexen, spezifischen Mustern (YAMASHITA et al. 2018). Seit den 2010er-Jahren erlangten künstliche neuronale Faltungsnetze große Verbreitung in zahlreichen Anwendungsfeldern, wie der Bildklassifizierung, Objekt- bzw. Gesichtserkennung oder Text- bzw. Spracherkennung (SMIRNOV, TIMOSHENKO & ANDRIANOV 2014). Hervorzuheben sind hierbei KRIZHEVSKY, SUTSKEVER & HINTON (2012), die zweimal in Folge den prestigeträchtigen Wettbewerb ILSVRC gewinnen konnten.

In diesem Kapitel werden zunächst die Grundelemente erklärt, aus denen ein CNN im Kontext der digitalen Bildanalyse aufgebaut werden kann. Außerdem wird anhand von ausgewählten Beispielen die Entwicklung verschiedener CNN-Architekturen aufgezeigt. Umfassende Informationen zu CNN sind u. a. bei AGHDAM & HERAVI (2017), YAMASHITA et al. (2018) und CHOLLET (2021) zu finden.

Zum Verständnis der Informationsverarbeitung in einem CNN sind neben der Eingabe- und Ausgabeschicht drei Grundbausteine von essenzieller Bedeutung: Convolution-Schichten, Pooling-Schichten und Fully-Connected-Schichten. Die Eingabeschicht eines CNN bilden, wie in Kapitel 2.4.1.1 bereits erklärt, jene Neuronen, an die die Pixel-Werte des Eingabebildes angelegt werden. Ein 224×224 Pixel großes Bild mit drei Farbkanälen entspräche demnach 150528 Neuronen in der Eingabeschicht.

Convolution-Schichten enthalten einen Faltungsoperator gefolgt von einer nichtlinearen Aktivierungsfunktion. Der Faltungskernel definiert eine lokal begrenzte Region auf dem Eingabebild. Typischerweise verarbeitet eine ReLU Aktivierungsfunktion die lineare Aktivierung der Faltungsausgabe weiter (siehe Kapitel 2.4.1.1). Wie in Abbildung 53 dargestellt, wird ein Kernel auf eine Eingangsmatrix (2D-Tensor) angewendet. Ein Tensor bezeichnet in diesem Fall ein Array von Zahlenwerten. Im Zuge des Faltungsprozesses wird ein elementweises Produkt zwischen jedem Kernel-Element (Wert einer Kernel-Zelle) und dem Eingangstensor an jeder Stelle des Kernels auf dem Eingabebild berechnet und summiert, um den Ausgangswert an der entsprechenden Stelle des Ausgangstensors zu erhalten. Dieser Ausgangstensor stellt die Aktivierung dar und wird auch als Merkmalskarte (engl.: Feature-Map) bezeichnet. Die beschriebene Operation wird innerhalb einer Schicht mit mehreren Kernels wiederholt, um eine ganze Reihe von linearen Aktivierungen zu generieren. So kann eine beliebige Anzahl von

Merkmalskarten erzeugt werden, wobei jede etwas andere Merkmale des Eingabetensors kodiert. Die Kernel einer Convolution-Schicht können somit als Merkmalsextraktoren angesehen werden (YAMASHITA et al. 2018). Anzahl und Größe der Kernel innerhalb einer Convolution-Schicht variieren je nach Netzarchitektur. Hinsichtlich der Größe sind 3×3 , 5×5 oder 7×7 Pixel Kernel üblich. Die Kernelanzahl ist eine Frage des Architektur-Konzepts und hat direkten Einfluss auf die Tiefe der ausgegebenen Feature-Maps.

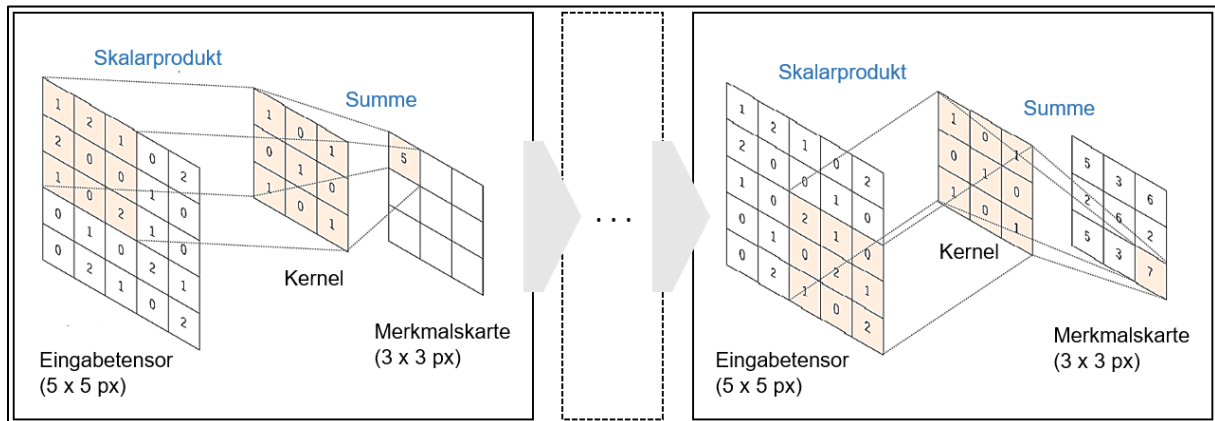


Abbildung 53: Faltung als zentrale Operation in einer Convolution-Schicht. Ein Kernel wird schrittweise über eine Eingabematrix geschoben und erzeugt so eine Merkmalskarte als Ausgabe (nach YAMASHITA et al. 2018).

Der Abstand zwischen zwei Kernel-Positionen wird als Stride bezeichnet. Ein Stride von 1 bedeutet, dass der Kernel auf jeden Matrixwert des Eingabetensors zentriert, also in Einer-Schritten verschoben wird. Stride-Werte > 1 werden nur gewählt, wenn die Ausgabe deutlich kleiner werden soll als die Eingabe. Damit die Eingabe nicht automatisch bei jeder Faltung auch bei Stride = 1 kleiner wird, findet typischerweise eine Randbehandlung statt – das sogenannte Padding. Beim Zero Padding wird beispielsweise um den Eingabetensor temporär ein Rand mit Nullen erzeugt, sodass ein Kernel auch auf die äußersten Werte des Eingabetensors zugreifen kann und die Feature-Map dieselbe Größe hat wie der Eingabetensor (siehe Abbildung 54).

Eine Schlüsseleigenschaft von Convolution-Schichten ist das sogenannte Weight Sharing. Es bezeichnet den Umstand, dass bei einem CNN die Kernel-Werte Gewichte repräsentieren und jedes Kernel-Element an jeder Position des Eingabetensors angewendet wird. Daraus ergeben sich Vorteile, die YAMASHITA et al. (2018) wie folgt beschreiben: CNN können effizienter trainiert werden, als reine vollverbundene KNN (engl.: Fully Connected Neural Networks), da wesentlich weniger Gewichte angelern werden müssen. Im Beispiel von Abbildung 53 wären in einem KNN die 25 Neuronen der Eingabe mit den neun Neuronen der Ausgabe verbunden. Es wären hier also 225 Gewichte zu optimieren. Im CNN dagegen müssen hier lediglich die 9 Kernel-Elemente als Gewichte angepasst werden. Über diesen Effizienzaspekt hinaus, ist die

Invarianz der Merkmale, die durch die schrittweise über die Eingabe laufenden Kernel erfasst werden, ein entscheidender Vorteil. Als drittes Charakteristikum von Convolution-Schichten nennen die Autoren, dass durch Faltungsschichten, vor allem wenn sie mit Pooling-Schichten kombiniert eingesetzt werden, räumliche Hierarchien von Merkmalsmustern gelernt werden können.

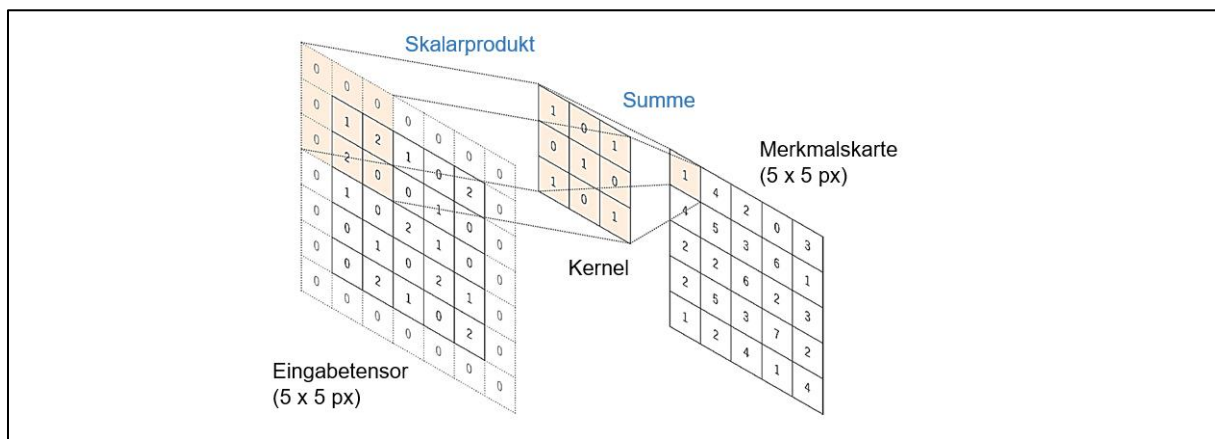


Abbildung 54: Zero Padding als eine Möglichkeit, Merkmalskarten mit denselben Dimensionen des Eingabetensors zu erzeugen (nach YAMASHITA et al. 2018).

Während Convolutional-Schichten Informationen in Form von Merkmalen aus einem Eingangstensor extrahieren, dienen Pooling-Schichten dazu, den Informationsgehalt ihrer Eingabe zu verdichten und Merkmale der Vorgängerschicht entsprechend zu generalisieren. Hierzu verkleinert die Pooling-Schicht die gegebene Eingabe entlang ihrer räumlichen Dimensionalität (engl.: Downsampling). Um den Informationsverlust zu begrenzen, wird in der Regel nur ein 2 x 2 Pixel Pooling verwendet. Am gebräuchlichsten ist die Max-Pooling-Operation, wobei der Maximalwert im Sichtfeld einer Kernel-Position für die Ausgabe übernommen wird (siehe Abbildung 55). Typischerweise werden Max-Pooling-Schichten mit Convolution-Schichten in einem CNN im Wechsel aufeinanderfolgend kombiniert. Eine Sonderform stellt das Global Average Pooling dar. Hierbei wird die Eingabe auf ein 1 x 1 Pixel Array verdichtet, indem der Mittelwert des gesamten Eingabetensors berechnet wird. Die Tiefe bleibt dabei erhalten. Liegen beispielsweise 32 Feature-Maps von 5 x 5 Pixel Größe vor, so ist das Ergebnis nach dem Global Average Pooling ein Tensor mit den Dimensionen 1 x 1 x 32 (Höhe x Breite x Tiefe). Diese Sonderform des Poolings wird in CNN nur direkt vor einer Fully-Connected-Schicht verwendet und hat den Vorteil, dass es das CNN befähigt, auch mit variablen Bildgrößen in der Eingabeschicht umgehen zu können.

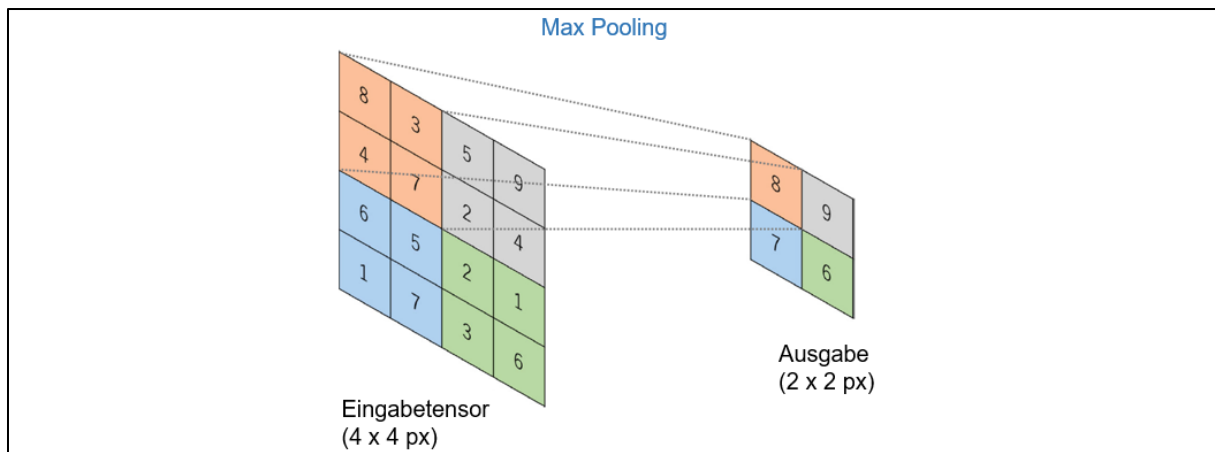


Abbildung 55: Beispiel für eine Pooling Operation mit einem 2 x 2 Pixel Filterelement, das ohne Padding und mit Stride = 2 angewendet wird. Die Ausgabe enthält nur die Maximalwerte im Sichtfeld des Kernels und ist deutlich kleiner als die Eingabe (nach YAMASHITA et al. 2018).

Die Ausgabe der letzten Convolution- oder Pooling-Schicht eines CNN wird typischerweise in einen eindimensionalen Zahlenvektor umgewandelt. Dieser Vorgang wird Flattening genannt. Die Elemente des resultierenden 1D-Vektors werden dann mit einer oder mehreren Fully-Connected-Schichten (auch: Dense Layers) verbunden. Dabei handelt es sich um schichtweise arrangierte Neuronen analog der versteckten Schichten, die in Kapitel 2.4.1.1 beschrieben sind. Ziel von Fully-Connected-Schichten ist es, die letzten Aktivierungen aus der Abfolge von Convolution- und Pooling-Schichten (Merkmalsextraktion) auf die Netzwerkausgabe (Klassifikation) abzubilden. Während zwischen verschiedenen Fully-Connected-Schichten oftmals auch ReLU als Aktivierungsfunktion eingesetzt wird, nutzt die letzte Fully-Connected-Schicht in einem CNN typischerweise eine Softmax-Aktivierung.

Das Lernen in einem CNN funktioniert grundlegend nach dem selben Prinzip, wie in Kapitel 2.4.1.2 beschrieben. In Abbildung 56 sind Aufbau und Training eines CNN schematisch dargestellt. Ein wesentlicher Unterschied zu Fully Connected Networks wurde bereits im Kontext des Weight Sharings angerissen: im Zuge des Backward Passes werden nicht nur die Gewichte der Verbindungen zwischen den Neuronen der Fully-Connected-Schichten modifiziert, sondern es werden auch die Werte in den Filter-Kernels so optimiert, dass sie eine auf die Problemstellung und Trainingsdaten angepasste Merkmalsextraktion umsetzen. Die Kernel-definierenden Werte (z. B. Kernelgröße, Stride und Padding) zählen bei CNN auch zu den Hyperparametern. Wie auch die Lernrate müssen sie vor dem Training definiert werden. Ziel des Trainingsprozesses ist es, die passenden Werte für die Kernel-Elemente in Faltungsschichten sowie die optimalen Gewichte in den vollständig verknüpften Schichten zu finden, die die Unterschiede zwischen den Ausgabevorhersagen und den gegebenen Labels in einem Trainingsdatensatz minimieren (YAMASHITA et al. 2018).

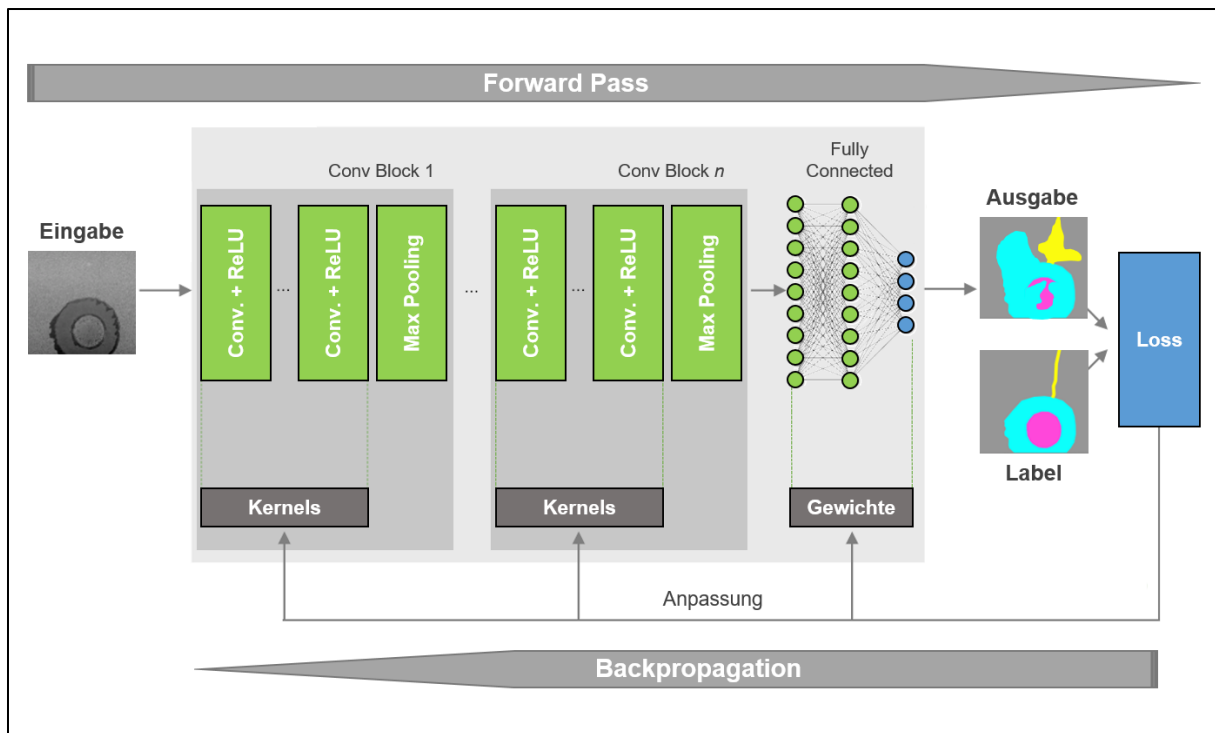


Abbildung 56: Überblick über die Architektur und das Training eines Convolutional Neural Network (nach YAMASHITA et al. 2018).

2.4.2.2 Evolution von CNN Architekturen

Das im Rahmen der vorliegenden Arbeit genutzte Deep-Learning-System nutzt wahlweise eine ResNet- oder eine Xception-Architektur als sogenannten CNN Backbone (siehe Kapitel 4.4.2). Aus diesem Grund werden nachfolgend ausgewählte CNN-Architekturen überblicksartig beschrieben, auf die ResNet und Xception aufbauen.

Das wohl bekannteste CNN ist das sogenannte AlexNet, weil es den Ausgangspunkt des CNN-Ansatzes im Bereich Computer Vision bildet. Das von KRIZHEVSKY, SUTSKEVER & HINTON (2012) entwickelte Netzwerk ist in Abbildung 57 dargestellt. Es besteht aus fünf Convolution-Schichten, von denen einige von Max-Pooling-Schichten gefolgt werden, sowie drei Fully-Connected-Schichten. Am Ende des Netzes steht eine tausendfache Softmax-Ausgabe, um die 1000 Klassen des ImageNet-Datensatzes (DENG et al. 2009) abzubilden. Insgesamt besteht dieses Netz aus 650000 Neuronen mit 60 Millionen Parametern, die im Zuge des Trainings angepasst werden müssen. Im ersten Convolution-Block werden 96 Filterkernel mit einer Größe von 11 x 11 Pixel optimiert. Im zweiten Block sind es bereits 256 Filterkernel mit einer Größe von 5 x 5 Pixel und in den folgenden Blöcken werden bis zu 384 Kernel mit einer Größe von 3 x 3 Pixel erzeugt. Die Anzahl der Kernel ist eine Design-Entscheidung.

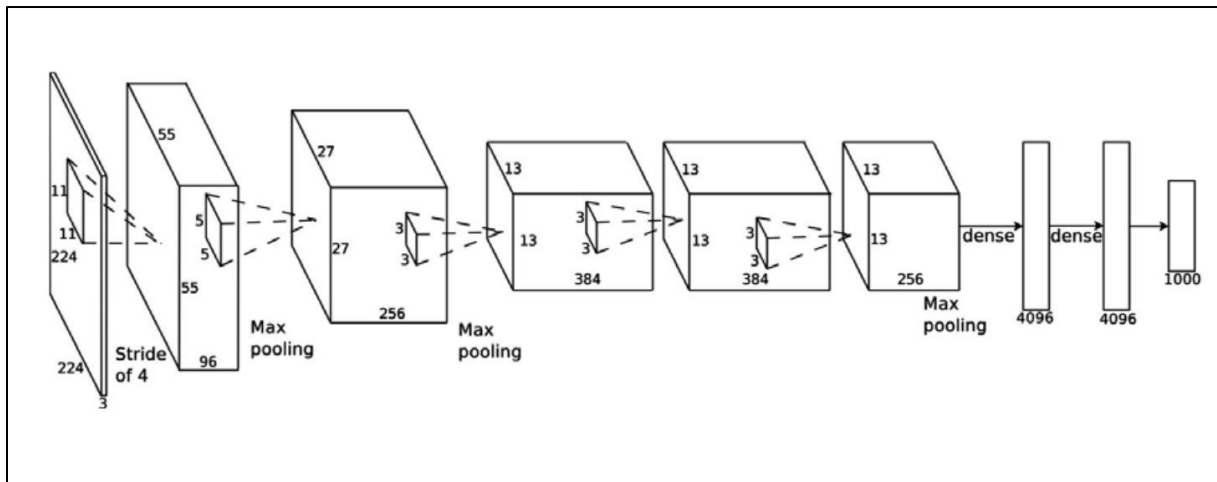


Abbildung 57: Schematische Darstellung der AlexNet-CNN-Architektur (SMIRNOV, TIMOSHENKO & ANDRIANOV 2014).

In den Jahren nach der Veröffentlichung der AlexNet-Architektur wurden immer tiefere Netze entwickelt, wie beispielsweise das VGG-Netz, das schon 16 bzw. 19 Schichten umfasste (SIMONYAN & ZISSERMAN 2015). Es gewann den ILSVRC-Wettbewerb im Jahr 2014. Im Unterschied zu AlexNet wurden hauptsächlich 3 x 3 Kernel in den Convolution-Schichten verwendet und aufgrund der hohen Anzahl von Schichten mussten verstärkt Regularisierungstechniken eingesetzt werden, um der steigenden Komplexität der Modelle entgegenzuwirken und Overfitting zu verhindern.

HE et al. (2016) stellten fest, dass das Ansteigen der Fehlerraten beim Versuch, noch tiefere CNN zu realisieren, auch mit Regularisierung nicht zufriedenstellend zu kontrollieren ist (siehe Abbildung 58). Die bestimmende Frage ihrer Forschung war, ob die Lösung für leistungsfähigere Netze lediglich darin bestehen kann, dass immer mehr Schichten eingebaut werden. Das grundsätzliche Problem bei tiefen und sehr tiefen künstlichen neuronalen Netzen ist, dass die Gradienten, die vom Optimierer im Backward Pass berechnet werden, um die Gewichte der Neuronenverbindungen anzupassen, zunehmend gegen Null tendieren, je weiter ein Neuron von der Ausgabeschicht entfernt ist. Somit können die Gewichte am Anfang des Netzes ab einer bestimmten Schichtanzahl nicht mehr effektiv optimiert werden. Dies wird als das Problem des verschwindenden Gradienten bezeichnet (engl.: Vanishing Gradient Problem) (CHOLLET 2021).

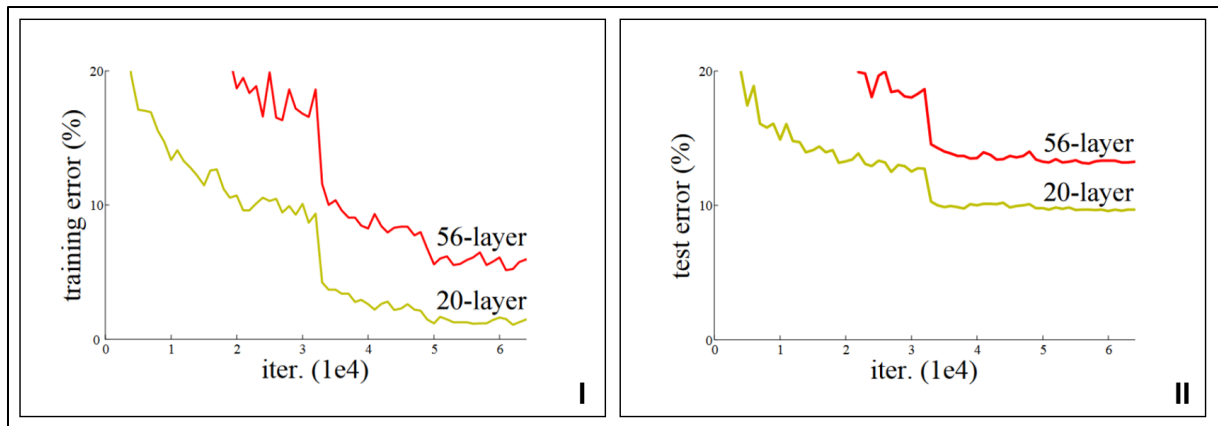


Abbildung 58: Trainings- (I) und Testfehler (II) eines CNN mit 20-Schichten und mit 56-Schichten. Mit zunehmender Tiefe werden Netze schwieriger zu trainieren, was sich in höheren Fehlerraten sowohl auf den Trainings- als auch auf den Testdaten äußert (nach HE et al. 2016).

Die grundlegende Idee von HE et al. (2016) aus dem Microsoft Research Team zur Lösung des Problems ist, dass sie in ihren Residual Neural Networks (ResNet) sogenannte Shortcut-Connections vorsehen. Das Netzwerk wird dabei in Blöcken von Faltungsschichten organisiert, wobei jeweils zwischen zwei aufeinanderfolgenden Blöcken eine zusätzliche Shortcut-Connection existiert. Dadurch wird der Gradientenfluss während des Trainings verbessert. Da Residual Networks außerdem weitestgehend auf Fully-Connected-Schichten verzichten, kommen sie, trotz größerer Tiefe, mit deutlich weniger Gewichten aus, als VGG-ähnliche Architekturen. Abbildung 59 zeigt zur besseren Darstellbarkeit ein ResNet mit nur 18 Schichten. ResNets mit mehr als 100 Schichten sind in der Praxis aber keine Seltenheit.

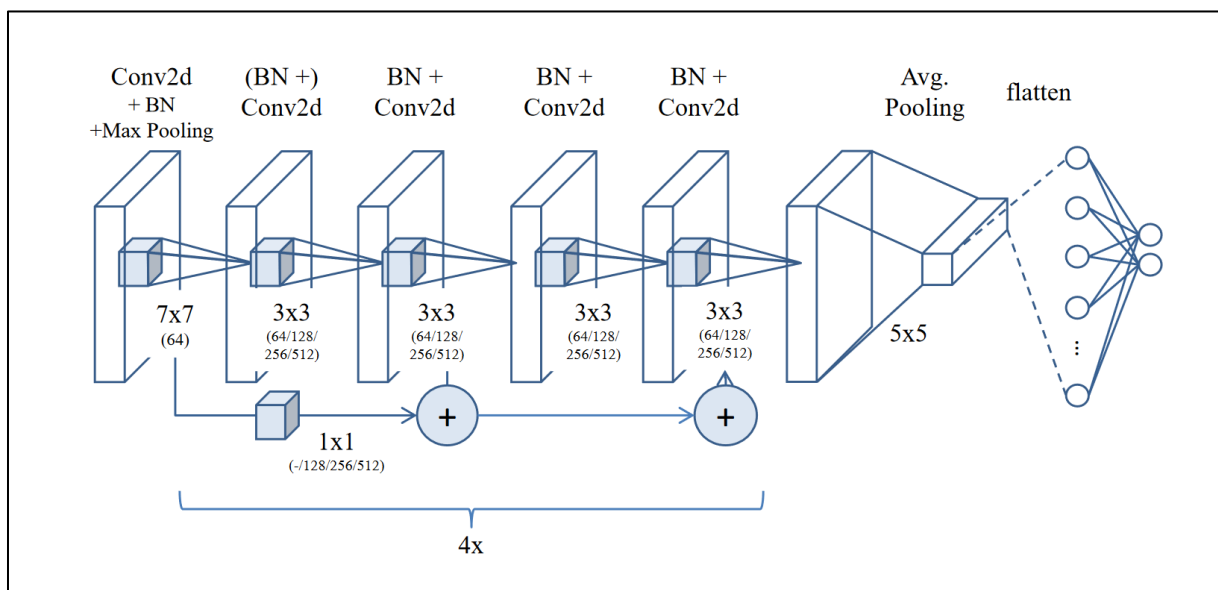


Abbildung 59: Struktur eines 18 Schichten umfassenden ResNets (STRICKER et al. 2019).

Ein wesentliches Element von Residual Neural Networks sind die sogenannten Residualblöcke, die im Beispiel aus zwei Abfolgen jeweils einer Convolution und einer Batch-Normalization-Schicht bestehen. Am unteren Rand der Abbildung 59 ist zu sehen, dass die Ausgaben des Vorgängerblocks immer mit der Ausgabe des aktuellen Blocks durch eine Additionsschicht verbunden sind (Shortcut-Connection). Diese eliminiert im Training das Problem des verschwindenden Gradienten.

Die vom Google Team entwickelte Xception-Architektur greift u. a. die Innovationen von AlexNet, VGG und ResNets auf (CHOLLET 2017). Der Name Xception steht für Extreme Inception und verweist auf die Nähe zu den ebenfalls bei Google entwickelten Vorgänger Architekturen der Inception-Reihe. Xception führt ein neues Konzept ein: die Depthwise Separable Convolution. Im Unterschied zu einer normalen Faltung teilt die Depthwise Separable Convolution die Faltung in zwei Teilprozesse auf: eine Depthwise und eine Pointwise Convolution. Anhand des Beispiels in Abbildung 60 lassen sich die unterschiedlichen Konzepte verdeutlichen. Bei einer gewöhnlichen Faltung wird eine Eingabe mit den Dimensionen $12 \times 12 \times 3$ (Höhe, Breite, Anzahl der Farbkanäle) mit einem Kernel von $5 \times 5 \times 3$ gefiltert. Dabei entsteht, bei Stride = 1 und ohne Zero Padding, eine Feature Map mit den Dimensionen $8 \times 8 \times 1$. Bei der Depthwise Convolution werden dagegen drei einzelne Kernels mit den Dimensionen $5 \times 5 \times 1$ verwendet. Jeder Kernel wird dabei nur auf einen der drei Farbkanäle des Eingabebildes angewendet und erzeugt drei Feature Maps, die gestapelt mit den Dimensionen $8 \times 8 \times 3$ vorliegen. Anschließend wird bei der Pointwise Convolution ein Kernel mit den Dimensionen $1 \times 1 \times 3$ verwendet, der aufgrund seiner Höhe und Breite auf jedem Pixel der $12 \times 12 \times 3$ Eingabe quasi punktwise arbeitet. Die Ausgabe hat die Dimensionen $8 \times 8 \times 1$. Die Standardfaltung und die Depthwise Separable Convolution führen zunächst also zum selben Ergebnis. Der Vorteil ist jedoch, dass durch den Zerlegungsprozess in der Depthwise Separable Convolution weniger Multiplikationen durchgeführt werden müssen. Dies führt angesichts der vielen Millionen Gewichte, die während des Trainings anzupassen sind, zu deutlich effizienteren Trainingsprozessen.

Mit der Xception-Architektur endet die Entwicklung von CNN nicht. Weitere wichtige CNN-Architekturen sind bei ALZUBAIDI et al. (2021) erklärt.

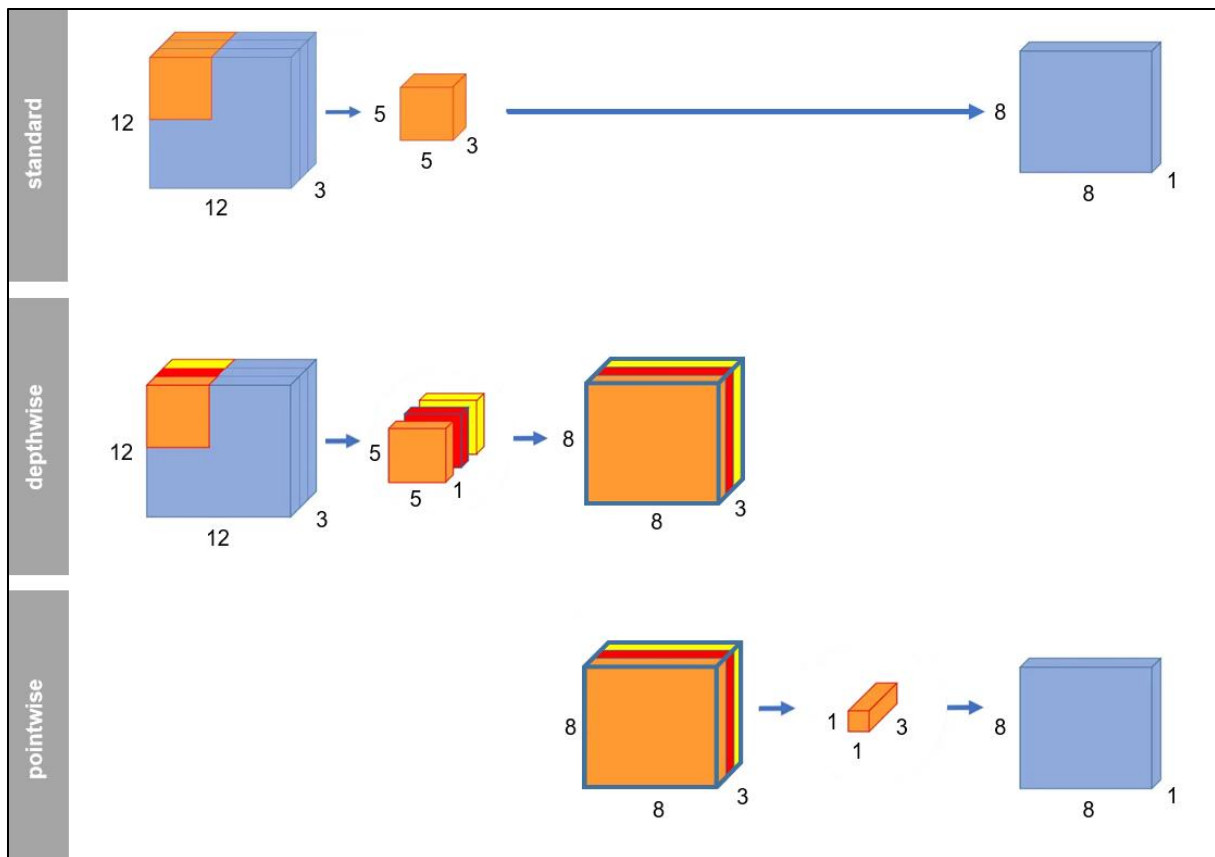


Abbildung 60: Prinzip der Depthwise Separable Convolution in Abgrenzung zur oben abgebildeten standardmäßigen Faltung (nach WANG 2018).

2.4.2.3 Fully Convolutional Networks

Die Grundidee, eine pixelweise Klassifikation mithilfe eines CNN zur semantischen Segmentierung eines Eingabebildes zu nutzen, wurde von LONG, SHELHAMER & DARRELL (2015) publiziert. Die Autoren schlugen mit ihrem Ansatz des sogenannten Fully Convolutional Networks (FCN) vor, die letzten Schichten eines klassischen CNN so zu modifizieren, dass sie nicht mehr nur eine Klassenentscheidung für das gesamte Bild ausgeben, sondern stattdessen einen entsprechenden Klassenwert für jede Pixelposition des Eingabebildes. Durch das Weglassen der für klassische CNN-Architekturen typischen Fully-Connected-Schicht am Ende des Netzes, kann eine FCN-Architektur beliebig dimensionierte Eingabebilder verarbeiten.

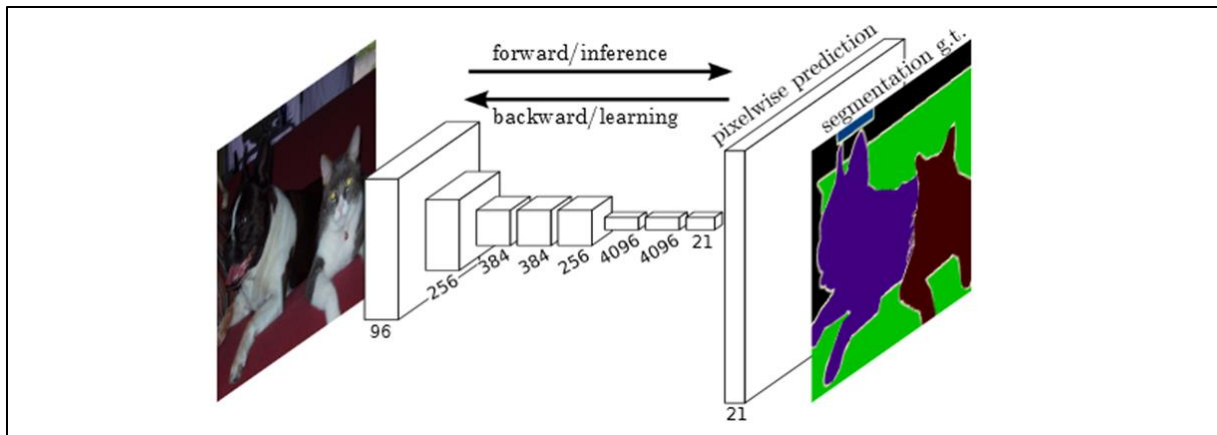


Abbildung 61: Schematische Darstellung eines Fully Convolutional Networks. Die Zahlen repräsentieren die Anzahl der Feature-Maps bzw. Filter pro Schicht (LONG, SHELHAMER & DARRELL 2015).

Wie Abbildung 61 schematisch zeigt, reduziert sich die Auflösung der Ausgabe-Feature-Maps im Zuge der Propagierung des Eingabebildes durch die mehrfach gestaffelten Faltungs- und Pooling-Schichten, weshalb die direkten Ausgaben des FCN in einer gegenüber dem Eingabebild deutlich verringerten Auflösung erfolgen. Dies wiederum führt zu relativ unscharfen Objektgrenzen, wenn die Prädiktionsmaske durch Upsampling wieder auf die Dimension des Eingabebildes zurück transformiert wird. Beim Upsampling handelt es sich um einen Prozess, bei dem leere Zeilen und Spalten zwischen den vorhandenen Matrixwerten eingefügt und anschließend mittels Interpolation aufgefüllt werden (siehe Abbildung 62).

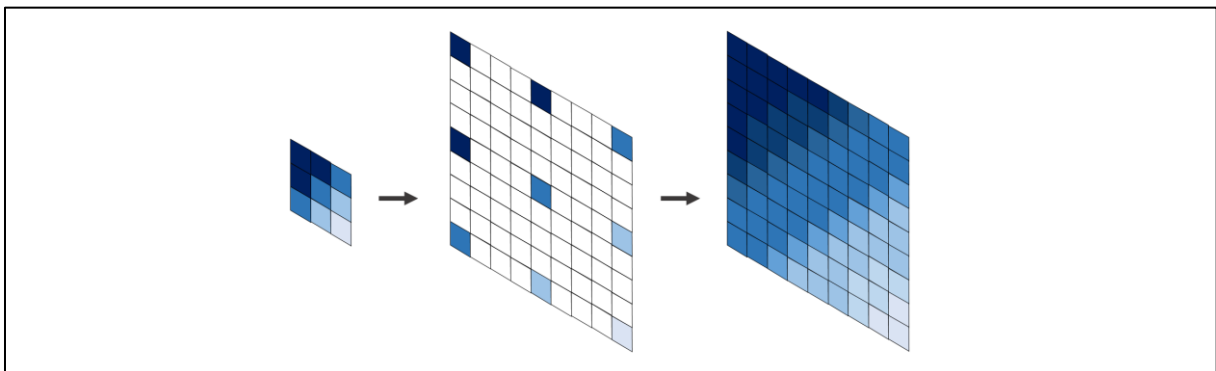


Abbildung 62: Prinzip des Upsamplings durch Einfügen leerer Pixel und anschließender Interpolation.

Um auch Objektgrenzen für feinere Objekte im Bild adäquat erhalten zu können, machten sich LONG, SHELHAMER & DARRELL (2015) bei der Implementierung des FCN Folgendes zu Nutzen: die Ausgabe von Schichten am Anfang des Netzes enthalten noch relativ viel Ortsinformation bzw. Information über den Verlauf von Objektgrenzen. Diese geht über die schrittweise Reduktion der Bildgröße im Netz immer mehr verloren. Dafür enthalten tiefere Schichten essenziellere semantische Information.

Die Experimente von LONG, SHELHAMER & DARRELL (2015) zeigten, dass durch die Fusion der Ausgaben flacherer und tieferer Schichten das Segmentierungsergebnis zugunsten einer besseren Lokalisierung von Objektgrenzen verbessert werden kann. In Abbildung 63 sind zur Verdeutlichung der Funktionsweise der FCN-Architektur lediglich die Eingabe- bzw. Ausgabeschichten sowie die Pooling-Schichten dargestellt. Die zwischen den Pooling-Schichten liegenden Faltungsschichten und auch die vollverschalteten Schichten sind hier weggelassen. Die durchgehende Linie repräsentiert das Basisnetz (FCN-32s), an dessen Ende eine Prädiktion mit einer noch groben Lokalisierung der Objektgrenzen steht. Das grobe Erscheinungsbild der Objektgrenzen resultiert daraus, dass die Ausgabe aus der Pooling-Schicht pool5 durch ein Upsampling mit einer Schrittweite von 32 auf die Größe des Eingabebildes zurücktransformiert werden muss. Für die beiden verfeinerten FCN-Netze FCN-16s und FCN-8s, die durch die gestrichelte und gepunktete Linie dargestellt sind, werden die Ausgaben der Pooling-Schicht pool3 und pool4 durch elementweise Addition mit den Ausgaben des Basisnetzes kombiniert. Dabei werden für das Upsampling kleinere Schrittweiten von 16 und 8 verwendet, was zu einem höheren Detailgrad der Segmentierungsergebnisse führt, wie im rechten oberen Teil der Abbildung am Beispiel des Radfahrers zu sehen ist.

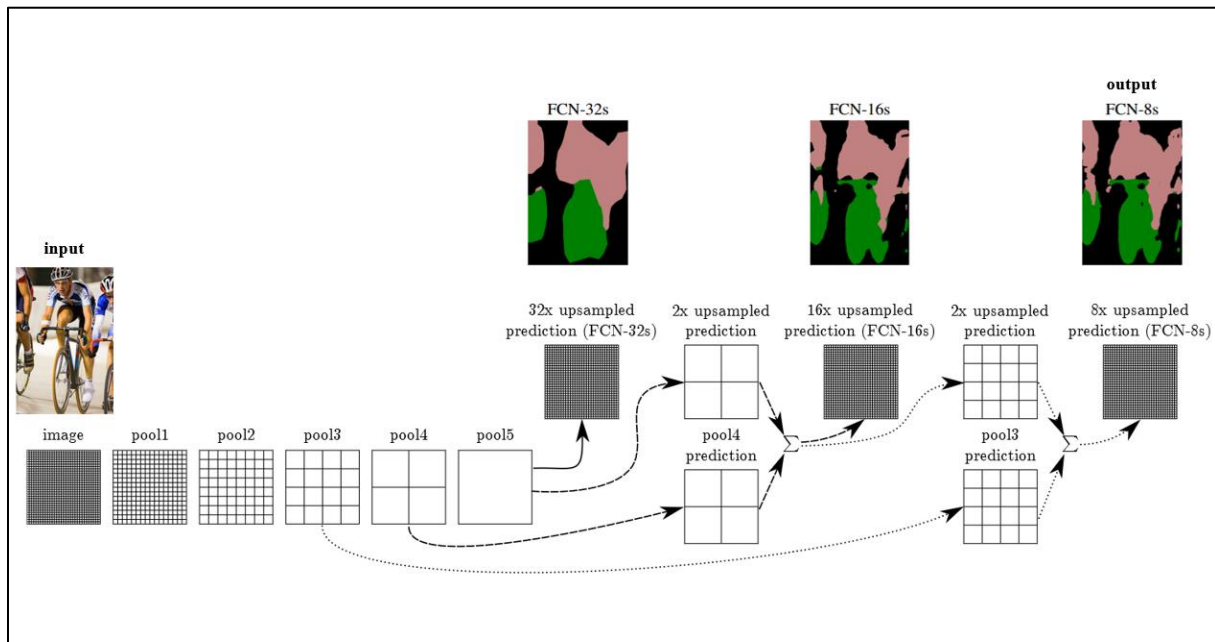


Abbildung 63: Schematische Darstellung der Kombination von Ausgaben an unterschiedlich tiefen Stellen im Netz. Die Kombination flacherer und tieferer Schichten führen zu einer besseren Segmentierung (nach LONG, SHELHAMER & DARRELL 2015).

2.4.2.4 Encoder/Decoder Networks

Eine Weiterentwicklung des Ansatzes von LONG, SHELHAMER & DARRELL (2015) ist das sogenannte SegNet von BADRINARAYANAN, KENDALL & CIPOLLA (2017) und soll hier beispielhaft für die Gruppe der Encoder/Decoder-Netzwerke stehen, zu denen auch das häufig verwendete U-Net von RONNEBERGER, FISCHER & BROX (2015) oder das PSP-Net von ZHAO et al. (2017) zählen. Encoder und Decoder beschreiben hierbei jeweils eine Makro-Funktion im künstlichen neuronalen Netzwerk: Der Encoder-Teil ist dafür zuständig, ein Eingabebild mithilfe eines CNN in eine dichte Repräsentation mit niedriger Dimensionalität zu überführen. Die Ausgabe des Encoder-Teils erlaubt bereits theoretisch eine semantische Segmentierung, jedoch aufgrund der schrittweisen Dimensionsreduktion zulasten von Details der Bildinformation. Daher wird im Anschluss ein Decoder-Teil genutzt, um aus den letzten Aktivierungen des Encoder-Teils mittels Upsampling-Techniken die Eingabebildauflösung wiederherzustellen.

Mit ihrer SegNet-Architektur entwickeln BADRINARAYANAN, KENDALL & CIPOLLA (2017) einen neuen Ansatz, Merkmale mit geringer Auflösung auf die Eingabeauflösung abzubilden, um die Genauigkeit der Lokalisierung von Objektgrenzen im Segmentierungsergebnis weiter zu verbessern. Wie in Abbildung 64 dargestellt, verzichtet auch SegNet auf Fully-Connected-Schichten und ist im vorderen Teil aus Faltungs- und Pooling-Schichten aufgebaut. In diesem Encoder genannten Teil des Netzes findet, analog der Funktionsweise eines CNN, eine schrittweise Verdichtung der Bildinformation durch Faltung und Pooling statt. Die Faltungsschichten entsprechen dabei den ersten 13 Faltungsschichten der VGG-Architektur von SIMONYAN & ZISSERMAN (2015). Jede Encoder-Schicht hat eine entsprechende Decoder-Schicht und somit hat das Decoder-Netzwerk ebenfalls 13 Schichten. Im Decoder-Teil wird ein Upsampling seiner Eingabe unter Verwendung der übertragenen Pooling-Indizes aus dem Encoder durchgeführt. Anschließend folgen Faltungen mit trainierbaren Kernels, um die Feature Maps schrittweise wieder auf die Auflösung des Eingabebildes zurück zu bringen. In einem letzten Schritt werden die Feature Maps der Decoder-Ausgabe an eine Softmax-Klassifikator-Schicht zur pixelweisen Klassifizierung übergeben. Im Unterschied zur FCN-Architektur, bei der die Filter im Zuge des Upsamplings fest implementiert sind, können die Filter im Decoder-Teil der SegNet-Architektur während des Trainings optimiert werden.

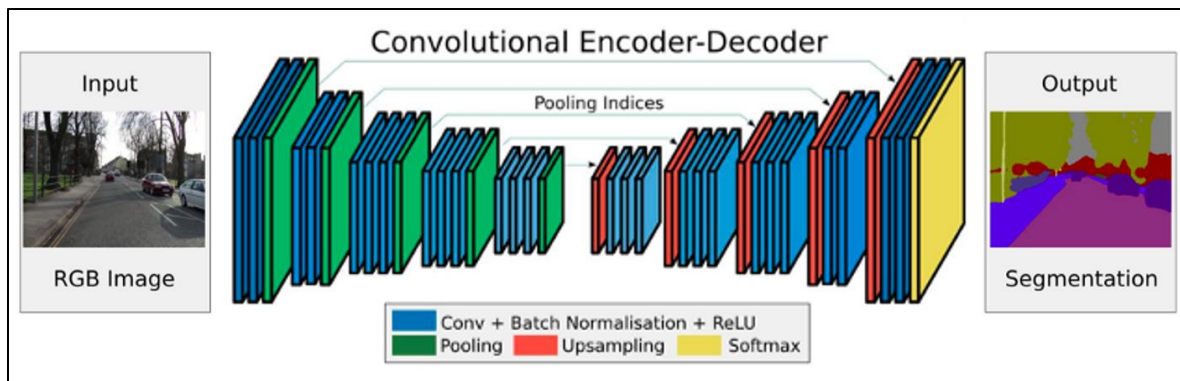


Abbildung 64: Schematische Darstellung der SegNet-Architektur (BADRINARAYANAN, KENDALL & CIPOLLA 2017).

Der Vorteil von SegNet gegenüber U-Net liegt darin, dass nicht alle Feature Maps der Encoder-Schichten zu den entsprechenden Decoder-Schichten transferiert werden, sondern nur die Pooling-Indizes, wodurch SegNet hinsichtlich des Speicherplatzes viel effizienter arbeitet.

Die in Kapitel 4.4 beschriebenen DeepLab-Architekturen greifen verschiedene Konzepte der SegNet-, U-Net- und PSP-Net-Architekturen auf und kombinieren bzw. ergänzen diese zu einem leistungsfähigen Encoder/Decoder-Netz, das in verschiedensten Anwendungsdomänen erfolgreich eingesetzt wird.

2.4.3 Transfer Learning

Wie in den Kapiteln 2.4.1 und 2.4.2 dargestellt, ermöglichen viele und diverse Daten das automatisierte Erlernen einer auf die Problemstellung angepassten Extraktion relevanter Merkmale in künstlichen neuronalen Netzen. Deshalb eignen sich Deep-Learning-Ansätze vor allem dann zur Bearbeitung von Fragestellungen, wenn eine große Menge an Daten verfügbar ist. In der Praxis ist es jedoch nicht immer möglich, auf entsprechend umfangreiche und qualitativ hochwertig annotierte Daten zugreifen zu können. Aus diesem Grunde wurden Techniken entwickelt, die es ermöglichen, auch auf kleineren Datensätzen tiefe neuronale Netze zu trainieren (YAMASHITA et al. 2018). Dazu zählen Data Augmentation und Transfer Learning. Bereits in Kapitel 2.4.1.3 wurde Data Augmentation im Rahmen von Regularisierungsstrategien als eine Möglichkeit beschrieben, robustere Modelle zu trainieren, indem die verfügbare Datenbasis künstlich erweitert wird.

In diesem Kapitel liegt der Fokus auf dem Transfer Learning. Dabei handelt es sich um ein zentrales Konzept, bei dem Wissen über eine Ursprungsdomäne auf eine Zieldomäne übertragen wird (WEISS, KHOSHGOFTAAR & WANG 2016). Damit wird es möglich, Deep-Learning-Verfahren auch dann zu nutzen, wenn semantisch annotierte Daten nur in begrenztem Umfang

verfügbar sind. Dieser Strategie liegt die Annahme zugrunde, dass generische Merkmale, die auf einem sehr großen Datensatz gelernt wurden, übergreifend nutzbar sind. Wie in Abbildung 65 zu sehen ist, werden in den vorderen Schichten eines CNN universelle Merkmale, wie Kanten, Ecken oder Gradienten extrahiert, die unabhängig von der Anwendungsdomäne bei der semantischen Interpretation von Eingabedaten bedeutsam sind (ZEILER & FERGUS 2014). Deshalb wird beim Transfer Learning davon ausgegangen, dass diese Merkmale auch für die Zielanwendung relevant sein könnten und nur die spezifischen Merkmale in tieferen Schichten explizit auf die Daten der Zieldomäne angepasst werden müssen. Diese Übertragbarkeit von erlernten generischen Merkmalen ist ein wesentlicher Vorteil des Deep Learning, der sich bei verschiedenen Aufgaben mit kleinen Datensätzen als hilfreich erwiesen hat (YAMASHITA et al. 2018).

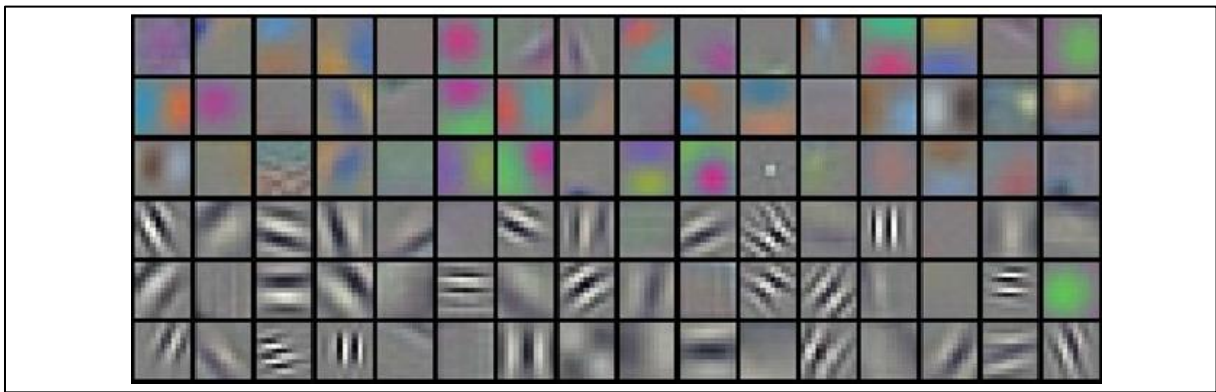


Abbildung 65: Visualisierung der 96 Faltungskernels, die in der ersten Schicht der AlexNet-Architektur gelernt werden (SMIRNOV, TIMOSHENKO & ANDRIANOV 2014).

Im Bereich Computer Vision wurden verschiedene Benchmarkdatenätze aufgebaut, die es Forschern und wissenschaftlichen Gruppen erlauben, neue Ansätze und Algorithmen zu entwickeln und vergleichend zu testen. Dazu zählen beispielsweise ImageNet (DENG et al. 2009) und COCO (LIN et al. 2014), die mehr als eine Millionen Bilder und Annotationen für 1000 Klassen bzw. 300000 Bilder und Annotationen für 80 Klassen umfassen. Selbst vergleichsweise kleine Benchmarkdatensätze, wie PASCAL VOC (EVERINGHAM et al. 2015) oder CityScapes (CORDTS et al. 2016) enthalten 15000 bzw. 5000 annotierte Bilder. Auf Basis solcher Datensätze wurden im Rahmen wissenschaftlicher Wettbewerbe, wie beispielsweise der ILSVRC, innovative Ansätze im Bereich Deep Learning entwickelt. Typischerweise werden der Quellcode sowie die trainierten Modelle zusammen mit einem Fachartikel veröffentlicht. Es ist prinzipiell möglich, diese vortrainierten Modelle zu nutzen und auf andere Fragestellungen hin anzupassen, sofern die jeweiligen Lizenzmodelle dies nicht untersagen.

Im einfachsten Fall wird ein auf den ImageNet Datensatz vortrainiertes CNN-Modell ohne Modifikationen lediglich auf Bilder eines anderen Kameratyps angewendet (siehe Abbildung 66 links). In der Regel werden in der Zielanwendung aber andere Klassen benötigt, als beispielsweise im ImageNet-Datensatz vorkommen. Deshalb besteht die Möglichkeit, die Gewichte in den Schichten des CNN, die die Merkmalsextraktion repräsentieren, unangetastet zu lassen (engl.: Freeze) und nur die letzte vollverschaltete Schicht des Netzes zu trainieren (siehe Abbildung 66 Mitte). Dabei werden nur die Gewichte dieser letzten Schicht zufällig initialisiert und dann im Training angepasst. Der Begriff Fine-Tuning beschreibt im Transfer-Learning-Kontext, dass zusätzlich zur Klassifikationsschicht auch Teile der Merkmalsextraktion oder die Merkmalsextraktion als Ganzes nachtrainiert werden (siehe Abbildung 66 rechts). Dabei werden beispielsweise auch die Filter der Faltungs- und Poolingschichten auf die Daten der Zielanwendung hin angepasst. In der Regel werden eher tiefere Schichten der Merkmalsextraktion beim Fine-Tuning-Ansatz nachtrainiert, um von den vortrainierten generischen Merkmalen der vorderen Schichten zu profitieren.

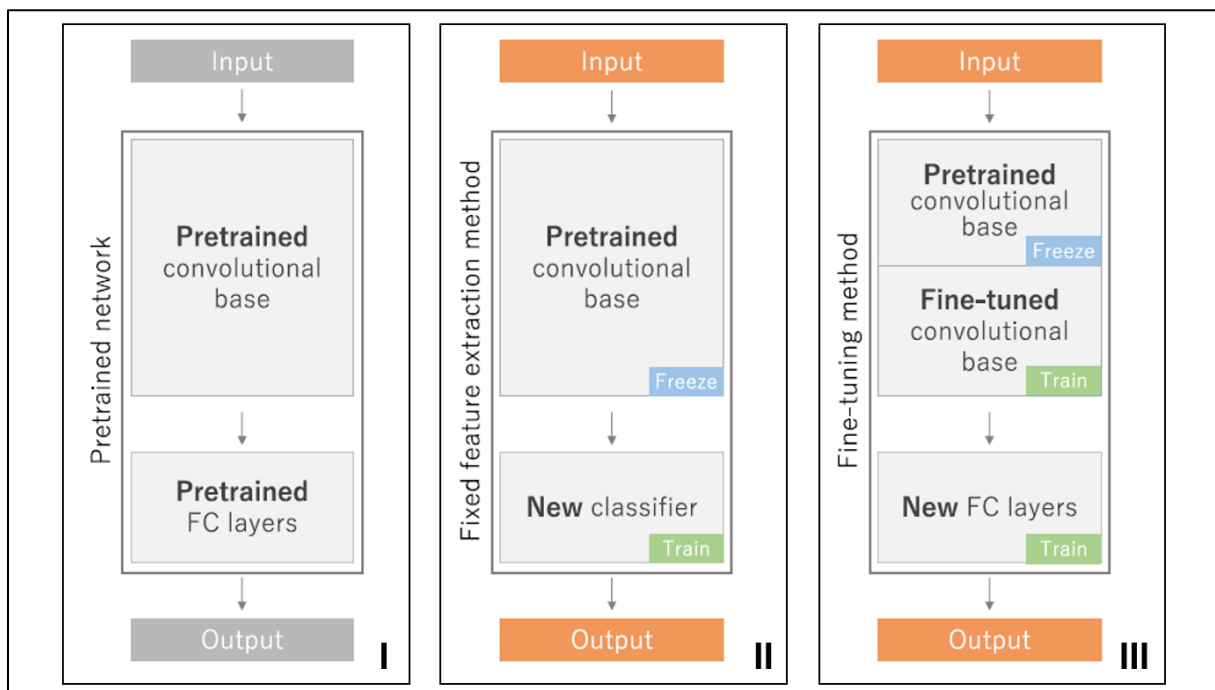


Abbildung 66: Verschiedene Transfer-Learning-Arten: Verwenden eines vortrainierten Netzes ohne jegliche Anpassung (I), Verwenden einer vortrainierten Merkmalsextraktion mit Anpassung der vollverschalteten Schichten (II) und Verwenden einer vortrainierten Merkmalsextraktion mit Fein-Anpassung von Teilen derselben und Anpassung der vollverschalteten Schichten (III) (nach YAMASHITA et al. 2018).

Transfer Learning ist ein mächtiges Werkzeug, auch im Hinblick auf die Demokratisierung der Deep-Learning-Technologie, da sie es ermöglicht, auch mit begrenzten Daten und Hardware-Ressourcen leistungsfähige Modelle zu trainieren. Transfer Learning sollte jedoch nicht blind

genutzt werden. Wie vorstehend beschrieben, ist eine Auseinandersetzung damit sinnvoll, welcher Teil des Wissens überhaupt übertragen werden kann, um zur gewünschten Leistungssteigerung in der Zieldomäne beizutragen. Damit verbunden ist auch die Frage danach, worin sich die Problemstellungen von Quell- und Zieldomäne überhaupt ähneln. Auch eignen sich nicht alle Architekturen gleichermaßen gut für Transfer Learning (KORNBLITH, SHLENS & LE 2019; ZHANG & DAVISON 2020). ResNet- und Xception-Modelle sind im Allgemeinen geeignet für eine Transfer-Learning-Strategie.

2.4.4 Geospatial Artificial Intelligence

GeoAI ist ein Kompositum aus den Worten Geospatial und Artificial Intelligence und beschreibt den Einsatz von Methoden der Künstlichen Intelligenz, einschließlich Machine Learning und Deep Learning, um Wissen durch die Analyse von räumlichen Daten zu erzeugen (siehe Abbildung 67). ALASTAL & SHAQFA (2022) beschreiben GeoAI als eine noch junge und in der Entwicklung befindliche multidisziplinäre Fachrichtung, die Informatik, Ingenieurwesen, Statistik und Geowissenschaften umfasst.

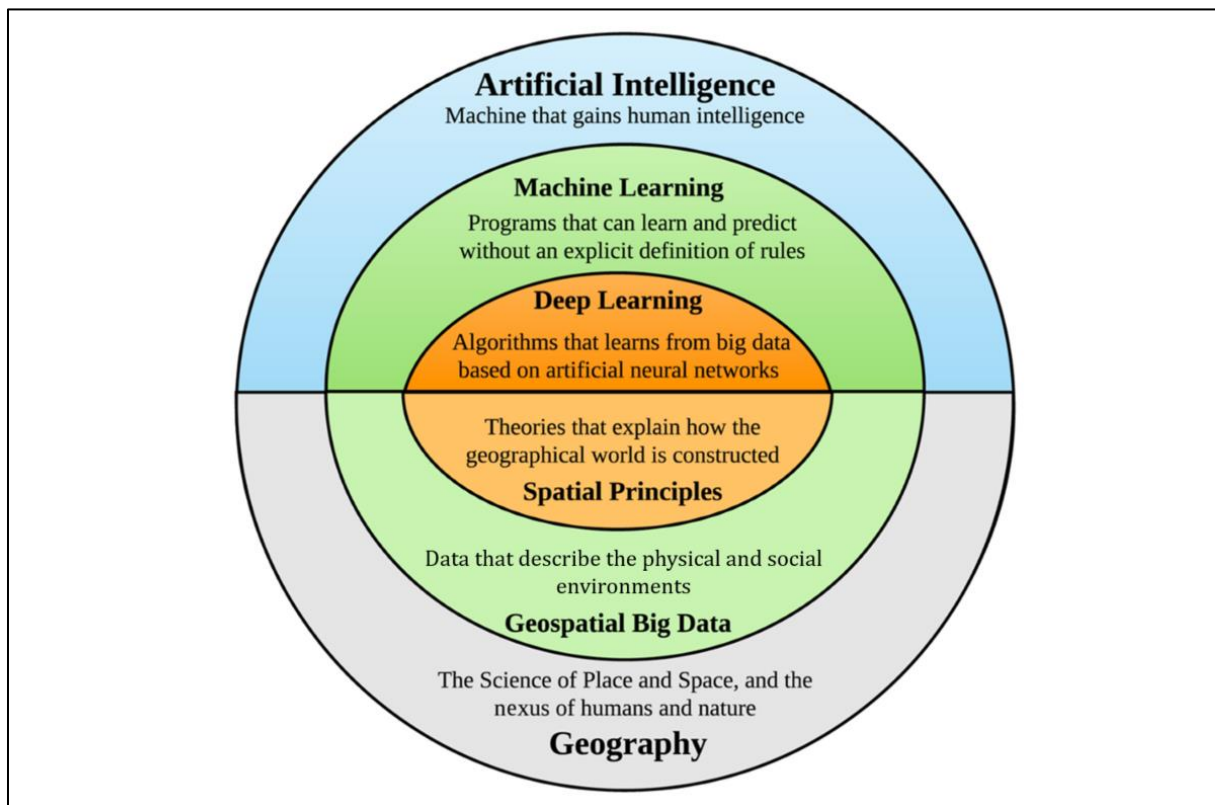


Abbildung 67: Schema zur Veranschaulichung, wie die Disziplinen Geographie und künstliche Intelligenz in der Fachrichtung GeoAI ineinandergreifen (LI & HSU 2022).

Als Begriff wird GeoAI seit dem Jahr 2017 verwendet (LI & HSU 2022). Dabei ist der Einsatz von KI-Techniken in der Geographie und den Geowissenschaften nicht neu. Bereits in den 1980er und 1990er-Jahren beschäftigten sich Forscher damit, wie Künstliche Intelligenz für geographische Problemlösungen verwendet werden könnte (SMITH 1984; OPENSHAW & OPENSHAW 1997). Analog zu den Entwicklungen auf dem Feld der künstlichen neuronalen Netze, die zeitweise stagnierte und erst seit den 2010er-Jahren wieder einen immensen Innovationsschub verzeichnet, verlief auch die Entwicklung im Bereich GeoAI. JANOWICZ et al. (2020) begründen dies nicht ausschließlich mit den rasanten Weiterentwicklungen auf Seiten der Hardware im Consumer-Bereich (leistungsfähige Grafikkarten), der Entwicklung neuer Methoden (Transfer Learning) oder der on-demand Verfügbarkeit sehr großer Rechenkapazitäten über Cloud-Dienste wie Amazon Web Services oder Microsoft Azure. Die Autoren sehen einen weiteren Einflussfaktor für den jüngsten Erfolg von KI-Techniken in einer neuen Kultur der Datenerstellung und der gemeinsamen Datennutzung. Beispielhaft sei hier das Social Sensing genannt, also die Nutzung von Daten, die aktiv oder passiv von Menschen über körpernahe Geräte wie Smartphones oder Fitness-Tracker abgegeben werden.

Im Zeitalter von Big Data werden täglich 2,5 Exabyte an Daten produziert (VOPHAM et al. 2018). Es wird postuliert, dass bis zu 80 % dieser Daten einen Raumbezug haben (LI et al. 2016). Auch wenn dieser Wert diskutiert werden darf, so verdeutlicht er doch, dass es einen hohen Bedarf nach automatisierter räumlicher Datenverarbeitung gibt, der weit über die klassischen GIS- und Fernerkundungsanwendungen hinaus geht. KI und insbesondere Deep Learning haben bereits in vielen Anwendungsbereichen ihre Leistungsfähigkeit unter Beweis gestellt, in denen große Datenmengen bewältigt werden müssen – von der Verbesserung der Customer-Experience durch Chatbots über Assistenzsysteme in Fahrzeugen bis hin zu kommerziell verfügbaren Robotern, wie Spot von Boston Dynamics. Angesichts räumlich immer hochauflösenderer Sensoren und zeitlich immer dichterere Datenreihen rückt das Thema KI in den letzten Jahren auch wieder verstärkt in den Fokus der Geowissenschaften (GRUNAU 2022). Das äußert sich nicht zuletzt darin, dass Deep-Learning-Modelle mittlerweile fester Bestandteil von GIS-Anwendungen sind. So kann ArcGIS beispielsweise Modelle auf Basis gängiger Deep Learning Frameworks oder Bibliotheken wie TensorFlow, CNTK, Pytorch, ScitLearn oder auch IBM Watson integrieren.

Aktuelle Anwendungsgebiete von GeoAI sind unter anderem die geomorphologische Kartierung bzw. das Erkennen von Geländeformen (GUYOT, LENNON & HUBERT-MOY 2021), die Modellierung fluvialer Systeme (GONZALES-INCA et al. 2022), die Kartierung von Gebäudegrundrissen in dicht bebauten Gebieten (XIE et al. 2020), die Extraktion von

Informationen aus digitalisierten historischen Karten (DUAN et al. 2020), die semantische Segmentierung von Mobile-Mapping-Daten (REITERER et al. 2020) und Erzeugung digitaler Zwillinge (DÖLLNER 2020), Verkehrsmodellierungen (POLSON & SOKOLOV 2017) und die räumliche Interpolation (ZHU et al. 2020). Bei LI & HSU (2022) ist eine Zusammenfassung von GeoAI-Anwendungen im Bereich fernerkundlicher Bildanalyse zu finden. Auch die Untersuchungen der vorliegenden Arbeit können dem Gebiet der GeoAI zugeordnet werden.

2.5 Forschungsstand zur Automatisierung im Bereich Straßenzustandserfassung

In diesem abschließenden Kapitel des Grundlagenteils wird der Stand der Forschung im Bereich automatisierter Straßenzustandserfassung zusammengefasst und dabei die Lücke aufgezeigt, die mit den Forschungsfragen der vorliegenden Arbeit adressiert wird: Im Bereich der kommunalen Straßenzustandserfassung werden MLS-Daten bislang nur für den Teilbereich Ebenheit genutzt. Eine automatisierte Ebenheitsbewertung ist simpel zu realisieren, da sich die entsprechenden Indikatoren, gemäß der Definition im jeweiligen Standard, durch geometrische Operationen oder Frequenzanalysen klassisch berechnen lassen (siehe Kapitel 2.5.1). So ist beispielsweise im amerikanischen Standard ASTM E1703 /E1703M-10 (2015) festgelegt, wie Spurrinnentiefen ermittelt werden. In Europa ist die Norm DIN EN 13036 (2020) für die Erfassung und Analyse von Straßenlängs- und Querprofilen maßgeblich, auf die sich wiederum die für Deutschland relevanten Regelwerke beziehen (FGSV 2006, 2012).

Oberflächenschäden werden im deutschen Straßennetz typischerweise aus Kamerabilddaten detektiert. Hierfür existieren verschiedenste Automatisierungsansätze (siehe Kapitel 2.5.2).

Eine automatisierte Erfassung von Schäden aus MLS-Daten würde, zusammen mit der bereits gelösten automatisierten Erfassung und Bewertung der Ebenheit, eine ganzheitliche messtechnische Zustandsbewertung für die Erfordernisse des kommunalen Straßennetzes aus einem gemeinsamen aktiven Sensor ermöglichen. Seit dem Jahr 2012 wurde in Deutschland nur einem Laserscanner eine Betriebszulassung seitens der BAST zur Erfassung von Ebenheiten gemäß der technischen Prüfvorschrift TP Eben (FGSV 2009a) erteilt und seither jährlich erneuert (siehe Kapitel 3.1.1.2). Bisher wurde noch nicht systematisch untersucht, ob auch eine automatische Erfassung von Oberflächenschäden aus den Messdaten dieses Laserscanners gemäß der Richtlinie E EMI (FGSV 2012) möglich ist.

2.5.1 Ebenheitsindikatoren aus Mobile Laserscanning Daten

Im Kontext der Straßenerhaltung werden MLS-Daten vor allem zur Erfassung der Quer- und Längsebenheit genutzt (GRÄFE 2009; REITERER et al. 2013; HAIGERMOSER et al. 2015; MIRALIAKBARI 2017; LUNDBERG et al. 2018; GÉZERO & ANTUNES 2019; GROßMANN et al. 2023).

Für Ebenheitsanalysen auf Basis von MLS-Daten werden Quer- und Längsprofile in der Regel nicht direkt aus der rohen 3D-Punktwolke abgegriffen, sondern es wird zunächst ein Oberflächenmodell auf Basis des Laserscans berechnet (GROßMANN et al. 2023). Der Grund hierfür ist, dass die damit einhergehende Glättung der Messdaten potentiell dazu führt, dass relative Messungen auf dem Oberflächenmodell in Relation zur Messunsicherheit der rohen Laserscans mit einer höheren Präzision möglich sind (GRÄFE 2009).

Aufgrund der in Kapitel 2.3.1.1 beschriebenen Eigenschaften des CRG-Formates, wird diese Art der Oberflächenmodellierung u. a. auch bei SKALECKI et al. (2021) und GENSER et al. (2022) im Rahmen von Ebenheitsanalysen verwendet. In beiden Fällen wird die CRG-Referenzlinie auf Basis der Fahrzeugtrajektorie automatisch mithilfe von Konstruktionselementen generiert. Um eine regelwerkskonforme Auflösung und Genauigkeit der Höhenprofile gemäß dem Regelwerk TP Eben (FGSV 2009a) zu erreichen, wird eine CRG-Gitterweite von 0,1 m x 0,1 m verwendet. Das so in der XY-Ebene konstruierte CRG-Grundgerüst wird anschließend mit der 3D-Punktwolke des Laserscanners räumlich verschnitten, wobei eine distanzgewichtete Mittelung der Höhenwerte in einem 0,1 m Radius stattfindet. Aufgrund der Matrixstruktur des CRG-Formates kann auf Höhenwerte eines Querprofils (Zeilen) und Höhenwerte eines Längsprofils (Spalten) zugegriffen werden. In der Folge können die relevanten Ebenheitsindikatoren berechnet werden.

Abbildung 68 zeigt exemplarisch die Berechnung der Spurrinntentiefe und fiktiven Wassertiefe auf einem CRG-Querprofil mit 0,1 m Punktabstand. Die Berechnungsvorschrift der Indikatoren ist in der Norm DIN EN 13036 (2020) bzw. der technischen Prüfvorschrift TP Eben (FGSV 2009a) festgelegt. Durch trigonometrische Berechnungen können alle für das kommunale Straßennetz relevanten Zustandsgrößen ermittelt werden (siehe Kapitel 2.1.3.1).

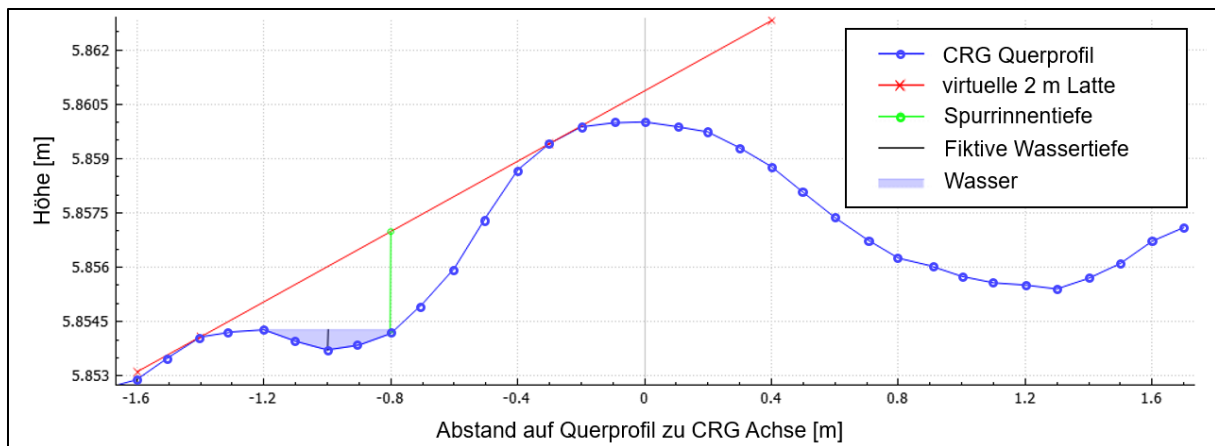


Abbildung 68: Visualisierung der Berechnung von Querebenenheitsindikatoren auf einem Laserscan-basierten CRG-Querprofil. Die Spurrinntiefe links beträgt hier 2,8 mm und die fiktive Wassertiefe 0,5 mm.

Im Jahr 2012 erhielt erstmals ein Laserscanner eine zeitbefristete Betriebszulassung der Bundesanstalt für Straßenwesen (BASt) für die Erfassung von Querebenenheitsdaten im Rahmen der Zustandserfassung auf Autobahnen (REITERER et al. 2013). Dabei handelte es sich um einen Pavement Profile Scanner (PPS) der ersten Generation (siehe Kapitel 3.1.1.2). In einer umfangreichen Studie stellten LUNDBERG et al. (2018) heraus, dass mit dem PPS erfasste Daten hinsichtlich Genauigkeit und Wiederholbarkeit mit bisher in skandinavischen Ländern eingesetzten berührungslosen Technologien, wie Triangulationslaser- oder Laserlichtschnittsensoren, vergleichbar sind und den Anforderungen an eine netzweite Querebenenheitserfassung gerecht werden. In den Jahren 2018 bis 2022 führte die BASt eine eigene Studie mit fünf verschiedenen Messsystemen durch, um neue Erfassungstechnologien im Bereich der berührungslosen Ebenheitsmesssysteme mit den bisherigen Technologien zu vergleichen (GROßMANN et al. 2023). Die Erkenntnisse von LUNDBERG et al. (2018) hinsichtlich der Querebenenheit konnten im Rahmen dieses BASt Forschungsprojektes bestätigt werden. Darüber hinaus kam die Studie zu dem Schluss, dass auch die auf dem PPS-Laserscan basierende Längsebenenheitsbewertung mit den Ergebnissen der bisher zugelassenen berührungslosen Methoden (Triangulationslasersensoren) vergleichbar ist. Die Ergebnisse stellen in Aussicht, dass künftig die technische Prüfvorschrift TP Eben (FGSV 2009a) aktualisiert und auch MLS-Systeme zur Erfassung der Längsebenenheit durch die BASt zertifiziert werden könnten.

Es ist festzuhalten, dass die Fragestellung der messtechnischen Gebrauchswertbestimmung (Ebenenheit) mithilfe von PPS-Laserscans im Kontext kommunaler Straßenzustandserfassung als gelöst betrachtet werden kann. Der Aspekt Ebenheit aus MLS-Daten wird in der vorliegenden Arbeit nicht weiter behandelt.

2.5.2 Automatisierte Erfassung von Oberflächenschäden

Das Themenfeld der automatisierten Straßenzustandserfassung ist Gegenstand zahlreicher Untersuchungen und Studien. In internationalen Publikationen wird hier oftmals der Begriff Road Condition Monitoring verwendet. Auch wenn der Straßenzustand durch weit mehr als nur durch Risse charakterisiert wird, beschäftigen sich der überwiegende Teil der Publikationen mit der sogenannten Crack Detection (MUNAWAR et al. 2021). Eine umfassende und systematische Literaturübersicht über Technologien aus dem Themenfeld Road Condition Monitoring ist bei RANYAL, SADHU & JAIN (2022) zu finden.

Die bisherigen Ansätze, eine automatisierte Schadenserkenkung auf Straßenoberflächen zu realisieren, unterscheiden sich zum einen hinsichtlich der verwendeten Sensorik und zum anderen hinsichtlich der verwendeten Analysemethoden (MIRALIAKBARI 2017; MOHAN & POOBAL 2018).

Bezüglich der verwendeten Sensoren kommen unter anderem Beschleunigungs-, Infrarot-, Ultraschall-, Laser-, Radarsensoren sowie Video- oder Einzelbildkameras (Mono oder Stereo) bei der Erfassung von Schadstellen zum Einsatz (KIM & RYU 2014; SESSELMANN 2018; DE BLASIIS, DI BENEDETTO & FIANI 2020; BRUNKEN 2021; RANYAL, SADHU & JAIN 2022). Typischerweise nutzen Dienstleister im Rahmen der Zustandserfassung und -bewertung von Straßen in Deutschland für die Detektion von Oberflächenschäden kamerabildbasierte Ansätze (FGSV 2006, 2012, 2015; SCHREIBER 2020).

In der globalen Perspektive ist das Laser Crack Measurement System (LCMS) der Firma Pavemetrics am weitesten verbreitet (LAURENT et al. 2012; PETITCLERC, LAURENT & HABEL 2018; DRENTH et al. 2020). Es arbeitet nach dem Laserlichtschnittverfahren und bietet neben der Sensorik auch die automatische Analysesoftware LcmsRoadInspect, die Oberflächenschäden wie Risse, Ausbrüche, Ausmagerungen und Flickstellen erkennen kann. Da es sich um eine proprietäre Software handelt, sind keine Details zu den verwendeten Algorithmen veröffentlicht. LAURENT (2010), LAURENT et al. (2012) und TSAI (2018) zufolge werden sowohl klassische Bildanalysesalgorithmen (Risserkennung), als auch klassisches Machine Learning (Ausmagerungen) zur Analyse der Sensordaten verwendet. Verschiedenste Dienstleister, wie beispielsweise SITECO (2022), NIEVELT LABOR GMBH (2022), LEHMANN+PARTNER POLSKA (2022), INFRALAB (2019), FUGRO N.V. (2022) oder DYNATEST US INC. (2022), haben das LCMS in ihre Mobile-Mapping-Systeme integriert. Während beispielsweise Oberflächenschäden auf den Nationalstraßen Polens seit dem Jahr 2012 mit dem LCMS automatisch erfasst und in Anlehnung an das deutsche Regelwerk ZTV ZEB-Stb (FGSV 2006) ausgewertet werden (SKAKUJ et al. 2019), verhindert die bislang in den LCMS-Sensoren

verwendete Laserklasse eine uneingeschränkte Nutzung des Systems in Deutschland im öffentlichen Raum. Eine Nutzung des LCMS im Rahmen von Zustandserfassungskampagnen im deutschen Straßennetz, wo ein Mitschwimmen des Messfahrzeugs im fließenden Verkehr eine zwingende Voraussetzung ist, ist derzeit ausgeschlossen.

Hinsichtlich der verwendeten Methoden lassen sich die bisherigen Ansätze, eine automatisierte Schadenserkennung auf Straßenoberflächen zu realisieren, in (1) Bildanalysemethoden, (2) maschinelle Lernverfahren mit manuell designter Merkmalsextraktion und (3) Deep-Learning-Verfahren gruppieren (siehe Tabelle 8).

Tabelle 8: Taxonomie ausgewählter Ansätze auf Basis der verwendeten Methoden.

Methoden/Ansatz	Schlagworte	Publikation (Auswahl)
Bildanalyse		(CHAMBON & MOLIARD 2011)
	▪ Filteroperationen	(CANZLER & WINKLER 2012)
	▪ Schwellenwertoperationen	(SALMAN et al. 2013)
	▪ Morphologische Operationen	(OLIVEIRA & CORREIA 2014)
	▪ graphenbasierte Analysen	(PENG et al. 2015)
		(AMHAZ et al. 2016)
Maschinelle Lernverfahren mit manuell designter Merkmalsextraktion		(OLIVEIRA & CORREIA, 2017)
	▪ Gray-Level Co-Occurrence Matrix (GLCM)	
	▪ Local Binary Patterns (LBP)	(VARADHARAJAN et al. 2014)
	▪ Histogram of Oriented Gradients (HOG)	(KAPELA et al. 2015)
	▪ Random Forest	(QUINTANA, TORRES & MENENDEZ 2016)
	▪ Support Vector Machines	(WANG et al. 2018)
Deep Learning Verfahren	▪ Ensemble Learning	
	▪ Convolutional Neural Networks (CNN)	(ZHANG et al. 2016)
	▪ Objekterkennung und Klassifikation	(EISENBACH et al. 2017)
	▪ semantische Segmentierung	(WANG & HU 2017)
		(PAULY et al. 2017)
		(FAN et al. 2020)
		(STRICKER et al. 2021b)

Den Bildanalyseverfahren liegt die Annahme zugrunde, dass sich Risse als lokale Intensitätsminima identifizieren und mittels Schwellenwertoperationen extrahieren lassen (JÄHNE 2012). Mithilfe von beispielsweise morphologischen Operationen oder graphenbasierten Analysen können die initialen Schadstellenkandidaten zu einer finalen Segmentierung des Schadens weiterverarbeitet werden (siehe Abbildung 69). Auch von der BAST wurde ein Forschungsprojekt betreut, das eine Automatisierung im Bereich TP3 zum Ziel hatte und bei

dem eine Kombination von verschiedenen Bildanalyseansätzen eingesetzt wurde, um Risse in Oberflächenbildern zu erkennen (CANZLER & WINKLER 2012).

Während zu Beginn der 2010er-Jahre noch die Verwendung von Bildanalysemethoden die wissenschaftlichen Publikationen dominierte, wandelte sich dies im Laufe des Jahrzehnts, sodass heute der überwiegende Teil der publizierten Ansätze Methoden des maschinellen Lernens und insbesondere des Deep Learnings nutzt (MUNAWAR et al. 2021).

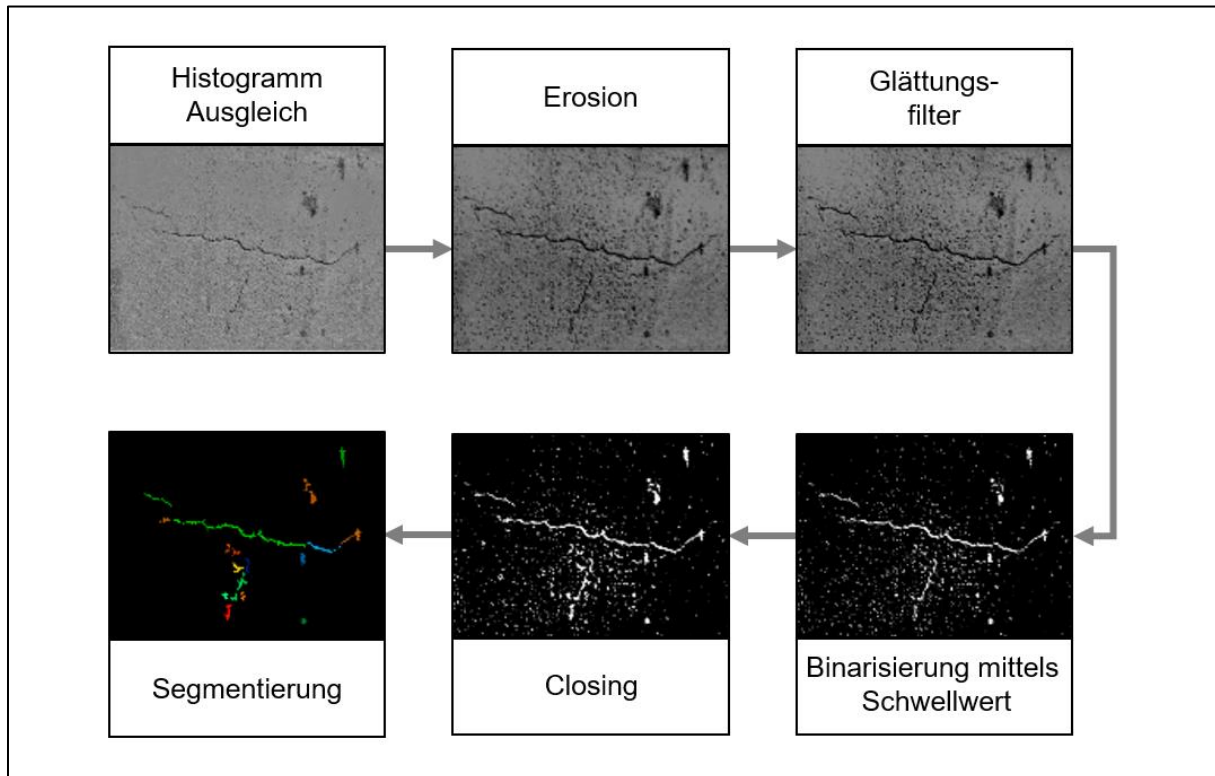


Abbildung 69: Beispiel für einen Bildanalyseansatz, bei dem versucht wird, mithilfe von gestaffelten Filter- und Schwellenwertoperationen Risse aus Bilddaten zu extrahieren (nach CHAMBON & MOLIARD 2011).

Konventionelle maschinelle Lernverfahren arbeiten in der Regel nicht pixelweise, sondern auf Bildausschnitten (siehe Abbildung 70). Für diese lokalen Bereiche (z. B. 32 x 32 Pixel Quadrate), werden zunächst Merkmale extrahiert und diese dann durch einen oder mehrere Klassifikatoren klassifiziert. Beispiele für häufig verwendete, explizit designte Merkmale sind GLCM-, LBP- oder HOG-Features (JÄHNE 2012). Wahl und Parametrisierung dieser Merkmale setzen ein hohes Maß an Fachwissen voraus. Im Unterschied dazu nutzen Deep-Learning-Verfahren CNN, bei denen sich die problemspezifischen Merkmalsdetektoren im Ergebnis eines datengetriebenen Lernprozesses selbständig herausbilden (STRICKER et al. 2021a).

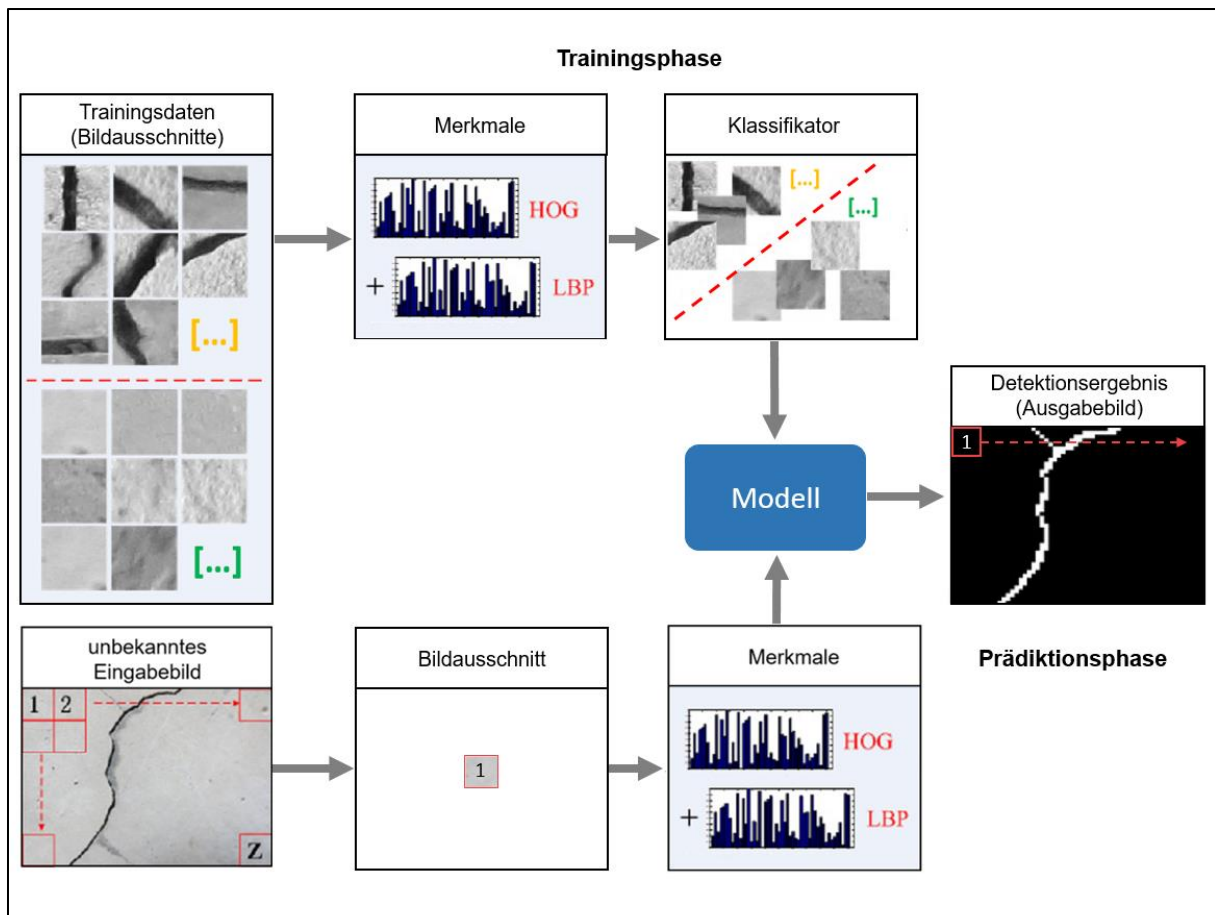


Abbildung 70: Beispiel für einen klassischen Ansatz des Maschinellen Lernens. Auf Basis von Bildausschnitten werden Merkmale berechnet und diese klassifiziert. Aus den bildausschnittswisen Detektionsergebnissen wird ein Ausgabebild zusammengesetzt (nach WANG et al. 2018).

Die Untersuchungen von EISENBACH et al. (2017) zeigen, dass bei der Detektion von Fahrbahnschäden auf CNN basierende Ansätze sowohl Bildanalyseansätzen als auch Verfahren mit klassischen, designten Merkmalsextraktion deutlich überlegen sind. Auch für andere Anwendungsdomänen, wie beispielsweise die medizinische Bildanalyse, konnten ähnliche Nachweise geführt werden (KARACA, CATTANI & MOONIS 2017). Zuletzt wurde im DACH-Raum ein mehrjähriges Forschungsprojekt abgeschlossen, bei dem die Möglichkeiten einer Deep Learning basierten Bildanalyse sowie Unsicherheitsmodellierung systematisch untersucht wurden (STRICKER et al. 2021a). Die wesentlichen Ergebnisse sind bei STRICKER et al. (2021b) beschrieben und legen den Schluss nah, dass sich eine pixelgenaue Segmentierung von Oberflächenschäden nach den länderspezifischen Schadenskatalogen der Schweiz, Österreich und Deutschland mit einem CNN basierten Encoder/Decoder-Netz realisieren lässt.

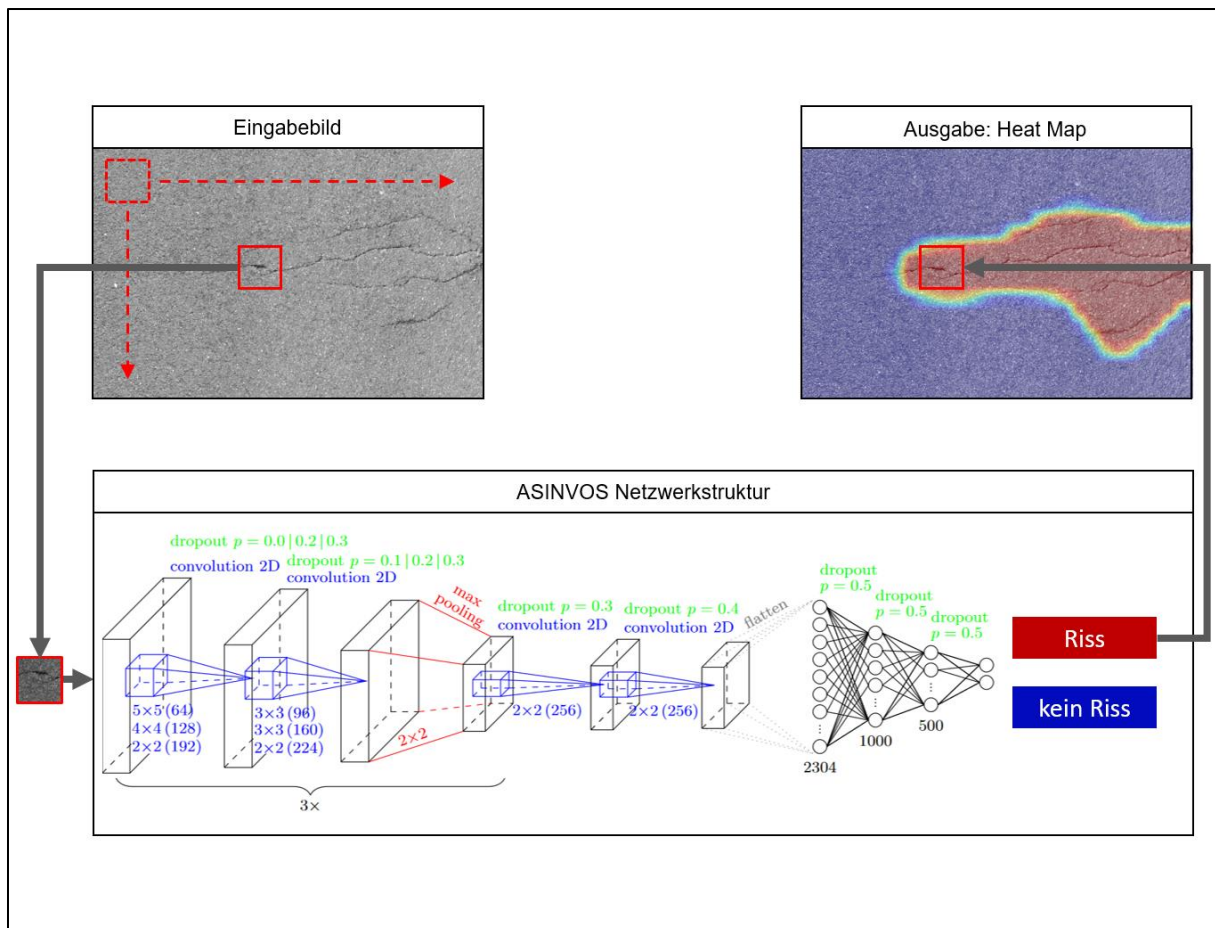


Abbildung 71: Beispiel für ein CNN, welches Risse auf Oberflächenfotos von Straßen erkennen kann (nach EISENBACH et al. 2017). Aus einem Eingabebild werden nach dem Moving-Window-Verfahren Bildausschnitte extrahiert und vom CNN klassifiziert. Die Netzwerkausgabe ist im dargestellten Beispiel eine binäre Klassenentscheidung. Für jeden klassifizierten Bildausschnitt wird der Konfidenzwert an die Extraktionsstelle des Bildausschnittes in ein Ausgabebild geschrieben. Im Beispiel wird so eine Konfidenzwertmaske für die Klasse Riss erzeugt (Heat Map).

Es lässt sich festhalten, dass die für die kommunale Straßenzustandserfassung relevanten bisherigen Automatisierungsansätze ausschließlich auf Kamerabilddaten basieren. Als passive Sensoren unterliegen Kameras jedoch einem entscheidenden Nachteil: Die vorherrschenden Wetterbedingungen können die Bildqualität bei ungünstiger Beleuchtung deutlich beeinflussen (SKAKUJ et al. 2019). Vor allem unterschiedliche Beleuchtungsstärken des natürlichen Tageslichtes können sich auf die Qualität der Kamerabilder und in der Folge auf die manuelle oder automatische Auswertung auswirken (SESSELMANN et al. 2020). In Abbildung 72 ist diese Problematik dargestellt.

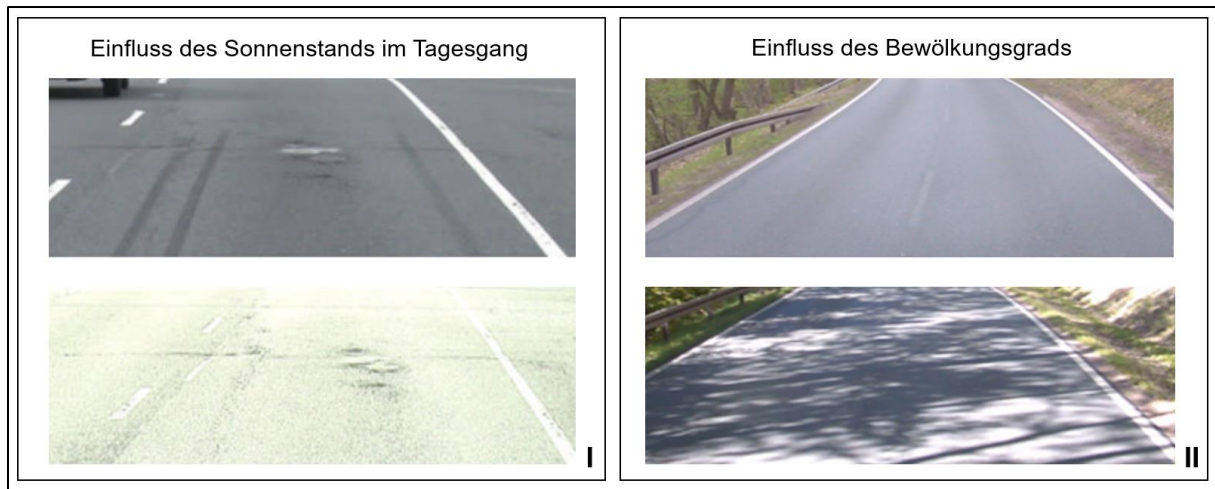


Abbildung 72: Einfluss des Umgebungslichtes im Tagesgang (I) und zu verschiedenen Wetterszenarien (II) auf Kamerabildern von Straßenoberflächen (nach SKAKUJ et al. 2019; BERTHOLD et al. 2023).

Wie in Kapitel 1 dargelegt, ist dies der Ansatzpunkt der vorliegenden Arbeit. Seit mehreren Jahren ist ein Laserscanner am Markt verfügbar, der wesentliche Sicherheits- und Qualitätsanforderungen der BAST erfüllt. Dieser Laserscanner ist bereits in diversen MLS-Systemen integriert, die sowohl im Bundesfernstraßennetz als auch im kommunalen Straßennetz Zustandsdaten erfassen. Deshalb ist es das Ziel dieser Arbeit, alle für die Straßenerhaltung relevanten Oberflächenschäden auf Basis derselben Daten zu erkennen und automatisiert auszuwerten, aus denen auch die Längs- und Querebenenheiten erfasst und bewertet werden können.

3 Datengrundlage

In diesem Kapitel wird erläutert, mit welchem mobilen Aufnahmesystem die Laserscans von Fahrbahnoberflächen erfasst wurden, die im Rahmen der Arbeit analysiert werden. Zunächst wird das Mobile-Mapping-System mit dem Fokus auf die für die Fragestellungen dieser Arbeit relevanten Primär- und Sekundärsensoren vorgestellt. Anschließend werden verschiedene Datensätze beschrieben, die zur Bearbeitung der Teilfragestellungen dieser Arbeit genutzt wurden.

3.1 Datenerfassung

Zur schnellen und flächendeckenden Erfassung hochauflösender Laserscans von Fahrbahnoberflächen werden kinematische Erfassungssysteme eingesetzt. Im Allgemeinen werden solche Multisensorsysteme als Mobile-Mapping-Systeme bezeichnet. Sind diese mit einem oder mehreren Laserscannern ausgestattet, spricht man von MLS-Systemen. Im Kapitel 2.2 wurden MLS-Systeme vorgestellt, die national und international eingesetzt werden.

In den nachfolgenden Passagen wird zunächst das von der Firma Lehmann+Partner GmbH entwickelte und betriebene MLS-System I.R.I.S vorgestellt, welches die Datengrundlage für die Arbeit in Form von 3D-Straßenoberflächenscans liefert (LEHMANN+PARTNER GMBH 2022). Anschließend wird das handgeführte GNSS-Messgerät beschrieben, mit dem die Passpunktaufnahmen zur Evaluierung der absoluten Lagegenauigkeit durchgeführt wurden (LEICA GEOSYSTEMS AG 2017).

3.1.1 Mobiles Laserscanning System I.R.I.S

Die Daten wurden mit dem kinematischen Erfassungssystem I.R.I.S (Integrated Road Information System) aufgenommen. I.R.I.S ist ein schnell fahrendes Mobile-Mapping-System zur bildhaften und dreidimensionalen Erfassung des Straßenraumes (siehe Abbildung 73). Das Messfahrzeug ist so konzipiert und dimensioniert, dass eine Befahrung sowohl des Fernstraßennetzes als auch des kommunalen Straßennetzes im fließenden Verkehr möglich ist. Die auf dem Fahrzeugdach montierte Rahmenkonstruktion trägt die Sensorik. Dazu zählen Umfeldkameras, Laserscanner und Komponenten des Positionierungssystems.



Abbildung 73: Das MLS-System I.R.I.S 16 der Firma Lehmann+Partner GmbH. Die Prinzipskizze (unten) zeigt den Messaufbau von der Seite und von hinten. Dargestellt sind die beiden für das Laserscanning relevanten Komponenten: inertiales Positionierungssystem (I) und Laserscanner (II). Die Montagehöhe h des Laserscanners beträgt ca. 3 m. Bei einem Sichtfeld von $\alpha = 70^\circ$ kann ein ca. 4,2 m breites Profil b der Fahrbahnoberfläche erfasst werden.

Der modulare Aufbau des I.R.I.S Systems ist in Abbildung 74 dargestellt. Die einzelnen Sensor-Komponenten lassen sich grob in drei Hauptgruppen gliedern: die primäre Sensorik umfasst die Sensoren zur Positionierung und Orientierung der Messplattform. Zur sekundären Sensorik gehören alle Sensoren, die Daten des Straßenraumes aufzeichnen (Kameras und Laserscanner). Daneben existiert eine dritte Systemkomponente, die im Gesamtsystem die Steuerung, Datenspeicherung und Energieversorgung sicherstellt. Die zeitliche Synchronisation aller drei Module und Sensoren ist wichtig, da bereits geringe Ungenauigkeiten in der Zuweisung der Zeitstempel bei hohen Fahrzeuggeschwindigkeiten oder hoher Fahrdynamik gravierende

Auswirkungen auf die Georeferenzierung der Sekundärsensordaten haben kann. Die eingesetzten Zeitserver ermöglichen eine Genauigkeit der Zeitbestimmung im Bereich von 0,000001 s.

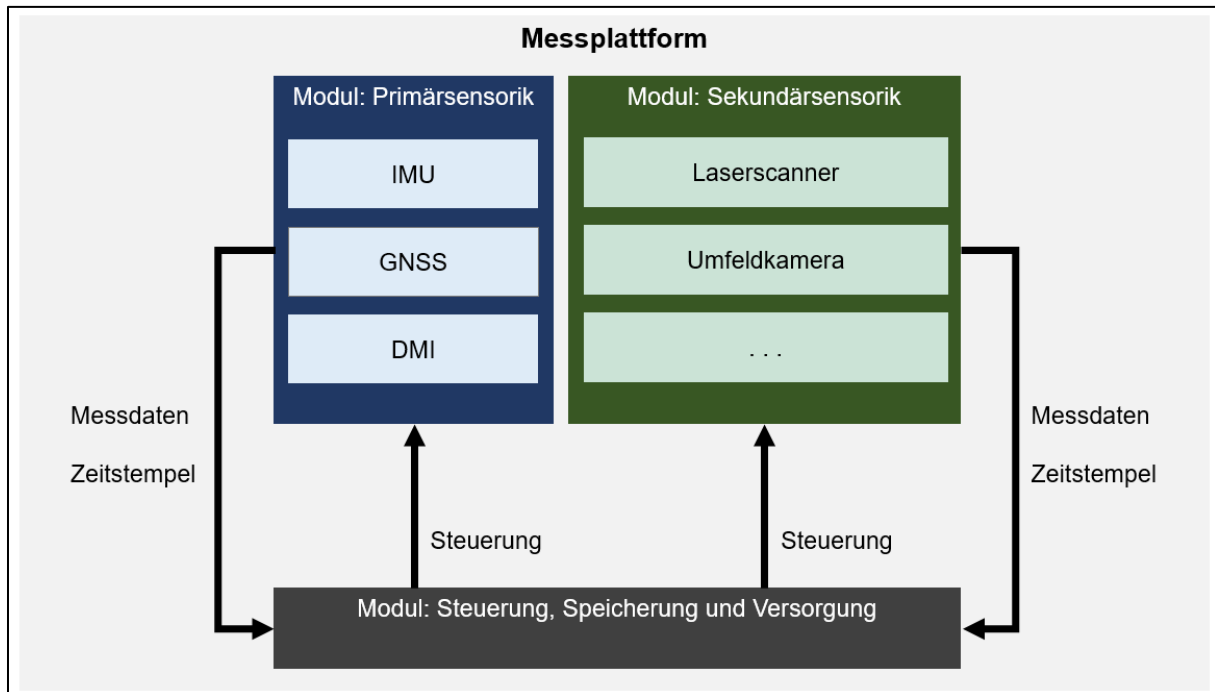


Abbildung 74: Übersicht zum modularen Aufbau des I.R.I.S Systems.

Im Folgenden werden die für die vorliegende Arbeit relevanten Komponenten des I.R.I.S Messsystems kurz vorgestellt. Dabei handelt es sich um ein inertiales Positionierungssystem des Herstellers Applanix aus der POS LV Reihe sowie den Oberflächenlaserscanner vom Typ Fraunhofer IPM Pavement Profile Scanner Plus. Über die zeitliche Synchronisierung der Messungen des Positionierungssystems und des Laserscanners entsteht das grundlegende Datenprodukt für die weiterführenden Analysen in Form von dichten 3D-Punktwolken.

3.1.1.1 Inertiales Positionierungssystem

Zu den Kernkomponenten des I.R.I.S Systems zählt das inertielle Positionierungssystem vom Typ Applanix POS LV 420, welches der Bestimmung der Fahrzeugtrajektorie im übergeordneten Koordinatensystem dient. Die Trajektorie stellt dabei die äußere Orientierung der sekundären Sensorik dar. Bei den Systemen der Applanix POS LV Serie handelt es sich um integrierte Systeme aus globalem Navigationssatellitensystem (GNSS), inertialer Messeinheit (IMU) und einem Wegstreckenmesser (DMI) (siehe Abbildung 75). Durch die

Kombination dieser Komponenten ist sowohl die Bestimmung der absoluten Position, als auch der relativen Positionsänderung möglich (APPLANIX 2019). Das GNSS dient dabei der Positionierung der Plattform im globalen Bezugssystem. Die IMU nutzt zur Bestimmung der Winkeländerungen und der relativen Orientierung ein faseroptisches Kreissystem und Beschleunigungssensoren, welche mit einer Frequenz von 200 Hz Daten aufzeichnen. Die eingesetzte IMU besitzt einen Winkelfehler von weniger als $0,03^\circ$ je Stunde, sodass auch bei schlechten GNSS-Bedingungen eine Positionsbestimmung mit einer geringen sensor-spezifischen Drift fortgeführt werden kann. Zusätzlich ist zur Erfassung des zurückgelegten Weges ein Wegstreckenmesser (Odometer) an einem Rad des Fahrzeuges montiert (engl.: Distance Measurement Unit, DMI). Je Radumdrehung werden etwa 2000 Pulse aufgezeichnet. Die Auflösung beträgt somit, je nach Radumfang, circa 0,5 mm je Puls.



Abbildung 75: Komponenten des POS LV Systems (APPLANIX 2019).

Aus den Rohdaten aller am Fahrzeug verbauten Sensoren wird während der Messung mit Hilfe eines Kalman Filters eine Realtime-Trajektorie berechnet, welche die Position und alle Raumwinkel enthält. Hierbei können auch Echtzeitkorrekturdaten eines NTRIP Casters verwendet werden, sofern im Einsatzgebiet eine Internetverbindung aufgebaut werden kann. Die Ermittlung der finalen Positionen und Winkel erfolgt typischerweise in einem Post-Processing-Schritt mit Hilfe der proprietären Applanix Software POSpac MMS unter Nutzung externer Korrekturdaten (LANDAU, VOLLATH & CHEN 2002). Bei den externen Daten handelt es sich zum einen um präzise Bahndaten der Satelliten (MONTENBRUCK, STEIGENBERGER & HAUSCHILD 2015), die u. a. Uhrzeit-Korrekturen enthalten, mithilfe derer die GNSS-Lösung im

Post-Processing der Trajektorie verbessert werden kann (siehe Kapitel 2.2.4.1). Zum anderen können GNSS-Referenzstationen von freien Anbietern sowie kommerziellen Diensten, wie beispielsweise SAPOS² oder VRS Now³, als externe Korrekturdaten im Post-Processing genutzt werden. Mithilfe der ortsfest erfassten GNSS-Daten eines Netzes aus Referenzstationen, welches die Fahrzeugtrajektorie räumlich und zeitlich einschließt, können atmosphärische, zeitliche und orbitale Fehlereinflüsse minimiert werden (siehe Kapitel 2.2.4.2).

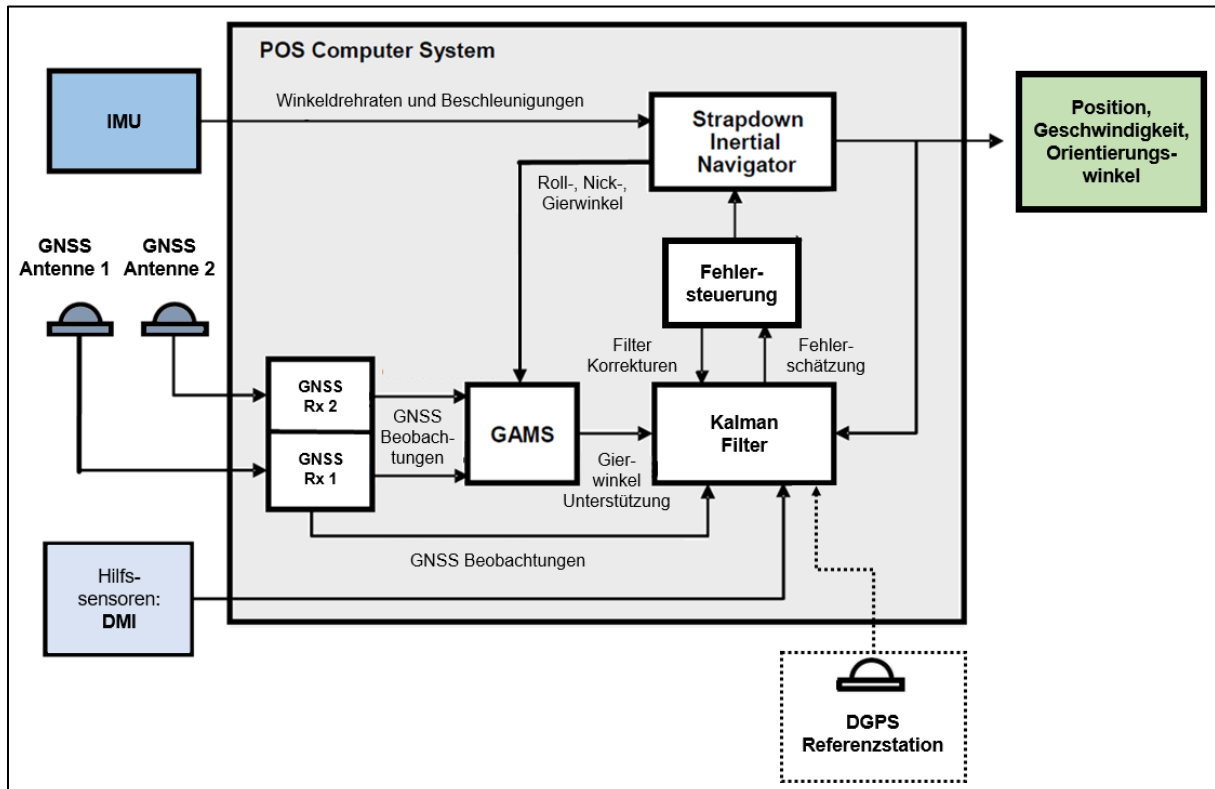


Abbildung 76: Prinzip der Applanix IN-Fusion Architektur, die in der Nachbearbeitungs-Softwareprodukt POSPac verwendet wird (nach NASTRO 2008).

In Abbildung 76 ist schematisch dargestellt, wie mithilfe eines Kalman Filters GNSS-, IMU- und DMI-Daten integriert und so hochgenaue Positions-, Orientierungs- und Geschwindigkeitsdaten erzeugt werden. Im Ergebnis liegt die Fahrzeugtrajektorie als Punktfolge mit allen relevanten Informationen, wie beispielsweise Zeitstempel, Beschleunigungen, Raumwinkel und Koordinaten im 200 Hz Intervall vor. Das eingesetzte Positionierungssystem ist bei gegebener GNSS-Abdeckung mit 2 cm und bei 60-sekündigem GNSS-Ausfall mit 12 cm Lagegenauigkeit (RMS) spezifiziert (APPLANIX 2019).

² Website: <https://sapos.de/>

³ Website: <https://www.vrsnow.eu/Map/SensorMap.aspx>

3.1.1.2 Laserscanner

Bei dem am Fahrzeugheck montierten und auf die Fahrbahnoberfläche ausgerichteten Laserscanner handelt es sich um einen Pavement Profile Scanner Plus (PPS-Plus), der vom Fraunhofer-Institut für Physikalische Messtechnik hergestellt wird (siehe Abbildung 77). Die erste Generation dieses Scannertyps wurde bereits im Jahr 2012 unter dem Namen Pavement Profile Scanner (PPS) vorgestellt (REITERER et al. 2013). Der PPS war damals das erste LiDAR-Instrument, welches von der BAST eine zeitbefristete Zulassung zur Erfassung von Querebenenheiten im Rahmen der Zustandserfassung und -bewertung von Straßen erhalten hat. Diese Zertifizierung wurde seither jährlich für den im Messfahrzeug S.T.I.E.R (speziell konfiguriertes I.R.I.S System) integrierten PPS bestätigt. Bis Ende 2022 erhielt kein anderes LiDAR-System diese Betriebszulassung in Deutschland.

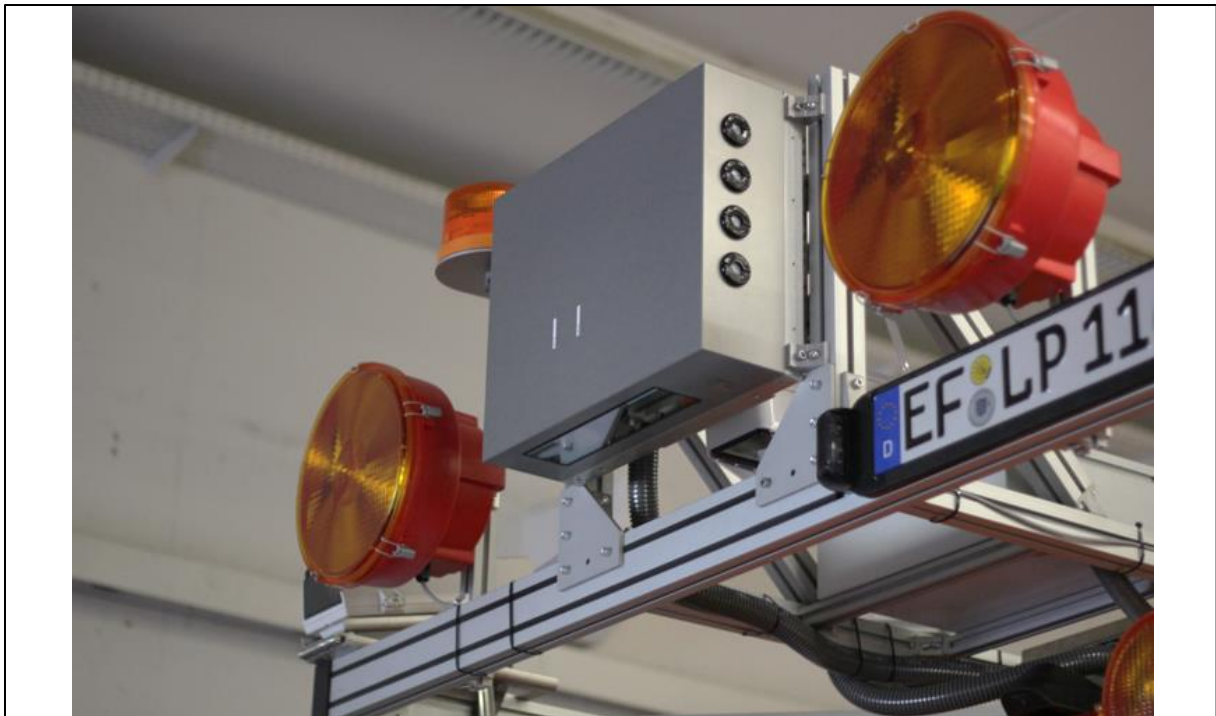


Abbildung 77: Laserscanner Pavement Profile Scanner Plus vom Fraunhofer IPM montiert am Dachträgerrahmen des I.R.I.S 16 Systems.

Im Jahr 2018 wurde der PPS-Plus als eine erweiterte Variante des ursprünglichen PPS präsentiert (FRAUNHOFER INSTITUTE FOR PHYSICAL MEASUREMENT TECHNIQUES IPM 2018). Dabei blieb die 3D-Einheit (PPS-Modul) funktional unangetastet und es wurde eine zusätzliche 2D-Einheit (MPS-Modul) entwickelt und in den neuen Scannertyp integriert.

Im PPS-Modul werden die Entfernungsmessungen mithilfe des Phasenvergleichsverfahrens und einem Infrarotlaser mit einer Wellenlänge von 1550 nm realisiert. Durch das Messprinzip

ist eine hohe Präzision der Entfernungsmessung bei hohen Messgeschwindigkeiten und gleichzeitiger Augensicherheit gewährleistet (siehe Tabelle 9). Der Messbereich liegt zwischen 1,2 m und 5 m. Ein rotierender Polygonspiegel mit acht Facetten bildet die Ablenkeinheit und lenkt den Laserstrahl profilweise über die Fahrbahnoberfläche (siehe Abbildung 78). Die Drehrate beträgt dabei ca. 6000 Umdrehungen pro Minute (100 Hz). Ein Öffnungswinkel von 70° ermöglicht bei einer Montagehöhe des Geräts von ca. 3 m über Grund die Erfassung eines ca. 4,2 m breiten Profils. Das PPS-Modul erfasst primär die Polarkoordinaten Entfernung und Scanwinkel. Zusätzlich werden Zeitstempel und die Reflexionsintensität pro Punktmessung aufgezeichnet. Die im Rahmen der vorliegenden Arbeit verwendeten Scanner sind mit dem Standard PPS-Modul ausgerüstet, das 1 Million Punkte pro Sekunde mit ca. 920 Punkten pro gescanntem Profil misst. Es besteht die Möglichkeit, herstellerseitig eine zweite Detektionseinheit zu verbauen, um ca. 1800 Punkte pro Profil aufzeichnen zu können. Die Messgenauigkeit des PPS wird vom Hersteller, gemittelt über ein $10\text{ cm} \times 10\text{ cm}$ großes Oberflächenelement, im Submillimeterbereich angegeben (FRAUNHOFER INSTITUTE FOR PHYSICAL MEASUREMENT TECHNIQUES IPM 2022). Zur Strahldivergenz werden weder in der Betriebsanleitung des Gerätes noch in den offiziellen Spezifikationen des Herstellers Angaben gemacht. Wie in Kapitel 2.2.1 beschrieben, bedingt die Strahldivergenz in Abhängigkeit zur Messdistanz die Größe des Laserspots auf dem Messobjekt. Es ist davon auszugehen, dass der Footprint des Lasers bei dem Entfernungsbereich, für den das Standard PPS-Modul entwickelt wurde sowie dem Punktintervall innerhalb eines Scanprofils von ca. 4,5 mm einen Durchmesser von ungefähr 5 mm oder weniger hat, wenn der Laser auf ein ebenes Messobjekt im Bereich direkt unter dem Scanner in 3 m Entfernung trifft. So ist gewährleistet, dass die einzelnen Messpunkte auf einem gemessenen Profil nicht miteinander korreliert sind.

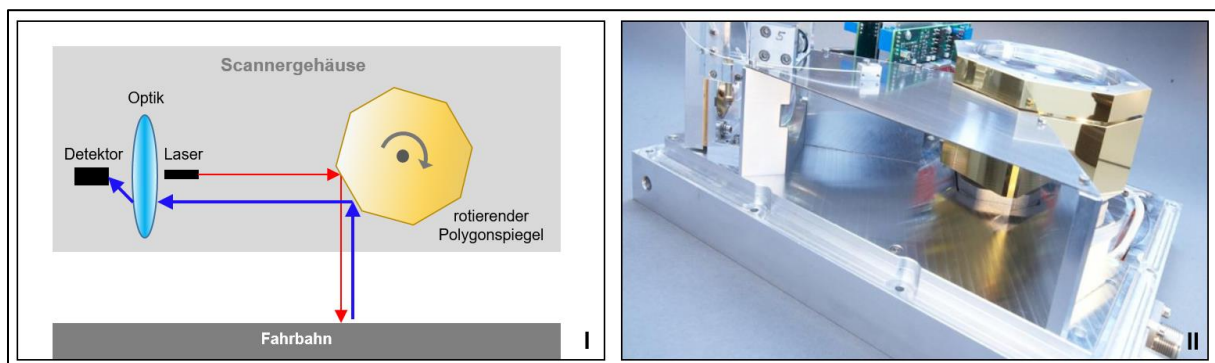


Abbildung 78: Schematische Darstellung des im PPS verwendeten rotierenden Polygonprismas, welches den Laserstrahl profilweise über das Messobjekt ablenkt (I). Das Foto rechts zeigt den Polygonspiegel eines PPS der ersten Generation (II nach REITERER 2013).

Mithilfe des sogenannten Boresight-Vektors (siehe Kapitel 3.1.1.3), der in einem Kalibrierungsprozess bestimmt werden muss, einer Transformation (siehe Kapitel 2.2.5) und dem jeweiligen Zeitstempel kann jede lokale Polarkoordinaten-Messung in den 3D-Raum des Weltkoordinatensystems überführt werden (REITERER et al. 2013). Somit entsteht im Zuge der Bewegung des Fahrzeuges aus den eigentlichen zweidimensionalen Messprofilen des PPS eine ungeordnete, dreidimensionale Punktwolke, bei der jeder 3D-Punkt direkt georeferenziert und mit seiner Reflexionsintensität attribuiert ist. Während der Abstand zwischen zwei Punkten innerhalb eines Scanprofils fix ist, hängt der Abstand zwischen zwei Profilen in der 3D-Punktwolke davon ab, wie schnell sich das Messfahrzeug bewegt. Bei einer Befahrungsgeschwindigkeit von 20 km/h beträgt der Profilabstand 7 mm, bei 50 km/h 17,5 mm und bei 80 km/h 28 mm.

Tabelle 9: Technische Spezifikation der 3D-Einheit des Pavement Profile Scanners Plus (nach FRAUNHOFER INSTITUTE FOR PHYSICAL MEASUREMENT TECHNIQUES IPM 2022).

Geräteeigenschaft		Spezifikation
Messprinzip		Phasenlaufzeitmessung
Lasertyp		cw
Laserwellenlänge (typisch)		1500 nm
Optische Laserleistung (typisch)		200 mW
Messbereich	Eindeutigkeitsbereich	1,2 m
	Mindestabstand	1,3 m
	Maximaler Abstand	5 m
Messrate (Entfernungsmessung)		1 MHz
Auflösung der Entfernungsmessung bei 3 m Messabstand	bei 90 % Remission	< 0,1 mm *
	bei 20 % Remission	< 0,3 mm *
Scanwinkel		70 °
Scanfrequenz **		800 Hz
Winkelauflösung (interpoliert)		0,018 °
Laserschutzklasse (System)		1
Sonstiges	Maße (B x H x T)	500 x 400 x 160 mm
	Gewicht	40 kg
	Schutzklasse IP	67
	EMV-Normen	EN 60825 - 1
		EN 61000 - 4 -2
		EN 61000 - 4 -4
		EN 61000 - 4 -6

* Standardabweichung des Mittelwerts über 100 Messungen

** gemessene Profile pro Sekunde

Zusätzlich zum PPS-Modul, verfügt der PPS-Plus über eine 2D-Messeinheit. Im MPS-Modul wird ein spezieller Linienlaser mit einer Wellenlänge von 830 nm genutzt, um hochauflösende Oberflächenbilder unabhängig von natürlichen oder zusätzlichen künstlichen Beleuchtungsquellen aufzunehmen. Anders als beim PPS-Modul, wird hier kein Laserpunkt, sondern eine Laserlinie auf das rotierende Polygonspiegelrad projiziert. Diese Laserlinie beleuchtet demnach einen Streifen auf der Fahrbahnoberfläche. Eine aus 16 CCD-Zeilen bestehende Detektoreinheit nimmt die rückgestreuten Photonen des durch den Laser beleuchteten Fahrbahnbereichs auf. Jede CCD-Zeile umfasst dabei 4000 Einzeldetektoren. Somit kann eine geschwindigkeitsunabhängige Bodenauflösung in Querrichtung von 1,2 mm und eine geschwindigkeitsabhängige Bodenauflösung von 1,7 mm in Fahrtrichtung erreicht werden (FRAUNHOFER INSTITUTE FOR PHYSICAL MEASUREMENT TECHNIQUES IPM 2022). In der vorliegenden Arbeit werden keine Daten des MPS-Moduls, sondern ausschließlich Daten des PPS-Moduls verwendet. Deshalb beziehen sich alle weiteren Aussagen zu dem Laserscanner auf die 3D-Einheit des PPS-Plus (nachfolgend: PPS). In Tabelle 9 sind die wichtigsten Spezifikationen des PPS aufgeführt.

3.1.1.3 Kalibrierung und Boresight-Alignment

Wie im vorherigen Kapitel beschrieben, muss im Rahmen eines Kalibrierungsprozesses der Boresight-Vektor bestimmt werden, um aus den Polarkoordinatenmessungen des Laserscanners 3D-Punktwolken im übergeordneten Koordinatensystem generieren zu können. Der Boresight-Vektor repräsentiert hierbei die Transformation vom Nullpunkt des PPS-Scannerkoordinatensystems zum Nullpunkt des Applanix-Positionierungssystems.

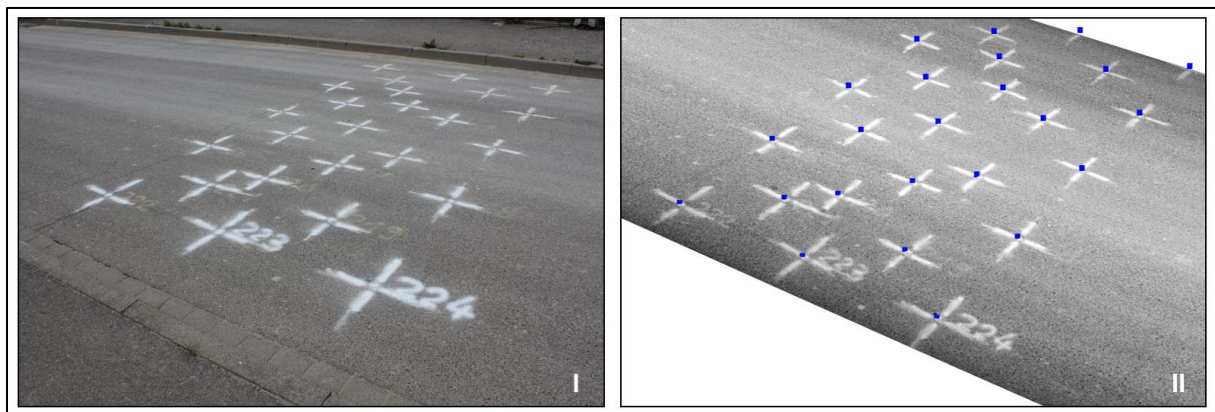


Abbildung 79: Kalibrierfeld zur Bestimmung des Boresight-Vektors. Die linke Teilabbildung (I) zeigt ein Foto des Kalibrierfeldes. Die rechte Teilabbildung (II) zeigt die tachymetrisch eingemessenen Referenzpunkte (blau) und den PPS-Laserscan in perspektivischer Ansicht nach der Reflexionsintensität in Graustufen eingefärbt.

Durch Messungen auf einem Kalibrierfeld aus verschiedenen Richtungen werden die Translations- und Rotationsanteile iterativ angepasst, bis die Klaffungen zwischen verschiedenen Befahrungsrichtungen und die Abweichungen zu tachymetrisch eingemessenen Zielmarken (Messnägeln) minimal sind (siehe Abbildung 79). Die Boresight-Parameter werden durch Geodäten der Firma Lehmann+Partner GmbH bestimmt.

Der Kalibrierungsprozess umfasst neben den beschriebenen dynamischen auch statische Messungen auf einem Prüflineal aus Stahl (siehe Abbildung 80). Dabei handelt es sich um ein Flachlineal von 5 m Länge, 0,16 m Höhe und einer 0,04 m breiten Oberfläche.



Abbildung 80: Kalibrierung des PPS mittels eines Stahl-Prüfkörpers (I). Die rechte Teilabbildung (II) zeigt auch den Infrarotsensor, mit dessen Hilfe der für das menschliche Auge unsichtbare Laser-Strahl detektiert und das Prüflineal in der Messlinie des PPS ausgerichtet werden kann.

Die Ebenheit der Prüfkörperoberfläche wurde von der Deutschen Akkreditierungsstelle GmbH gemäß der Richtlinie LW-K-Geradheit-1411-18 mittels Differenzmessverfahren unter Nutzung zweier elektronischer Neigungsmessgeräte vom Typ „Wyler BlueLevel“ bestätigt (siehe Anhang I). Die dabei festgestellten Abweichungen waren, wie in Abbildung 81 zu sehen ist, kleiner als 0,06271 mm.

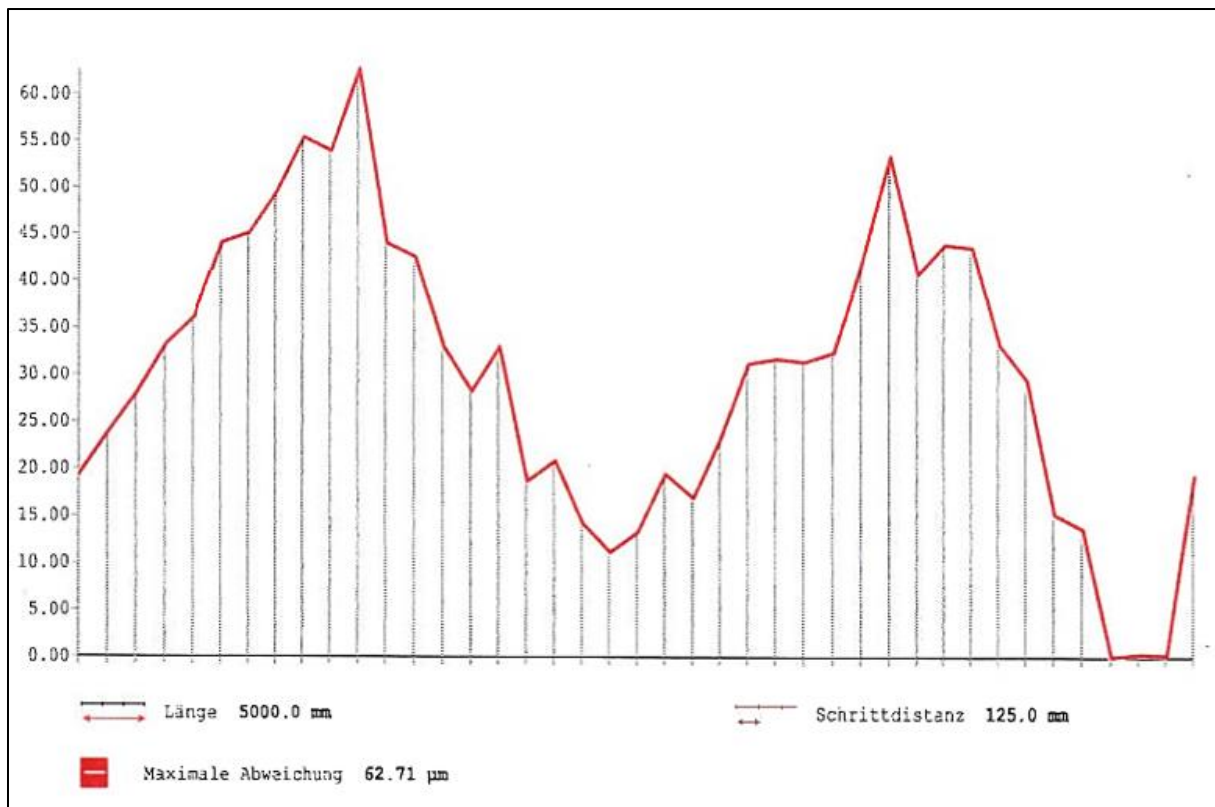


Abbildung 81: Auszug aus dem Kalibrierschein, der die Ebenheit des Prüflineals bescheinigt. Die maximale Abweichung des Stahl-Prüfkörpers beträgt 0,06271 mm (siehe Anhang I).

Durch die Messungen auf dem Prüflineal kann analysiert werden, wie genau die Laserscannermessungen die definierte Prüfkörperoberfläche abbilden können. Abbildung 82 zeigt die Streuung der Lasermesspunkte auf dem Prüfkörper. Die Streuung gemittelt über 10 cm liegt mit mittleren Residuen von 0,18 mm und einer Standardabweichung von 0,13 mm in einem Bereich, der mit den Herstellerangaben von FRAUNHOFER INSTITUTE FOR PHYSICAL MEASUREMENT TECHNIQUES IPM (2022) vergleichbar ist (siehe Tabelle 9).

Neben der genannten geometrischen Prüfung kann die Erfassung der Reflexionsintensität auf der ebenen und homogen gefärbten Oberfläche des Prüflineals kalibriert und korrigiert werden. Dies ist nötig, da sich der Auftreffwinkel des Laserstrahls, bedingt durch die Messanordnung im Bereich direkt hinter dem Fahrzeug (Austrittswinkel 0°) von den seitlichen Bereichen (Austrittswinkel -35° bzw. $+35^\circ$) unterscheidet. Je weiter sich der Austrittswinkel des Lasers den Austrittswinkelextremen annähert, desto flacher wird dessen Auftreffwinkel auf dem Prüfkörper und desto weniger Photonen werden in Richtung der Detektoreinheit des Scanners reflektiert. Im Ergebnis sind die linken und rechten Randbereiche eines Messprofils im Vergleich zu zentralen Bereichen durch eine niedrigere Reflexionsintensität geprägt. Die Intensitätswerte werden als Attribut der Polarkoordinaten Winkel und Entfernung von der scannereigenen Software mit aufgezeichnet.

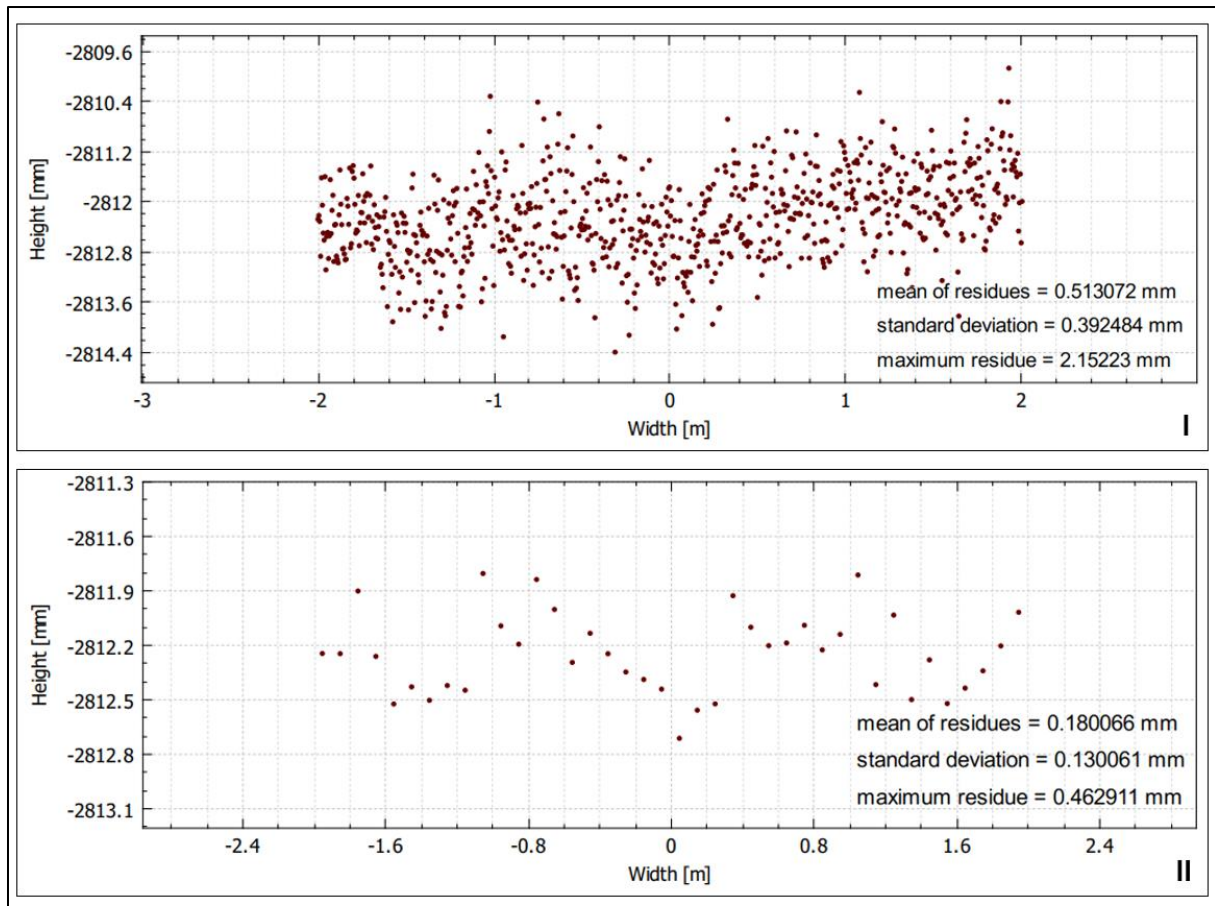


Abbildung 82: Darstellung der Streuung der Höhenwertmessungen auf dem Prüflineal im 5 mm Raster (I) und im 10 cm Raster (II). Entlang der X-Achse sind die Scanner-Messungen auf einer Breite von 4 m abgetragen. Die auf der Y-Achse abgetragene Höhenlage resultiert aus der Montagehöhe des Scanners und der Höhe des Messobjekts.

Die auf dem Prüflineal gemessene Intensitätskurve wird mit einem Polynom fünften Grades angenähert, wobei p_0 bis p_5 die zu bestimmenden Koeffizienten repräsentieren und x der Punkt-Index auf einem Scanprofil ist.

$$d_y = p_5 \cdot x^5 + p_4 \cdot x^4 + p_3 \cdot x^3 + p_2 \cdot x^2 + p_1 \cdot x + p_0 \quad (3)$$

Um in Abhängigkeit vom Scanwinkel schließlich einen korrigierten Intensitätswert zu berechnen, wird das auf dem Prüflineal bestimmte Polynom so angebracht, dass die Seitenbereiche der Kurve auf das Intensitätsniveau angehoben werden, welches dem Bereich direkt unter dem Scanner entspricht (siehe Abbildung 83). Auffällig ist auch die verminderte Streuung in den lateralen Bereichen. Eine Ursache dafür kann sein, dass sich der im zentralen Bereich punktförmige Footprint des Laserspots mit zunehmend flacher werdendem Auftreffwinkel zu den Seiten hin zu teilweise überlappenden Ellipsen verformt. Dadurch sind die Einzelmessung dort stärker korreliert und die Werteverteilung wirkt glatter.

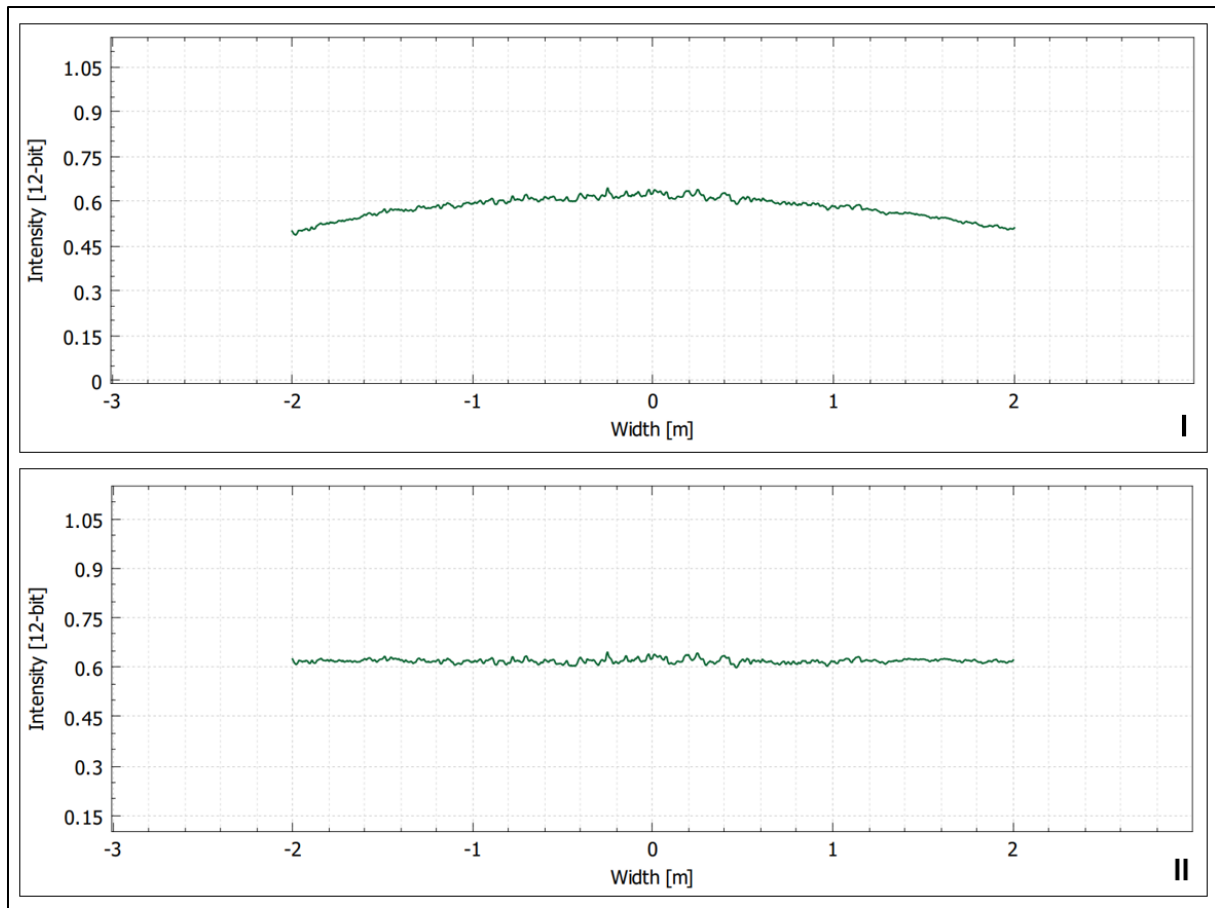


Abbildung 83: Darstellung der zwischen 0 und 1 normierten Intensitätswerte für die Lasermessungen auf dem Prüflineal. Entlang der X-Achse sind die Scanner-Messungen auf einer Breite von 4 m abgetragen. Der obere Graph (I) zeigt den Scanwinkel-abhängigen Intensitätsabfall hin zu den Profilrändern. Der untere Graph (II) zeigt die Kurve nach der Korrektur mittels eines Polynoms fünften Grades.

3.1.2 Handgeführte Satellitenvermessung

Neben den Laserscans des kinematischen Erfassungssystems I.R.I.S werden im Rahmen der vorliegenden Arbeit auch Daten verwendet, die mit einem handgeführten GNSS vermessungstechnisch aufgenommen wurden (siehe Kapitel 3.2.1). Typischerweise kommen solche Geräte in der Bauvermessung zum Einsatz – sowohl in stationären als auch kinematischen Szenarien. Für die Messungen im Rahmen dieser Arbeit wurde ein System des Herstellers Leica Geosystems verwendet. Wie in Abbildung 84 dargestellt, besteht es aus der Empfangseinheit (GNSS-RTK Rover Leica GS18), die oben auf einem Lotstab arretiert wird und einem seitlich am Lotstab angebrachten Bedienelement (Leica CS 20). Antenne und Controller stehen via Bluetooth miteinander in Verbindung.

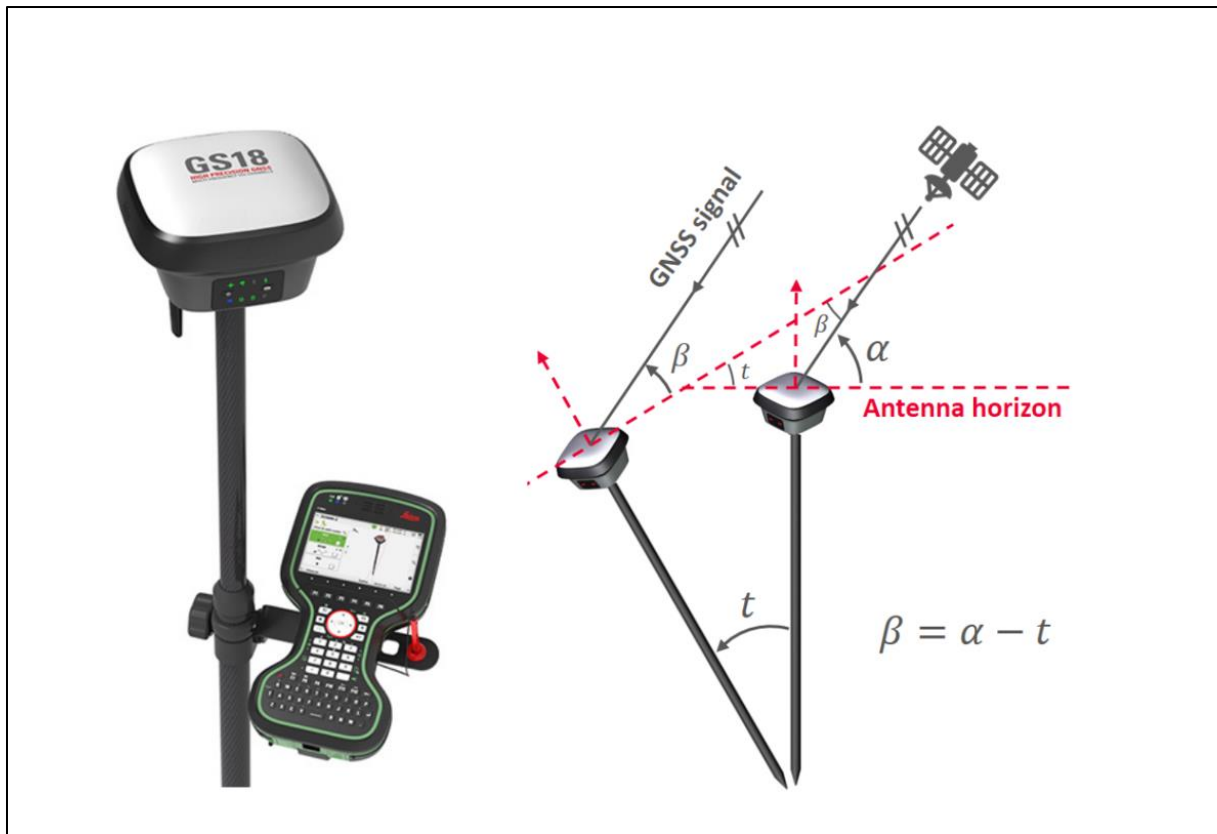


Abbildung 84: Abgebildet sind links der GNSS-RTK Rover Leica GS18 T und das Bedienelement Leica CS20 sowie rechts das Funktionsprinzip (LUO, SCHAUFLE & RICHTER 2018).

Im Zuge der Messung wird der Lotstab mit dessen unterer Spitze senkrecht auf das zu messende Objekt aufgehalten (siehe Abbildung 85). Dabei muss der Stab, klassischerweise unter Verwendung einer Libelle, für den Messzeitraum im Lot gehalten werden. Der verwendete GNSS-RTK Rover Leica GS18 T hingegen nutzt die Daten einer integrierten IMU, um die Bewegung und Orientierung der Antenne zu erfassen. Diese Technologie ermöglicht es auch, Punkte aufzuhalten, die nur mit einem schräg gestellten Teleskopstab erreichbar sind. Gemäß den technischen Spezifikationen sind nach einer Initialisierungszeit von ca. 4 Sekunden Positions-genauigkeiten in der Lage von 8 mm und in der Höhe von 15 mm möglich. Bei sehr schräg aufgehaltenem Lotstab kommt eine zusätzliche Lageungenauigkeit von $< 8 \text{ mm pro } 1^\circ$ Neigung hinzu (LEICA GEOSYSTEMS AG 2017).



Abbildung 85: Passpunktaufnahme mit einem Leica GNSS-RTK Rover.

Typischerweise werden Passpunkte im Rahmen von Mobile-Mapping-Kampagnen verwendet, wenn Ansprüche an die absolute Lagegenauigkeit der erfassten Sensordaten oder der daraus abgeleiteten Geodaten herrschen, die über die Spezifikation des verwendeten inertialen Positionierungssystems hinausgehen. In so einem Fall können Passpunkte verwendet werden, um im Post-Processing der Fahrzeugtrajektorie die Positionierung mit dieser zusätzlichen Datenquelle zu verbessern.

Im Rahmen der vorliegenden Arbeit jedoch dienen sie als Referenz- bzw. Passpunkte, um die absolute Lagegenauigkeit des mit dem I.R.I.S-System während der Fahrt erfassten Laserscans zu evaluieren. Dies ermöglicht in der Folge auch eine Aussage zur erreichbaren absoluten Lagegenauigkeit der automatisch aus den Laserscans segmentierten Geoobjekte, wie beispielsweise Oberflächenschäden, Einbauten oder Fahrbahnmarkierungen.

3.2 Datensätze

Die in Kapitel 5 beschriebenen Untersuchungen stützen sich auf Datensätze, die im Rahmen verschiedener Messfahrten mit den MLS-Systemen I.R.I.S 3, I.R.I.S 5, I.R.I.S 12, I.R.I.S 13 und I.R.I.S 16 aufgenommen wurden. Alle Messfahrzeuge waren zur Zeit der Datenaufnahme mit einem Applanix POS LV 420 und einem PPS-Laserscanner ausgerüstet. In den folgenden Abschnitten werden Details zu diesen Daten beschrieben. Die genauen Örtlichkeiten der jeweiligen Datenaufnahme werden bewusst nicht genannt, weil sie für die Untersuchungen und Erkenntnisse dieser Arbeit unerheblich sind und Teile der Daten nur unter der Bedingung der Anonymisierung verwendet werden durften.

3.2.1 Aspekt: absolute Positionierungsgenauigkeit

Um eine Aussage zur absoluten Lagegenauigkeit der Laserscans und der daraus abgeleiteten Geodaten treffen zu können, wurden am 28.07.2022 zwischen 8 und 17 Uhr im Umland von Bonn auf acht Kreisstraßen insgesamt 51 Passpunktmarkierungen auf die Fahrbahnoberfläche im potentiellen Sichtfeld des PPS-Laserscanners aufgebracht und mittels GNSS-RTK Rover (siehe Kapitel 3.1.2) eingemessen. Die Passpunkte decken eine Netzlänge von ca. 38 km ab (siehe Abbildung 86). Die Genauigkeit der Passpunktaufnahme ist Anhang II zu entnehmen.

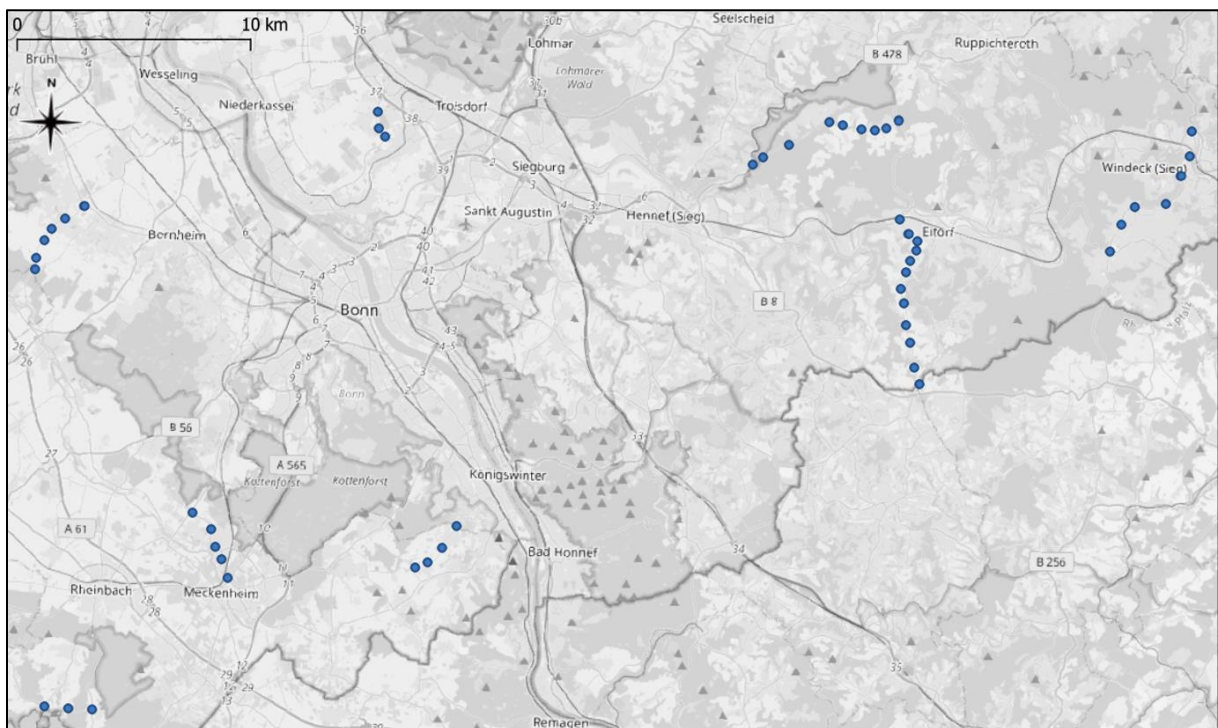


Abbildung 86: Lagekarte der Passpunkte (Basiskarte: © OpenStreetMap contributors).

Bei den mit Passpunkten markierten Straßenabschnitten im Untersuchungsgebiet handelt es sich mehrheitlich um Landesstraßen außerhalb geschlossener Ortschaften. Ein deutlich kleinerer Teil, ca. 20 % der Passpunktpositionen, entfällt auf Ortsdurchfahrten mit eher lockerer Bebauung. In Abbildung 87 ist ein Auszug der während der Befahrung aufgenommenen Frontkamerabilder dargestellt, um einen Eindruck der Messsituation im Untersuchungsgebiet zu vermitteln. Während einige Bereiche nahezu Open-Sky-Bedingungen bieten, verlaufen andere Straßenabschnitte durch Wälder, wo eine abschattungsbedingt verringerte GNSS-Abdeckung herrscht. Der Datensatz umfasst demnach sowohl Bedingungen, die für eine hochgenaue Positionierungslösung auf Basis einer kinematischen Erfassungsplattform im fließenden Verkehr günstig sind, als auch solche, die eine optimale Positionierungslösung erschweren.



Abbildung 87: Beispielhafte Aufnahmen der Frontkamera des I.R.I.S-Messfahrzeugs während der Aufnahme des Datensatzes zur Bestimmung der absoluten Genauigkeit der Fahrbahnoberflächenscans.

Nachdem die Passpunkte mit Sprühfarbe markiert und mit dem handgeführten RTK-GNSS vermessen wurden, fand die Messung der Fahrbahnoberfläche mit dem Messfahrzeug I.R.I.S 13 am Folgetag statt (siehe Abbildung 88).



Abbildung 88: Messfahrzeug I.R.I.S 13 bei der Messung der Passpunktstrecken am 29.07.2022 (Foto: Markus Löffler, Außendienst Lehmann + Partner GmbH).

Abbildung 89 zeigt beispielhaft den Abgleich einer händisch aufgenommenen Passpunktposition (blauer Punkt) und der entsprechenden Position im PPS-Laserscan. Anhand der Reflexionsintensität ist die Passpunktmarkierung im Laserscan deutlich zu erkennen. Dieselbe Szene ist im linken Teil der Abbildung aus der Perspektive der frontal am Messfahrzeug angebrachten Messbildkamera dargestellt. Der Abstand zwischen RTK-GNSS-Messpunkt und den aus dem Laserscan abgegriffenen 3D-Koordinaten quantifiziert die absolute Lagegenauigkeit des Laserscans im Bereich des Passpunktes.



Abbildung 89: Passpunktmarkierung im Frontkamerabild (I) und PPS-Laserscan (II). Rechts ist in blau der mittels RTK-GNSS eingemessene Passpunkt zu sehen. Im Beispiel beträgt die 3D-Distanz zwischen Laserscan und Passpunkt ca. 6,5 cm.

3.2.2 Aspekt: Präzision des Messsystems

Zur Bestimmung der Präzision (Wiederholgenauigkeit) des MLS-Systems wurde ein Prüfkörper wiederholt gemessen. Ziel ist es, anhand dieses Datensatzes die vom Messsystem produzierte Streuung auf einem bekannten Messobjekt zu quantifizieren.

Für diese Art der Analyse wurde ein Prüfkörper benötigt, der einerseits hinsichtlich seiner Dimensionen von einem mobilen Messsystem überrollt und aufgenommen werden kann (dynamisches Szenario). Andererseits musste dieser Prüfkörper dreidimensional nachmodelliert werden, um die Präzision des Messsystems auf Basis des digitalen Abbildes zu bestimmen. Für die Untersuchungen im Rahmen der Arbeit wurde deshalb ein Prüfkörper gewählt, der zur Kalibrierung eines Laserlichtschnittsensors der kanadischen Firma Pavemetrics dient (PETITCLERC, LAURENT & HABEL 2018). Dabei handelt es sich um Metallelemente, die auf einer 0,52 m x 0,47 m großen Grundplatte montiert sind. Ein Prüfzertifikat der Firma Pavemetrics bestätigt die Konformität des Prüfkörpers (siehe Anhang III). Der Prüfkörper ist, wie in Abbildung 90 auf der linken Seite dargestellt, pulverbeschichtet und hat eine matte, hellgraue Farbe.

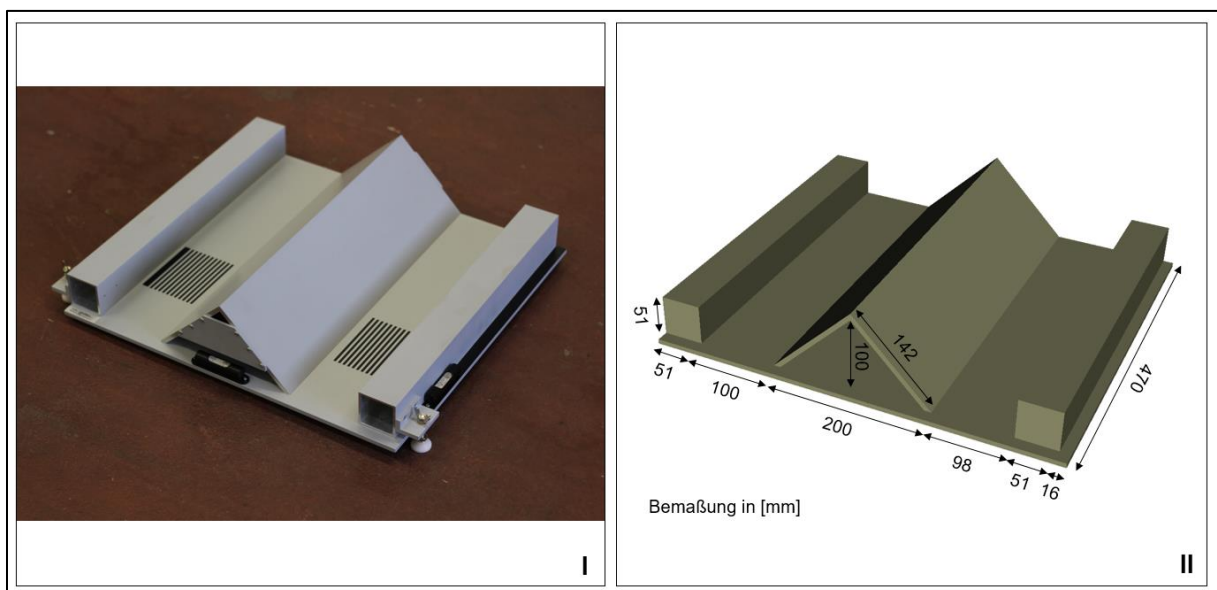


Abbildung 90: Darstellung der Konstruktion eines 3D-Abbildes (II) des realen Prüfkörpers (I). Details, wie Flügelmuttern oder Libellen, wurden nicht mit modelliert.

Die Datenaufnahme fand am 17.07.2020 bei trockenen Witterungsbedingungen mit dem Messsystem I.R.I.S 3 statt. Der Prüfkörper wurde mittig auf der Fahrbahnoberfläche ausgelegt und so ausgerichtet, dass die länglichen Rechteckelemente quer zur Fahrtrichtung verliefen. Abbildung 91 zeigt die Situation der Datenaufnahme als Kartendarstellung mit Fokus auf die

Lage des Prüfkörpers innerhalb des Laserscans und relativ zum Verlauf der Fahrzeugtrajektorien. Der Prüfkörper befand sich bei der Messung immer nahezu senkrecht unter dem Scanner. Abbildung 92 zeigt Fotos vom Prüfkörper und dem Messfahrzeug während der Datenaufnahme.

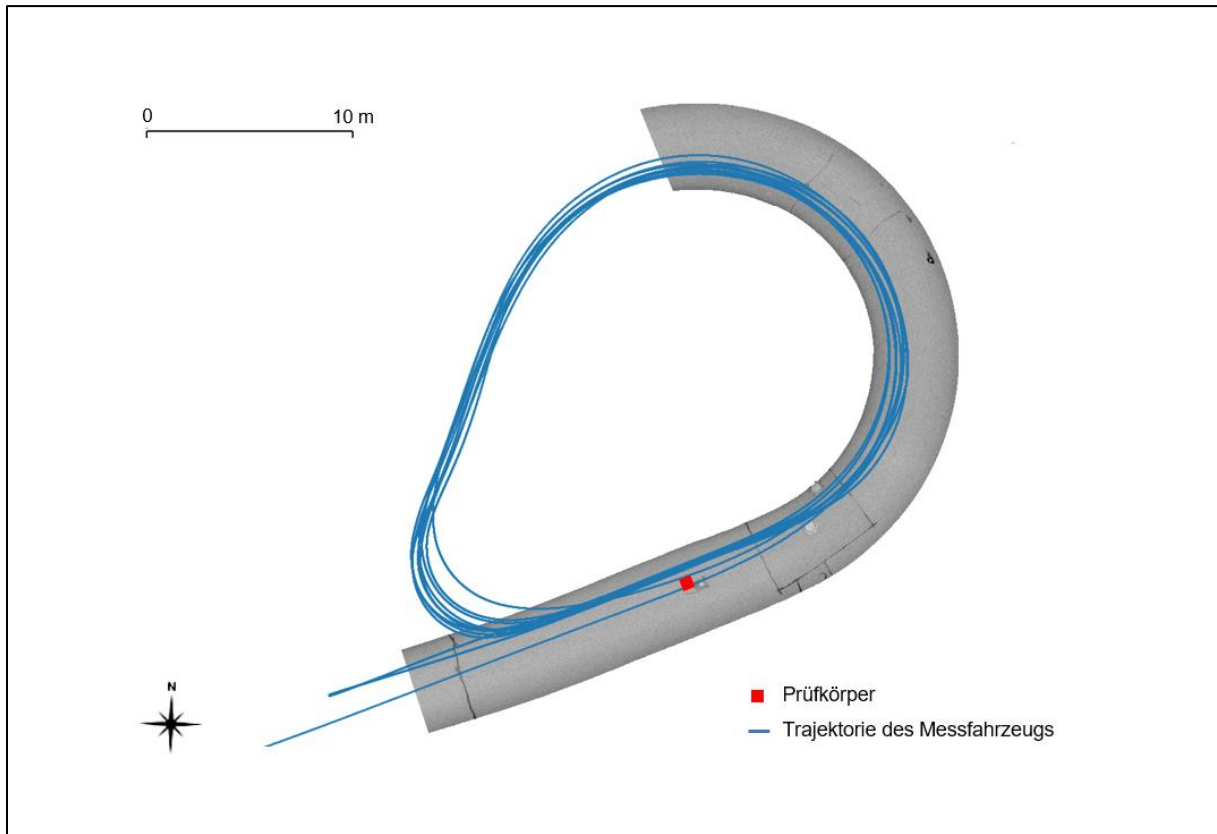


Abbildung 91: Darstellung der Messfahrzeug-Trajektorie im Zuge der Erfassung des Datensatzes zur Untersuchung der Präzision des Messsystems. Der virtuelle Nullpunkt des Positionierungssystems liegt am Kontaktpunkt des linken Hinterrads mit der Fahrbahnoberfläche. Im Hintergrund ist der Laserscan der ersten Messung mit in Graustufen kodierter Reflexionsintensität dargestellt.

Der mit dem mobilen Laserscanning System erfasste Datensatz beinhaltet insgesamt acht Messungen des Prüfkörpers, die bei einer jeweils sehr geringen Geschwindigkeit von ca. 8 km/h aufgenommen wurden. Die Punktdichten sind daher entsprechend hoch: Jeder Scan der Prüfkörperoberfläche besteht aus ca. 21000 3D-Punkten, was einer Punktdichte von ca. 200 Punktnachbarn in einer Kugel mit Radius $r = 0,03$ m entspricht.



Abbildung 92: Aufnahme des Datensatzes zur Analyse der Präzision des mobilen Laserscanning Systems. Insgesamt wurden acht Laserscans des Prüfkörpers in einem Zeitraum von ca. 4 Minuten aufgezeichnet. Der Prüfkörper (I) wurde so positioniert, dass die geneigten Seiten des mittleren Elements in Fahrtrichtung ansteigen und abfallen sowie die Quaderelemente quer zur Fahrtrichtung verlaufen (II und III).

Die Genauigkeit der Trajektorie wird von der Post-Processing Software für die Epochen der Datenaufnahme mit einem RMSE von ca. 0,013 m für die Lagekomponente und ca. 0,059 m für die Höhenkomponente angegeben (siehe Abbildung 93). Die einzelnen Messungen zeichnen sich in Abbildung 93 in Form von Spitzen ab, die sich in den RMSE-Kurven der Northing-, Easting und Down-Komponente wenige Millimeter über das vorherrschende Grundniveau des RMSE erheben. Die Peaks resultieren mutmaßlich aus der engen Kurvenfahrt, die das Messfahrzeug durchführen musste, um den Prüfkörper erneut anfahren und aufnehmen zu können. Zwischen den ersten beiden Messungen vergingen etwa 50 Sekunden. Die restlichen sieben Messungen wurden im direkten Anschluss aneinander eingefahren. Die gesamte Datenaufnahme dauerte ca. 4 Minuten.

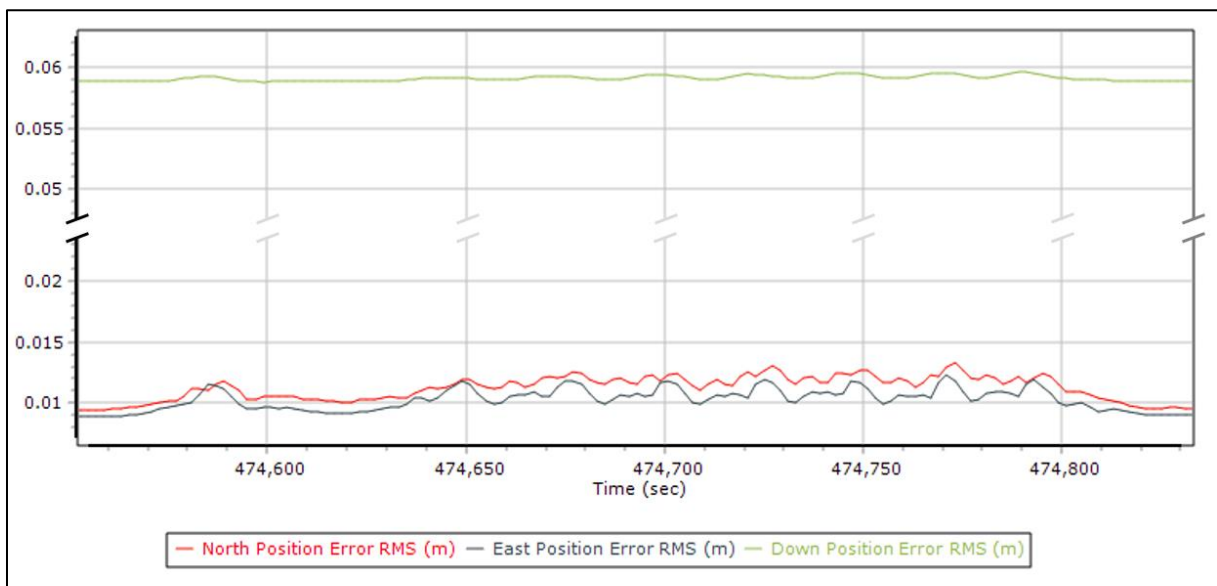


Abbildung 93: Darstellung der X,Y,Z-Komponente des Positionierungsfehlers in der Applanix Post-Processing Software POSPac MMS. Auf der Y-Achse ist der RMSE abgetragen und auf der X-Achse die Zeit in Form der sogenannten Epochen.

In Abbildung 94 ist verdeutlicht, wie sich die Genauigkeit der Trajektorie in der absoluten Genauigkeit der Laserscans widerspiegelt. Anhand des Längsschnitts durch die acht unabhängigen Laserscans ist zu sehen, dass ein Versatz der gescannten Prüfkörper zueinander sowohl in der Lage, als auch in der Höhe vorherrscht. Die Abweichungen entlang der drei Raumachsen entsprechen weitestgehend der in Abbildung 93 dargestellten Performance-Metrik der Trajektorien Post-Processing Software.

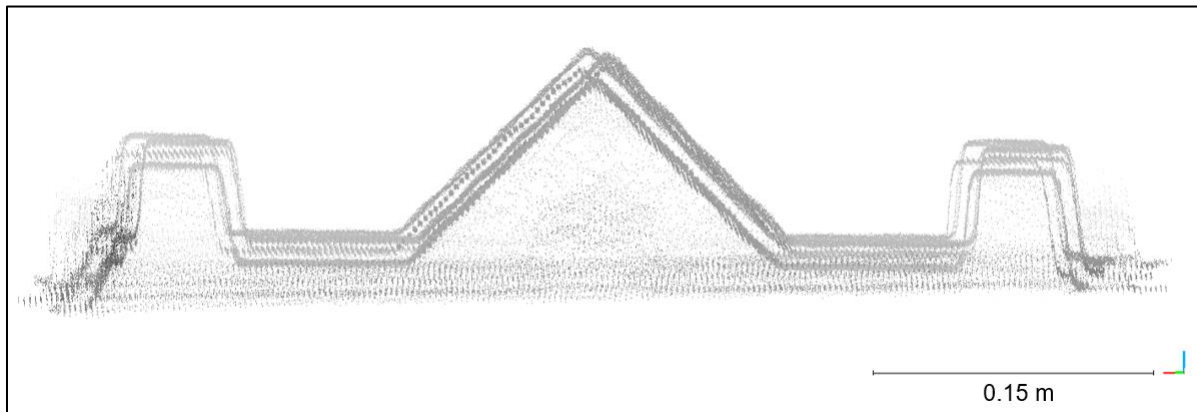


Abbildung 94: Längsschnitt durch die Punktwolken der acht Laserscans des Prüfkörpers. Die dargestellten Abweichungen repräsentieren die Schwankungen der absoluten Positionierungsgenauigkeit. In Anhang V sind statistische Werte zur Trajektoriengenauigkeit aufgeführt.

3.2.3 Aspekt: Oberflächenmodellierung

Für den Vergleich von Interpolationsmethoden, die zur Generierung von Oberflächenmodellen eingesetzt werden können, wurden zwei mit dem Fahrzeug I.R.I.S 13 erfasste Oberflächenscans von kommunalen, in Asphaltbauweise ausgeführten Straßen genutzt. Die Festlegung und Begrenzung der Testareale erfolgte so, dass sie jeweils ca. 100000 3D-Messpunkte über die gesamte erfassbare Fahrbahnbreite beinhalten. Die Auswahl der Testareale erfolgte willkürlich aus Befahrungsdaten des kommunalen Straßennetzes. Dabei sind die betreffenden Laserscans mit unterschiedlichen Geschwindigkeiten erfasst worden und beinhalten typische Schäden und Objekte (siehe Abbildung 95 und Abbildung 96).

In Abbildung 95 ist das ca. 4,30 m breite und ca. 0,8 m lange Testareal 1 dargestellt. Hellere Bereiche in der obersten Teilabbildung entsprechen Bereichen mit hoher Reflektanz, wie beispielsweise die Fahrbahnmarkierung im linken Bereich. In der mittleren und unteren Teilabbildung liegen rote Bereiche am höchsten und blaue am niedrigsten. Die Darstellung der relativen Höhe dient hier nur der besseren Visualisierung des Querrisses, der in der Darstellung der absoluten Höhe vom globalen Trend der Höhenwerte überprägt wird. Die Höhenwerte im

Testareal 1 weisen eine Spannweite von ca. 0,17 m auf, welche maßgeblich auf die Querneigung des gescannten Fahrstreifens von ca. 3 % zurückzuführen ist. Aufgrund der relativ geringen Befahrungsgeschwindigkeit von ca. 25 km/h liegt eine entsprechend hohe Dichte der 3D-Messpunkte im Testareal 1 vor. Der Punktabstand quer zur Fahrt- und Scanrichtung beträgt geschwindigkeitsunabhängig ca. 4,5 mm und der Profilabstand in Fahrtrichtung geschwindigkeitsabhängig ca. 9 mm. Im Mittel befinden sich innerhalb einer Kugel mit Radius $r = 0,03$ m um einen 3D-Messpunkt ca. 79 Nachbarpunkte.

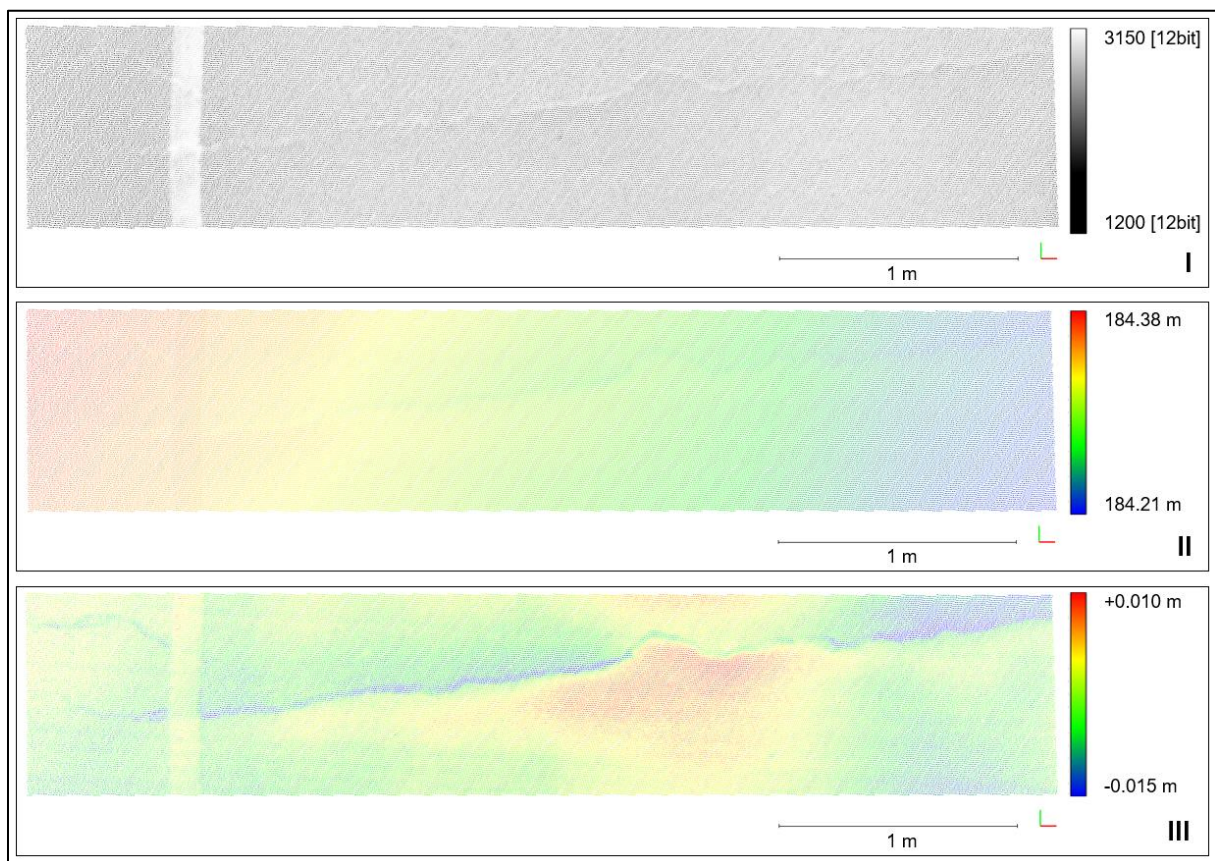


Abbildung 95: Testareal 1 für die Untersuchungen zur Oberflächenmodellierung aus dem Laserscan. Dargestellt ist von oben nach unten die Reflexionsintensität (I), die absolute Höhe (II) und die relative Höhe (III).

In Abbildung 96 ist das ca. 3,90 m breite und ca. 1,43 m lange Testareal 2 dargestellt. Der Laserscan enthält sowohl radiometrisch relevante Information in Form eines sanierten Risses, wie auch geometrisch relevante Information in Form eines Bordsteins und eines Schachtdeckels. Die Höhenwerte weisen eine Spannweite von ca. 0,14 m auf, wovon ca. 0,04 m auf den abgesenkten Bordstein entfallen. Da die Befahrungsgeschwindigkeit mit ca. 40 km/h deutlich höher war als in Testareal 1, ist auch der Profilabstand in Fahrtrichtung mit ca. 14 mm größer. Im Mittel befinden sich innerhalb einer Kugel mit Radius $r = 0,03$ m um einen 3D-Messpunkt ca. 52 Nachbarpunkte.

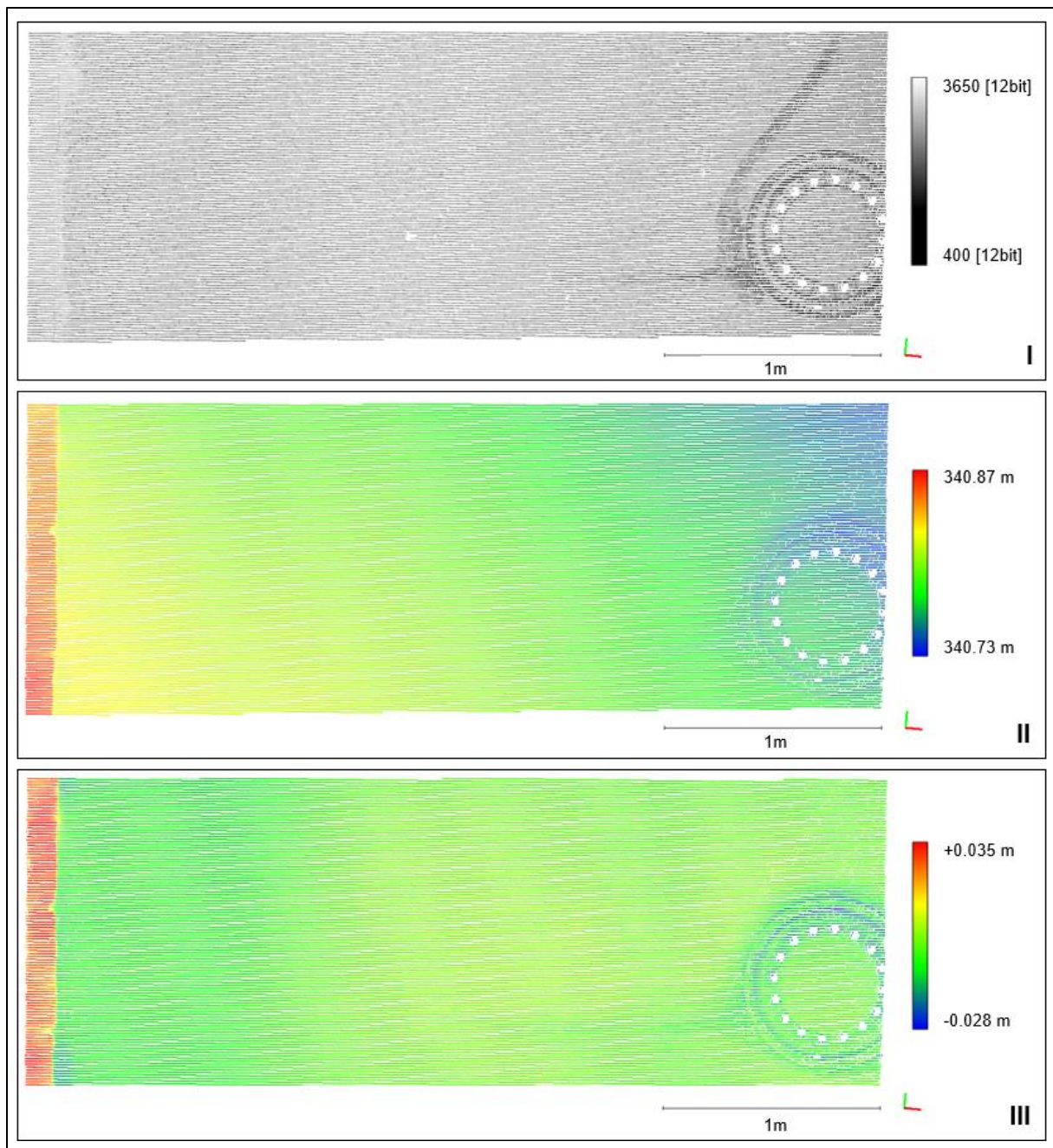


Abbildung 96: Testareal 2 für die Untersuchungen zur Oberflächenmodellierung aus dem Laserscan. Dargestellt ist von oben nach unten die Reflexionsintensität (I), die absolute Höhe (II) und die relative Höhe (III).

In Abbildung 97 ist die Punktdichtevertellung von Testareal 1 und Testareal 2 in Form von Histogrammen gegenübergestellt. Die x-Achse repräsentiert die Anzahl benachbarter Punkte innerhalb einer 0,03 m Nachbarschaft. Auf der y-Achse sind die Fallzahlen als Absolutwerte abgetragen. Geringe Punktdichten von < 30 Nachbarpunkten finden sich lediglich an den äußersten Rändern der Testareale.

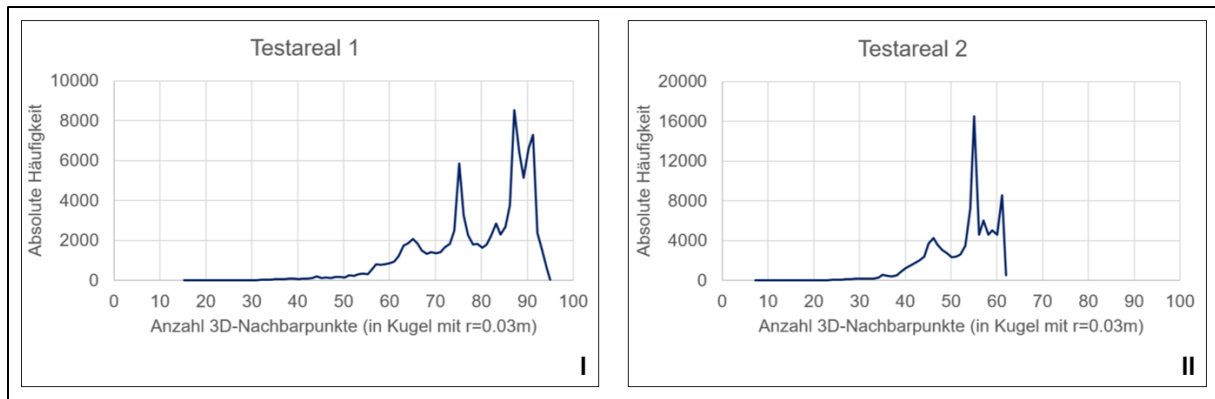


Abbildung 97: Histogramme der Punktdichte in den beiden Testarealen bei ca. 25 km/h (I) und bei ca. 40 km/h (II).

3.2.4 Aspekt: Erkennbarkeit von Oberflächendiskontinuitäten

Im Rahmen der Straßenzustandserfassung ist es vorteilhaft, Schädigungen der Substanz im frühestmöglichen Stadium erkennen zu können (siehe Kapitel 2.1). Aufgrund ihrer hervorgehobenen Bedeutung im Rahmen der Zustandsbewertung und im Erhaltungsmanagement bedeutet das, dass vor allem Risse bereits dann erkannt werden sollten, wenn sie noch nicht weit aufklaffen (STASCHKIEWICZ & RADENBERG 2022).

Um untersuchen zu können, bis zu welcher Ausprägung Risse in den Oberflächenmodellen erkannt werden können, die aus den mobilen Laserscandaten erzeugt wurden, wurde ein Prüfkörper angefertigt, bei dem Spalten definierter Tiefe und Breite die zu erkennenden Oberflächendiskontinuitäten repräsentieren (SESSELMANN et al. 2022). Der Prüfkörper besteht aus Fichtenholz-Rechteckleisten mit 20 mm Stärke und 60 mm Breite, die auf eine 9 mm starke Siedruckplatte aufgeklebt wurden. Insgesamt misst der Prüfkörper 623 mm x 647 mm x 30 mm (siehe Abbildung 98), sodass er berührungslos vom Messfahrzeug überrollt werden kann.

Die Rechteckleisten wurden so parallel auf der Siebdruckplatte angeordnet, dass zwischen den einzelnen Leisten Spalten mit definierter Breite entstanden. Geplant waren die Abstände 1 mm, 2 mm, 5 mm, 7 mm, 12 mm, 15 mm, 20 mm und 30 mm. Anschließend wurde der Prüfkörper mit grauem und schwarzem Sprühlack matt lackiert. Durch die Eigenschaften des verwendeten Klebstoffes und der Lackierung konnten die vordefinierten Spaltmaße nicht exakt eingehalten werden. Daher wurden die Spaltmaße nachträglich mit Stahl-Endmaßen der Toleranzklasse 1 bestimmt. Diese auch als Gauge Blocks bezeichneten Stahlkörper sind auf wenige Mikrometer genau dimensioniert (siehe Anhang IV).

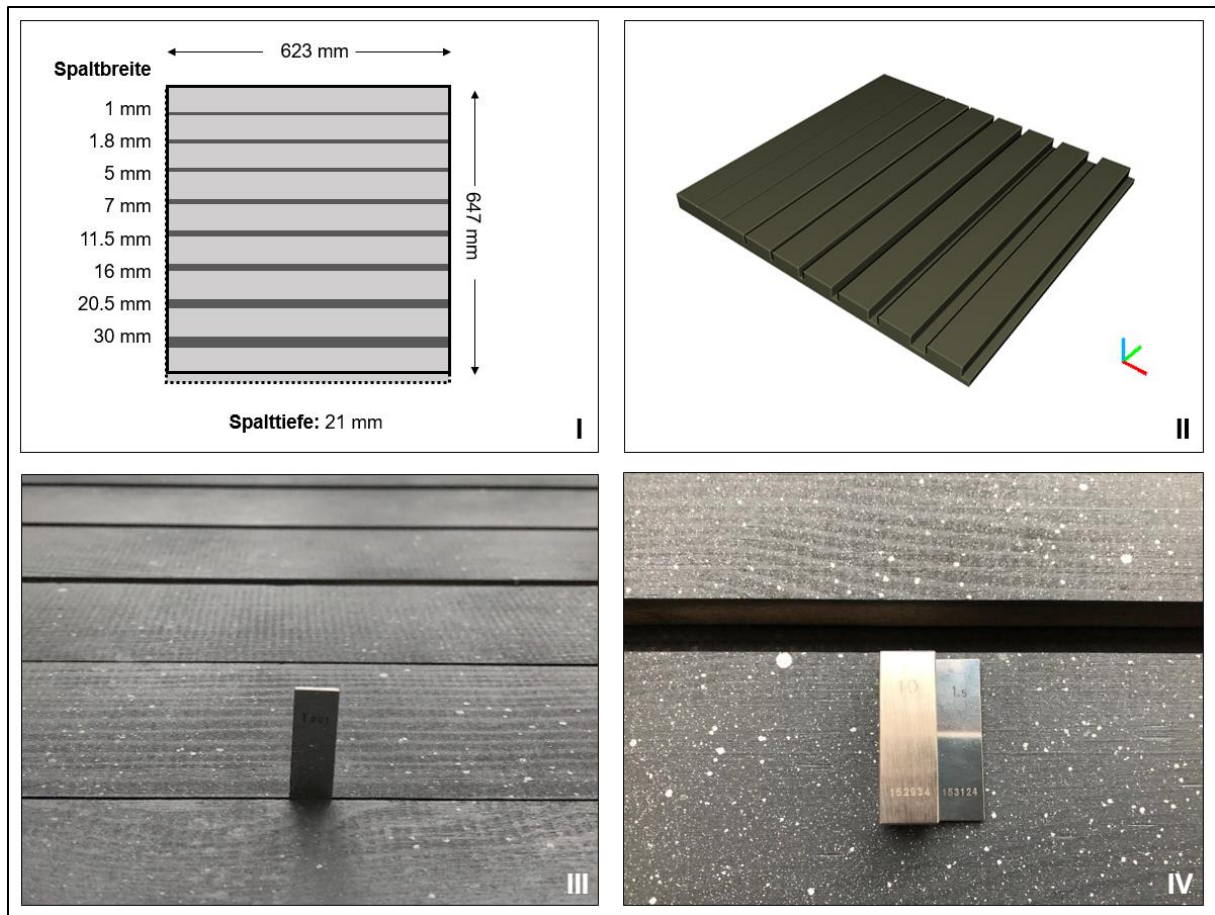


Abbildung 98: Die Teilabbildung I zeigt den Prüfkörper als Konstruktionsskizze in der Draufsicht inklusive Bemaßung. Die Teilabbildung II zeigt das darauf basierend konstruierte 3D-Modell in der Schrägansicht. Die unteren beiden Teilabbildungen III und IV zeigen die Bestimmung der Spaltbreite mithilfe von Stahl-Endmaßen.

Am 12.06.2020 fand die Messreihe mit dem Messfahrzeug I.R.I.S 16 statt, in deren Rahmen der Prüfkörper aufgenommen wurde. Abbildung 99 dokumentiert die Situation während der Aufnahme der Messreihe. Der Prüfkörper wurde zunächst mittig auf einem abgesperrten Feldweg bei 30 km/h gescannt. Es fanden je drei Wiederholungsmessungen statt, bei denen die Spalten des Prüfkörpers parallel und orthogonal zur Fahrtrichtung ausgerichtet waren. Nach derselben Vorgehensweise wurden im Anschluss bei 50 km/h weitere Messungen unternommen.



Abbildung 99: Messsystem I.R.I.S. 16 während der Messung des Prüfkörpers in der Messanordnung, in der die Spalten längs zur Fahrtrichtung verlaufen

Abbildung 100 zeigt exemplarisch links den Prüfkörper, der so ausgerichtet ist, dass dessen Spalten quer zur Fahrtrichtung, annähernd parallel zu den Scanprofilen, verlaufen. Rechts ist das Ergebnis der Messung in Form einer 3D-Punktwolke abgebildet. Die geplante Messreihe mit 80 km/h konnte aufgrund einsetzenden Regens nicht mehr realisiert werden. In Abbildung 99 ist die Regenfront im Hintergrund bereits zu sehen. Während der erfolgreich durchgeführten Messungen herrschten jedoch durchgängig trockene Verhältnisse.

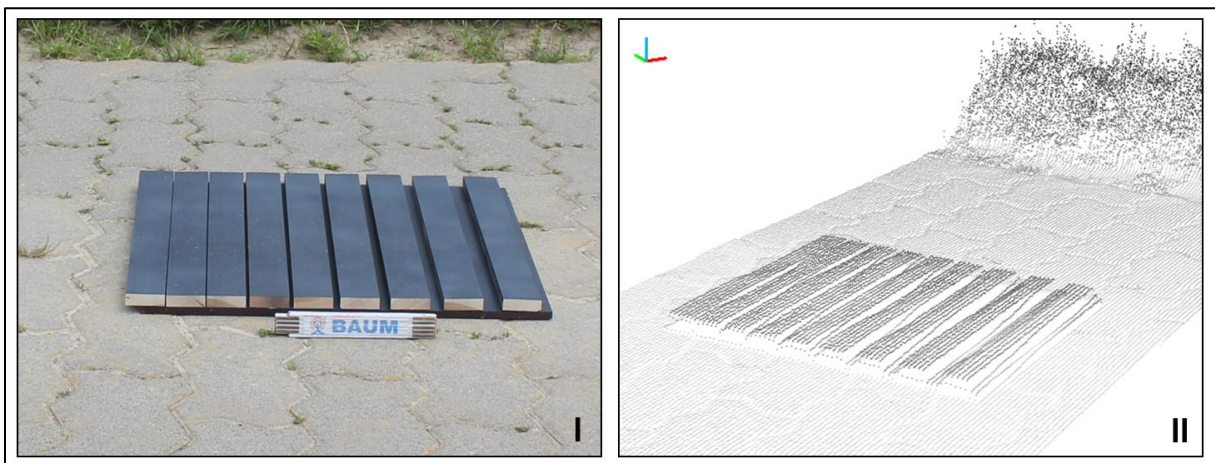


Abbildung 100: Die linke Teilabbildung (I) zeigt ein Foto des Prüfkörpers mit einem Zollstock im Vordergrund kurz vor der Messung. Die rechte Teilabbildung (II) zeigt eine Schrägansicht auf die mit dem PPS aufgenommene 3D-Punktwolke des Prüfkörpers bei 30 km/h.

3.2.5 Aspekt: Trainings- und Testdaten für Deep-Learning-Modelle

Für das Training eines robusten und generalisierungsfähigen Deep-Learning-Modells ist eine umfangreiche und möglichst divers ausgeprägte Datenbasis erforderlich. Deshalb wurden im Rahmen der Untersuchungen dieser Arbeit im Zeitraum von 2015 bis 2020 Straßenoberflächen in mehr als 30 verschiedenen Regionen aus elf Bundesländern Deutschlands mit dem I.R.I.S System dreidimensional aufgenommen. Der resultierende PPS-Datenstrom wurde, in Anlehnung an die ZTV ZEB-Stb (FGSV 2006) und TP Oberflächenbild-StB (FGSV 2020), in 10 m lange Punktwolkenabschnitte unterteilt. Abbildung 101 zeigt die räumliche Verteilung der verwendeten Daten. Die Datenauswahl erfolgte insofern zunächst zufällig, als dass überwiegend Überführungsfahrten zu Messeinsatzorten der I.R.I.S Systeme für die Datenaufnahme genutzt wurden. Die so generierte Datenbasis wurde, oft auch unter Zuhilfenahme der Frontkamerabilddaten, gesichtet und brauchbare Laserscans iterativ selektiert. Selektionskriterien waren das Vorhandensein von Schäden und Objekten sowie allgemeine Besonderheiten der Deckschichtmaterialien und Oberflächenanmutung, wie beispielsweise Verschmutzungen oder Alterungserscheinungen.

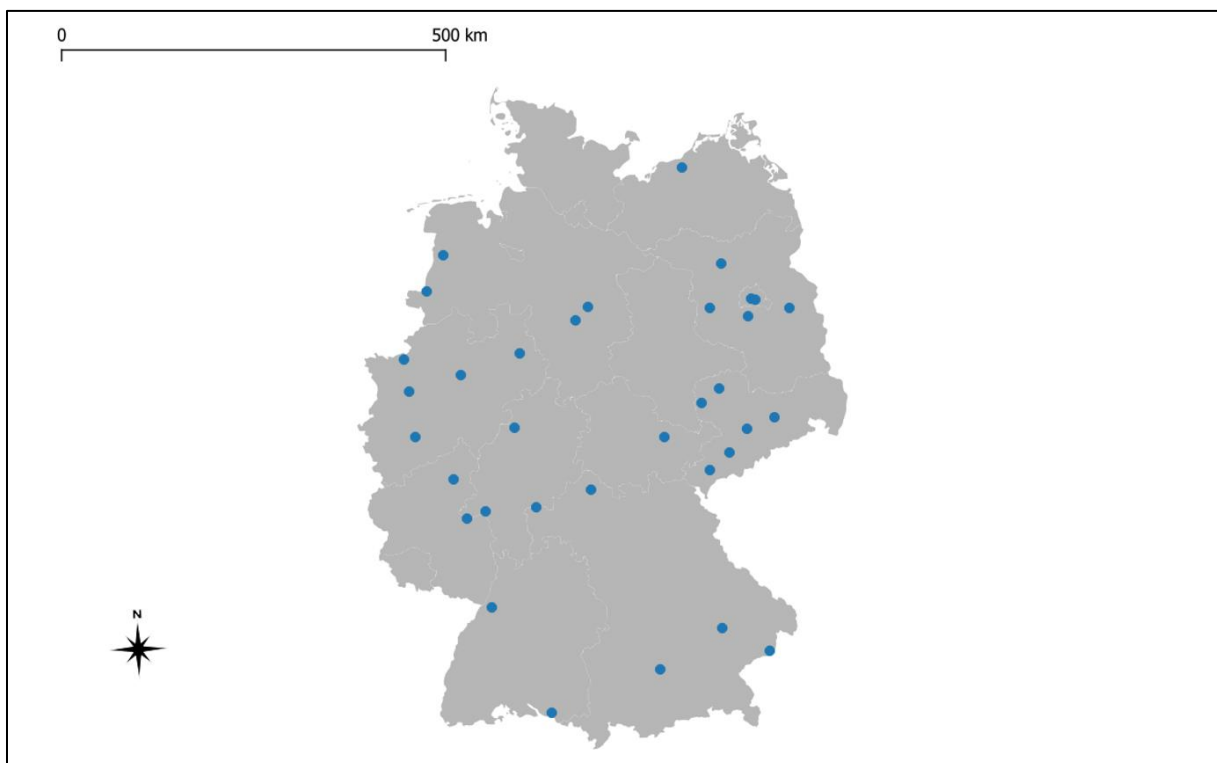


Abbildung 101: Übersichtskarte zur räumlichen Verteilung der Trainings- und Testdaten (Punkte). Für die Darstellung wurden im Hintergrund freie Geobasisdaten verwendet⁴.

⁴ Website: <https://opendata-esri-de.opendata.arcgis.com/datasets/esri-de-content::vg250-gemeindegrenzen/about>

Die so erzeugte Datenbasis umfasst insgesamt 2544 Laserscans. Das Verhältnis von Außerorts- zu Innerortsstrecken insgesamt beträgt etwa 1:2, wobei Innerortsstrecken zusätzlich wesentlich diverser vertreten sind. Der überwiegende Teil der im Datensatz vertretenen Außerortsstrecken entfällt auf den Landkreis Hameln-Pyrmont, den Rhein-Sieg-Kreis und den Bodenseekreis. Entsprechend der mittleren Geschwindigkeiten, die mit dem Messsystem im fließenden Verkehr gefahren werden können, liegen typische Punktdichten außerhalb geschlossener Ortschaften bei im Mittel 35 (σ : 5 Punkte) und innerorts bei im Mittel 85 Punkten (σ : 10 Punkte). Die Punktdichte bezieht sich hierbei auf die Anzahl von Nachbarpunkten in einer Kugel mit Radius $r = 0,03$ m. Der PPS-Laserscan mit der geringsten Punktzahl umfasst ca. 0,5 Millionen 3D-Punkte und weist eine mittlere Punktdichte von 34 Punkten auf. Der Laserscan mit der höchsten Punktzahl umfasst ca. 8,6 Millionen 3D-Punkte und seine Punktdichte beträgt 705 Punkte.

Die beschriebenen Laserscans bilden die Basis für die Berechnung von Oberflächenmodellen, die für das Training von tiefen neuronalen Netzwerkmodellen sowie deren Evaluierung verwendet werden (SESSELMANN et al. 2023).

3.2.6 Aspekt: Evaluierung des Deep Learning basierten Analysesystems

Um das auf Mobile Lasercanning und Deep Learning basierende Analysesystem dahingehend zu evaluieren, ob es die Anforderungen der geltenden Regelwerke und Richtlinien an die Auswertequalität erfüllen kann, wurde am 12.06.2020 mit den Messfahrzeugen I.R.I.S 3 und I.R.I.S 16 eine ca. 2,4 km lange, innerorts gelegene Strecke bei Dresden aufgenommen.

Die Straße ist insgesamt ca. 6,5 m breit und umfasst je einen Fahrstreifen pro Fahrtrichtung, die nicht durch eine Markierungslinie voneinander getrennt sind. Die Asphaltdecke der betreffenden Straße weist eine Vielzahl offener sowie bereits sanierter Schadstellen auf. Bis auf die Schadensarten Ausmagerung und Bindemittelaustritt kommen alle gemäß der Richtlinie E EMI (FGSV 2012) relevanten Schadenstypen auf der Messstrecke vor. Zur Entwässerung schließen sich an die Asphaltdecke seitlich gepflasterte Rinnen und bereichsweise Bordsteine an. Auch kommen zahlreiche Abflüsse im Bereich der Rinnen, Schachtdeckel und Wasserschieberkappen auf der Fahrbahnoberfläche vor. Die Bedingungen waren für die Datenaufnahme günstig: die Fahrbahnoberfläche war trocken und es war wechselnd sonnig und locker bewölkt.

Da das TP3-System (Oberflächenkamera mit Beleuchtungssystem) des Messfahrzeugs I.R.I.S 16 zum Messtermin noch nicht voll betriebsbereit war, wurde zusätzlich das Mess-

fahrzeug I.R.I.S 3 eingesetzt, welches über eine gültige zeitbefristete Betriebszulassung der BASt für das Teilprojekt TP3 gemäß ZTV ZEB-Stb (FGSV 2006) verfügte. Somit konnten mit dem I.R.I.S 3 System regelwerkskonforme Oberflächenbilder mit einer Pixelauflösung von 0,89 mm aufgenommen werden, wie sie auch im Rahmen von bauvertraglich festgelegten Abnahmeverfahren oder der ZEB von Autobahnen verwendet werden (siehe Abbildung 102). Auf dieser Grundlage wurde der Oberflächenzustand durch geschultes Auswertepersonal der Firma Lehmann+Partner GmbH gemäß der geltenden Richtlinie (FGSV 2006) erfasst und eine Referenzauswertung für die Evaluierungsstrecke geschaffen.

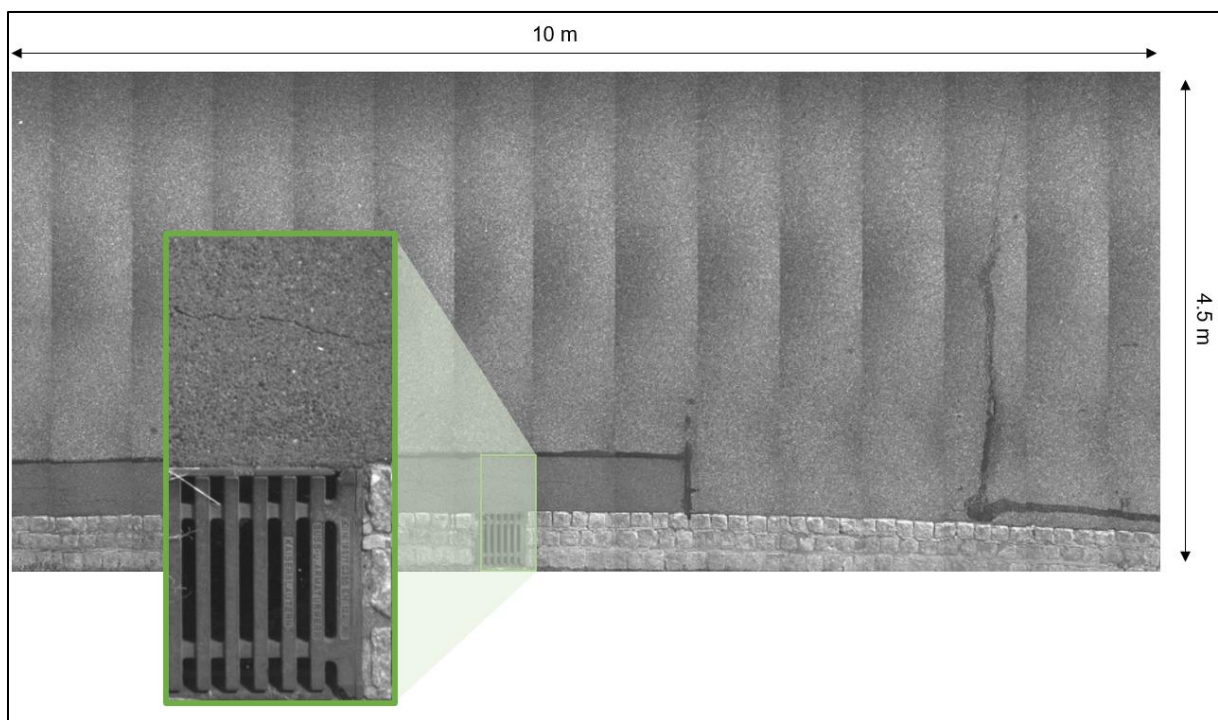


Abbildung 102: Beispiel eines TP3-Oberflächenbildes, das den Anforderungen der BASt für die Erfassung von Oberflächenschäden auf Bundesautobahnen entspricht.

Abbildung 103 zeigt das Messfahrzeug I.R.I.S 16 während der Datenaufnahme im fließenden Verkehr. Die Befahrungsgeschwindigkeit lag im Mittel bei 37 km/h und die Standardabweichung bei 5 km/h. Dementsprechend liegen die mittleren Punktdichten der aufgenommenen Laserscans im Bereich von 37 Punkten (σ : 5 Punkte) bis 75 Punkten (σ : 10 Punkte) in einer Kugel mit Radius $r = 0,03$ m. Die Profilabstände liegen im Bereich zwischen 9 mm und 17 mm.



Abbildung 103: Situation während der Datenaufnahme auf der Evaluierungsstrecke. Das Bild zeigt I.R.I.S 16 aus Sicht des hinterherfahrenden Fahrzeugs.

Im Umkreis von 50 km befinden sich keine Straßenabschnitte, die als Trainingsdaten des Deep-Learning-Verfahrens genutzt wurden. Insofern stellen die bei Dresden aufgenommenen Oberflächenscans unbekannte Daten für die im Rahmen der Untersuchungen trainierten tiefen neuronalen Netze dar. Abbildung 104 zeigt ein Beispiel für eine mit dem I.R.I.S 16 System aufgenommene 3D-Punktwolke der Evaluierungsstrecke.

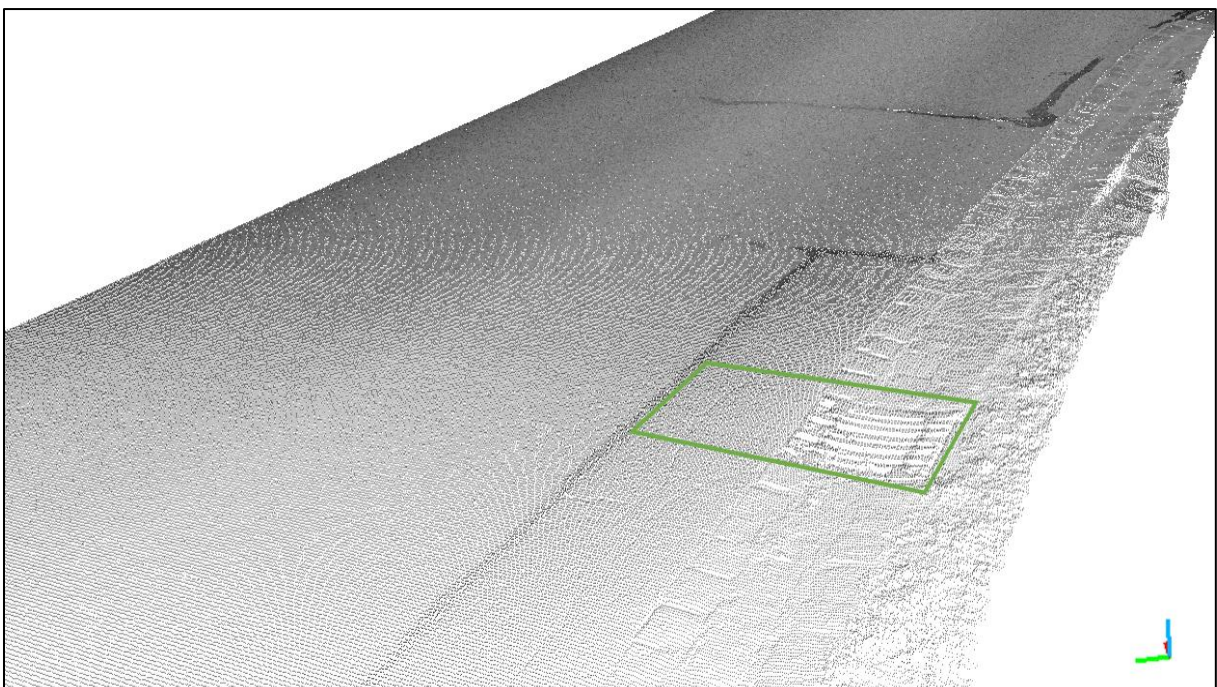


Abbildung 104: Beispiel für einen Laserscan der Evaluierungsstrecke. Die grüne Box markiert dieselbe Stelle, die auch in Abbildung 102 vergrößert dargestellt ist.

Die beiden beschriebenen Datensätze (PPS-Laserscans von I.R.I.S 16 und TP3-Oberflächenbilder von I.R.I.S 3) dienen dazu, die Qualität der Deep Learning basierten, automatischen Zustandserfassung und -bewertung auf Basis von PPS-Laserscans mit dem aktuellen deutschen State-of-the-Art-Verfahren einer manuellen Straßenzustandserfassung durch geschultes Auswertepersonal zu vergleichen.

Zusätzlich konnten auf der Evaluierungsstrecke am 14.07.2020 mit dem Messsystem I.R.I.S 17, welches von der Firma Nievelt Labor GmbH (NIEVELT LABOR GMBH 2022) betrieben wird, Daten mit dem LCMS2-System (LAURENT et al. 2012; PAVEMENTRICS 2018) aufgenommen werden. Aus logistischen Gründen konnte keine zeitgleiche Befahrung mit dem I.R.I.S 17 System erfolgen und die befahrene Strecke deckt nur einen ca. 1,2 km langen Teil der Strecke ab, die mit I.R.I.S 3 und I.R.I.S 16 etwa einen Monat zuvor befahren wurde.

Das verwendete LCMS2-System besteht aus zwei baugleichen Lasersensoren, die am Fahrzeugheck links und rechts, auf die Fahrbahnoberfläche ausgerichtet, montiert sind (siehe Abbildung 105). Mithilfe des Laserlichtschnittverfahrens wird die Oberflächenmorphologie und die Reflexionsintensität erfasst und in Form von Bilddaten mit einer Pixelauflösung von 1 mm überführt. Insgesamt decken die beiden LCMS2-Sensoren eine Breite von 4 m ab (siehe Abbildung 106). Weitere technische Spezifikationen sind bei PAVEMENTRICS (2018) zu finden. Für den Zeitraum der ca. zehnminütigen Messung musste der Straßenabschnitt für den öffentlichen Verkehr gesperrt werden, da die LCMS2-Sensoren nicht augensichere Laser der Klasse 3B verwenden.



Abbildung 105: Vergleichsmessung mit dem Messfahrzeug I.R.I.S 17 der Firma Nievelt Labor GmbH (NIEVELT LABOR GMBH 2022).

Alle beteiligten Messfahrzeuge waren mit Applanix POS LV Positionierungssystemen ausgerüstet und die Messstrecke war vergleichsweise kurz. So konnte die räumliche Synchronisierung der Sensordaten der einzelnen Messfahrzeuge auf wenige Dezimeter genau erreicht werden.

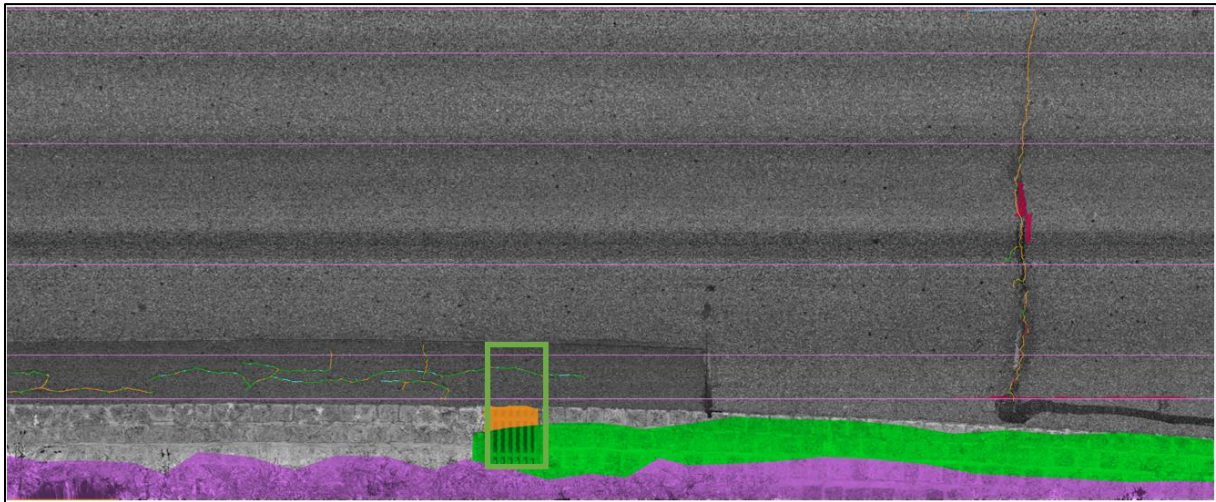


Abbildung 106: Oberflächenbild aus den LCMS2-Sensoren des I.R.I.S 17 Systems. Überlagert sind die Ergebnisse der automatischen Schadenserkenennung (LcmsRoadInspect Software). Die grüne Box markiert dieselbe Stelle, die auch in Abbildung 102 vergrößert dargestellt ist.

Mithilfe dieses Datensatzes kann das Deep Learning basierte Analysesystem auf die PPS-Laserscans angewendet werden und die Ergebnisse der automatischen Schadensdetektion gegen Referenzauswertungen verglichen werden, die sowohl den nationalen als auch internationalen Standard repräsentieren.

4 Methodik

In diesem Kapitel werden die Werkzeuge und Methoden erklärt, mithilfe derer die Untersuchungen der Arbeit durchgeführt wurden.

4.1 Arbeitsumgebung

Zur Bearbeitung der Fragestellungen dieser Arbeit wurden verschiedene Softwarelösungen genutzt: von 3D-Punktwolken Analysesoftware über GIS-Anwendungen, Datenbankmanagementsysteme, bis hin zu Entwicklungsumgebungen und Deep Learning Frameworks. Die Auswahl der verwendeten Werkzeuge erfolgte unter der Prämisse, weitgehend freie und quelloffene Software, sowie offene Standards zu verwenden. STEINIGER & BOCHER (2009) geben einen Überblick über verschiedene Lizenzmodelle freier und quelloffener Software mit Fokus auf den GIS-Bereich. In den folgenden Unterkapiteln werden die wichtigsten verwendeten Softwarelösungen und Standards gruppiert nach dem Anwendungskontext vorgestellt.

Alle Berechnungen und Arbeiten, die nicht direkt mit dem Training und der Evaluation der Deep-Learning-Modelle in Verbindung standen, wurden unter dem Betriebssystem Windows 10 mit einem Intel i7-1165G7 2,8 GHz Prozessor, 16 GB RAM und einer NVIDIA T500 Grafikkarte durchgeführt. Während die in Kapitel 4.1.1, 4.1.3 und 4.1.5 beschriebenen Anwendungen keine hohen Hardwareanforderungen stellen, ist für das Training und die Anwendung von Deep-Learning-Modellen spezielle Hardware nötig (siehe Kapitel 4.1.4). Es wurde hierfür eine Workstation mit einer Intel Core i9-10900X 10x 3,7 GHz CPU, 32 GB DDR4 RAM und einer NVIDIA Titan RTX Grafikkarte konfiguriert. Zum Zeitpunkt ihrer Markteinführung Ende 2018 war sie die schnellste im Consumer-Bereich verfügbare Grafikkarte (siehe Tabelle 10). Sie ist optimiert für Raytracing und KI-Anwendungen.

Tabelle 10: Technische Daten der verwendeten GPU⁵.

Name	NVIDIA TITAN RTX
Architektur	NVIDIA Turing
Framebuffer	24 GB GDDR6
Boost-Takt	1770 MHz
Tensor-Recheneinheiten	576
CUDA-Recheneinheiten	4608

⁵ Website: <https://www.nvidia.com/de-at/titan/titan-rtx/>

4.1.1 Trajektorienoptimierung und 3D-Punktwolkenerzeugung

Das Post-Processing der Trajektorien des Messfahrzeugs für die verschiedenen Datenaufnahmekampagnen (siehe Kapitel 3.2) wurde mit der proprietären Software POSpac MMS durchgeführt, die Teil des Applanix POS LV Positionierungssystems ist. Abbildung 107 zeigt das Nutzerinterface der Software POSpac MMS in der Version 8.6.

Für jede Messfahrt wurden die Rohdaten des inertialen Positionierungssystems zunächst verbessert, indem präzise Ephemeriden bezogen und integriert wurden. Außerdem wurde jeweils ein sogenanntes SmartBase Netz aus DGNSS-Referenzstationen berechnet. Der SmartBase Ansatz basiert auf dem Konzept der virtuellen GNSS-Referenzstation (SCHERZINGER & HUTTON 2023). Dabei werden die Beobachtungen realer Referenzstationen in einem Netz ausgeglichen und dann virtuelle Stationen in optimaler Nähe zur Rover-Position erzeugt. Damit kann die Positionierungslösung des Messfahrzeugs verbessert werden, solange dessen Trajektorie innerhalb des Netzes liegt. Aus den Rohdaten der Sensorkomponenten (GNSS, IMU, DMI) sowie den Korrekturdaten des SmartBase-Netzes berechnet POSpac MMS mithilfe eines Kalman Filters eine optimierte Positionierungslösung. Das Ergebnis sind verbesserte Trajektorien, die zur Georeferenzierung der Laserscannermessungen und somit zur 3D-Punktwolkenerzeugung verwendet werden. Die Erzeugung von 3D-Punktwolken aus dem scanner-spezifischen iaps-Datenformat und den verbesserten Trajektorien geschieht mithilfe einer internen Software der Firma Lehmann + Partner GmbH. Sie setzt die in Kapitel 2.2.5 beschriebene Vorgehensweise um.

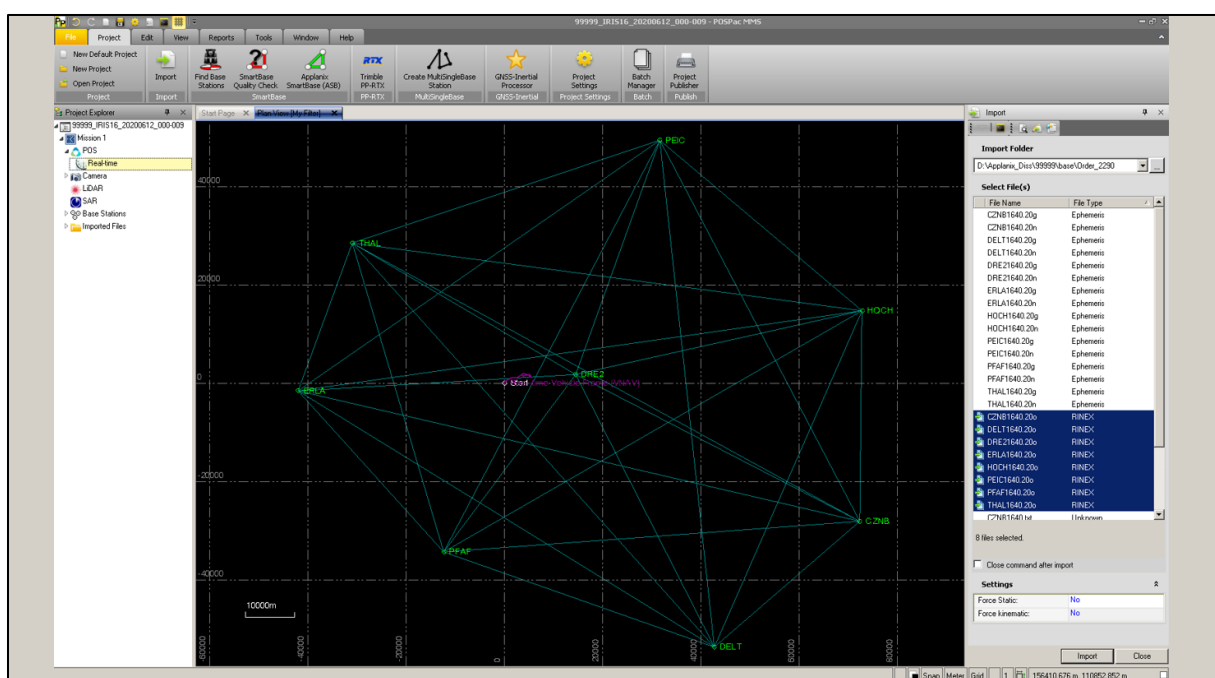


Abbildung 107: User Interface der Software Applanix POSpac MMS 8.6.

4.1.2 3D-Prozessierung und Visualisierung

Für die Arbeit mit dichten 3D-Punktwolken wurde das Open-Source-Projekt CloudCompare in der Version 2.12.4 genutzt (siehe Abbildung 108). CloudCompare ist eine Bearbeitungs-, Analyse- und Visualisierungssoftware für 3D-Punktwolken und Dreiecksvermaschungen. CloudCompare ist in C++ entwickelt und auf Windows, Linux und Mac OS und für 32 bit und 64 bit Architekturen kompiliert. Seit 2009 ist dieses Tool unter der GPL-Lizenz öffentlich verfügbar (CLOUDCOMPARE 2022). Im Kontext dieser Arbeit wurde CloudCompare für 3D-Visualisierungen im Allgemeinen sowie 3D-Distanzbestimmungen, 3D-Modellierungen und 3D-Registrierungen (Iterative Closest Point, ICP) verwendet.

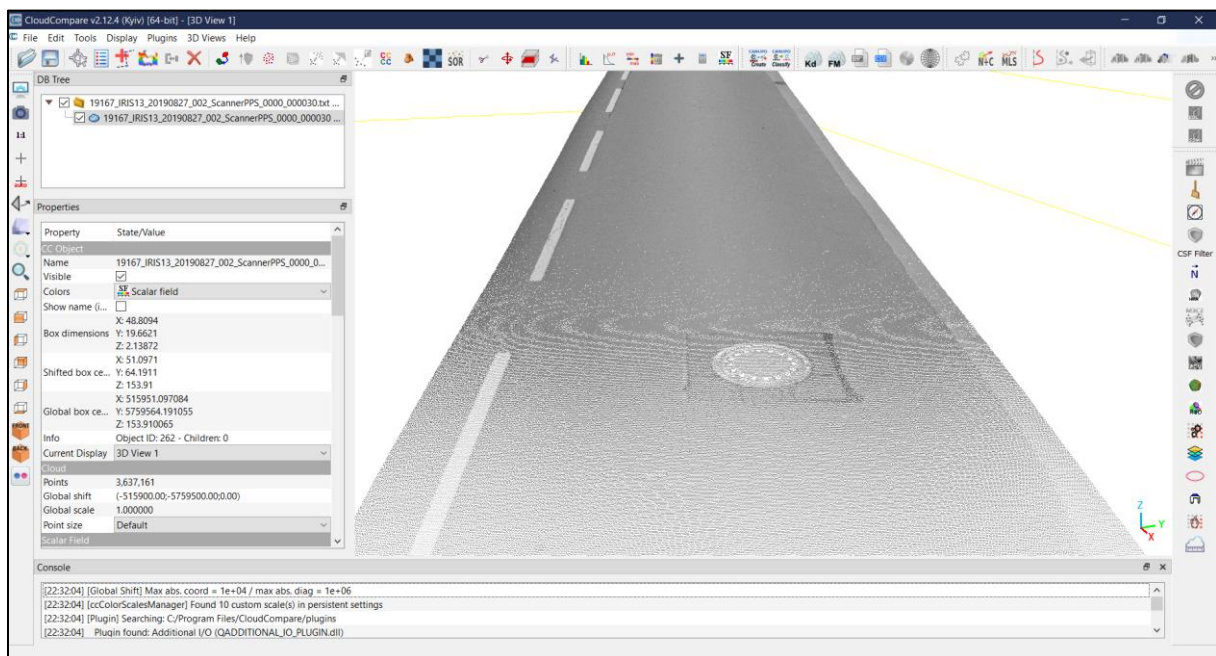


Abbildung 108: User Interface der Software CloudCompare.

4.1.3 Geoprozessierungen

SAGA GIS

Für Aufgaben aus dem Bereich Geoprozessierung und räumliche Interpolation wurde das freie Open-Source Geo-Informationssystem SAGA (System für Automatisierte Geowissenschaftliche Analysen) verwendet (CONRAD et al. 2015). SAGA ist objektorientiert in C++ programmiert und bietet eine effektive Anwendungsprogrammierschnittstelle (API). Alle Funktionen sind als Module in Framework-unabhängigen Modulbibliotheken organisiert und können sowohl über die grafische Benutzeroberfläche (siehe Abbildung 109) als auch aus

verschiedenen Skriptumgebungen, wie beispielsweise Shell, Python, R und Java, aufgerufen werden. Daher eignet sich SAGA besonders zur automatisierten Verarbeitung von Geodaten und zur geowissenschaftlichen Prozessmodellierung (BÖHNER & CONRAD 2011). Die ersten SAGA Module, die im Jahr 2001 am Geographischen Institut der Universität Göttingen entwickelt wurden, waren für die Arbeit mit Rasterdaten konzipiert. Mit seinem Hintergrund in der physischen Geographie liegt ein wesentlicher Schwerpunkt der GIS-Funktionalität von SAGA auf der digitalen Geländeanalyse und -visualisierung (STEINIGER & BOCHER 2009). Heute bietet die Software über 70 Modul-Bibliotheken mit insgesamt über 700 Geoprozessierungs-Tools und Workflows. Der Quellcode ist über SourceForge veröffentlicht (SAGA GIS 2022). Mit Ausnahme der API sind die meisten SAGA-Quellcodes unter der GNU General Public Licence lizenziert. Die API und die darauf aufbauenden SAGA-Module hingegen sind unter der weniger restriktiven GNU Lesser General Public Licence (LGPL) lizenziert.

Für die Untersuchungen dieser Arbeit wurde SAGA GIS in der Version 7.2.0 als Kommandozeilen-Tool für die Oberflächenmodellierung auf Basis von 3D-Punktwolken und Geländemodellanalysen genutzt.

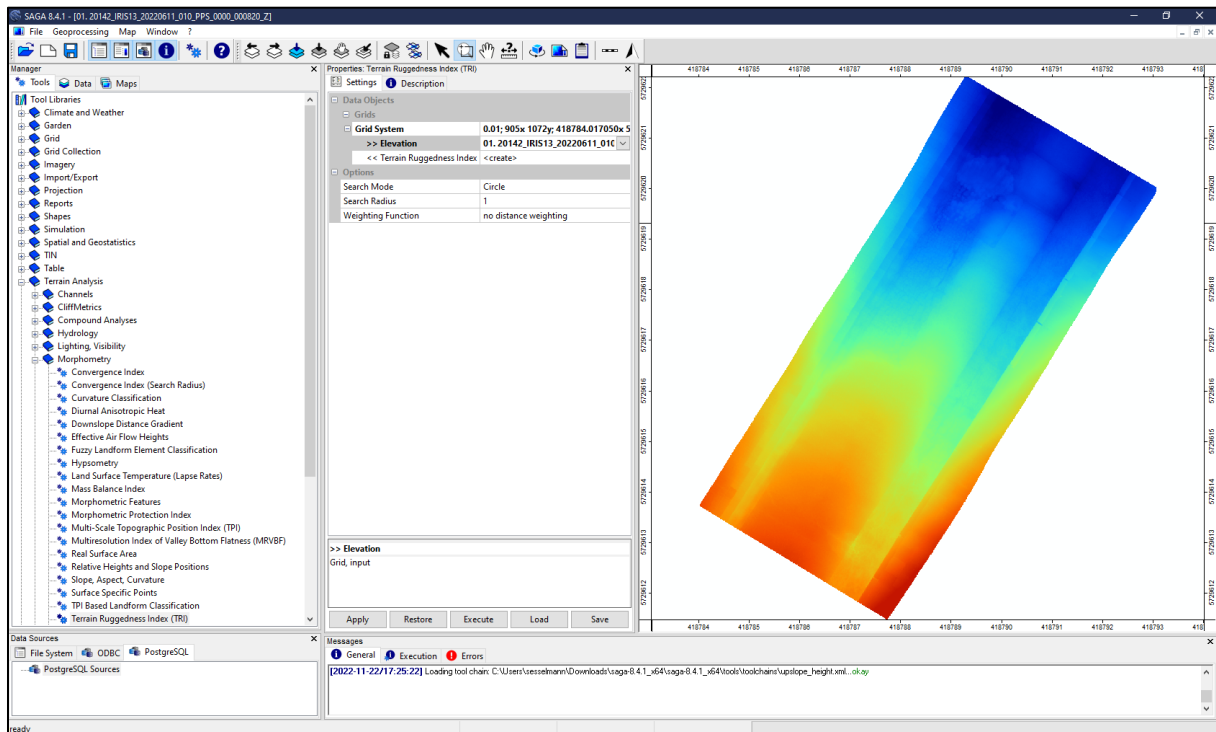


Abbildung 109: User Interface der Software SAGA GIS.

QGIS

QGIS ist eine freie und quelloffene, plattformübergreifende GIS-Anwendung, die das Erstellen, Bearbeiten, Analysieren und Anzeigen von Geoinformationen unterstützt (QGIS.ORG 2022). Die Entwicklung von QGIS begann im Jahr 2002 als Reaktion auf das expertenorientierte Design der GIS Software GRASS (STEINIGER & BOCHER 2009). So fungierte QGIS (damals: Quantum GIS) zunächst als User Interface für GRASS Werkzeuge. Heute kann QGIS auf allen gängigen Betriebssystemen, wie Windows, Mac OS, Linux, BSD und Android als Desktop-, Server- oder Browserlösung genutzt werden und ist unter GNU-GPL lizenziert (siehe Abbildung 110). Für Entwickler ist vor allem die Möglichkeit interessant, QGIS als Software-Bibliothek einzubinden. QGIS ist in C++ programmiert und modular aufgebaut. Zusätzliche Module können in C++ und Python programmiert und als Erweiterung über die Hauptsoftware sehr einfach installiert und verwaltet werden. Es stehen über 1700 Erweiterungen im offiziellen QGIS-Erweiterungsrepositorium zur Verfügung. Heute arbeitet eine Community von über 700 Entwicklern weltweit an der Programmierung von QGIS (ROSAS-CHAVOYA et al. 2022). Es ist unter den freien und quelloffenen Anwendungen das GIS mit einer sehr großen Anwender-Community.

Im Rahmen der Untersuchungen dieser Arbeit wurde QGIS zur visuellen Inspektion von Oberflächenmodellen in Verbindung mit der Erweiterung „Profile Tool“ (JURGIEL et al. 2022) sowie zur Darstellung und räumlichen Analyse der georeferenzierten Segmentierungsergebnisse der Deep-Learning-Modelle verwendet.

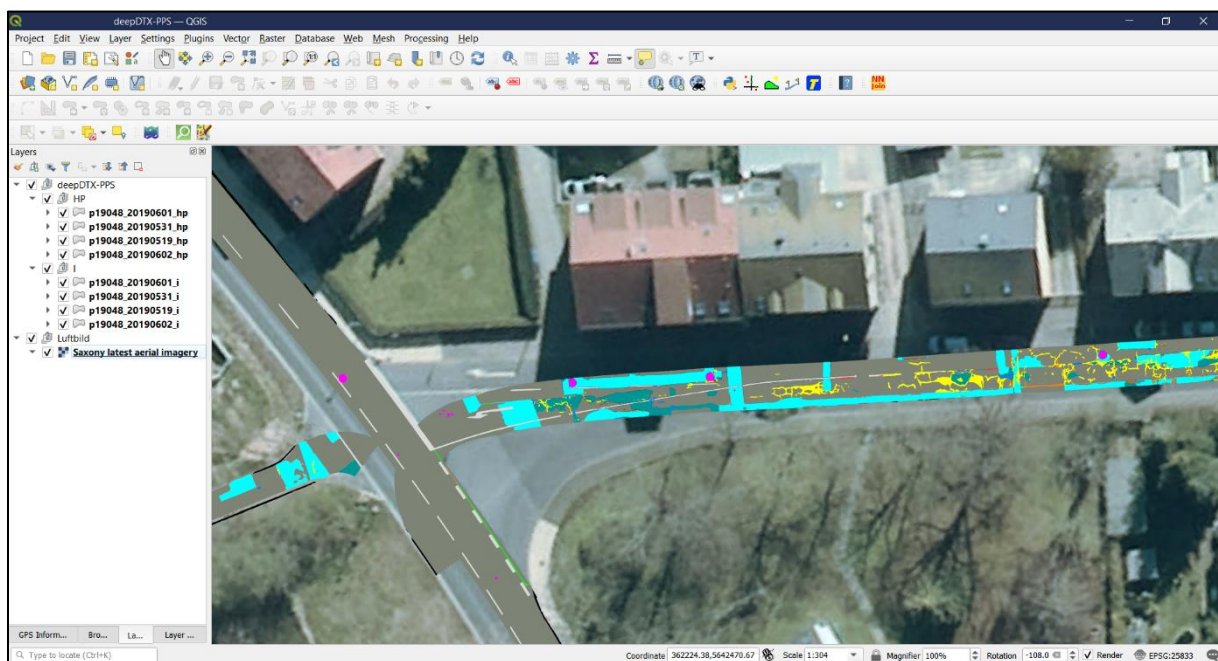


Abbildung 110: User Interface der Software QGIS mit Ergebnis-Polygonen der Deep Learning basierten Analyse von PPS-Oberflächenmodellen.

GeoPackage

GeoPackage ist ein vom Open Geospatial Consortium (OGC) entwickelter Encoding-Standard für Geodaten (siehe Abbildung 111). Bei GeoPackage handelt es sich um ein offenes, standardbasiertes, plattformunabhängiges, portables, selbstbeschreibendes und kompaktes Format für den Austausch und die gemeinsame Nutzung von Geodaten (FREEMAN et al. 2019). Der GeoPackage-Standard beschreibt eine Reihe von Festlegungen für die Speicherung von Elementen, wie Vektordaten, Kachelarchiven (engl.: Tile Matrix Sets) und sonstigen Attributen, in einer dateibasierten Datenbank. Die relationale Datenbank SQLite ermöglicht dabei schnelle Lese- und Schreibvorgänge. Alle gängigen Geometrien, darunter Punkte, Line-Strings, Polygone, MultiPoints, MultiLineStrings, MultiPolygons und GeomCollections, werden unterstützt. Attribute können als Text, reelle Zahlen, Boolesche Zahlen, aber auch Raster/Fotos gespeichert werden. Ein Vorteil von GeoPackage ist, dass mehrere Datensätze, Datentypen und Metadatensätze in derselben Datei gespeichert werden können.

Im Rahmen der Untersuchungen wurde GeoPackage als Ausgabeformat für die geo-referenzierten Segmentierungsergebnisse der Deep-Learning-Modelle gewählt, da es einerseits eine wesentlich schnellere Visualisierung und Interaktion in GIS-Umgebungen ermöglicht, als beispielsweise das Shapefile- oder GeoJSON-Format und andererseits dateibasiert auf allen gängigen Betriebssystemen arbeitet, was ein Höchstmaß an Flexibilität im Hinblick auf den Datenaustausch bietet.

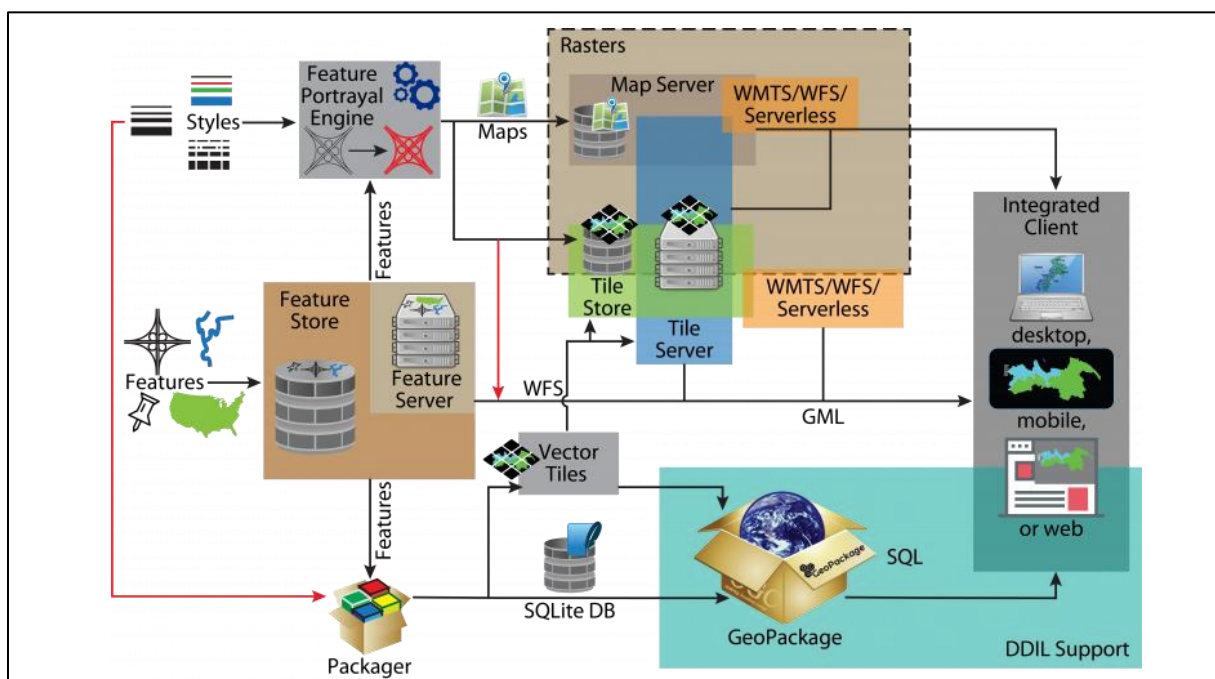


Abbildung 111: GeoPackage als Teil einer integrierten Geoinformationsarchitektur (YUTZLER 2020).

4.1.4 Training und Evaluierung von Deep-Learning-Modellen

ScaLabel

Das Training von Deep-Learning-Modellen zur Bildsegmentierung erfordert das Vorhandensein von semantisch annotierten Bilddaten. Um Bilder semantisch zu annotieren, kann auf eine Vielzahl von Softwarelösungen zurückgegriffen werden, da eine semantische Segmentierung prinzipiell mit jedem Zeichenprogramm durchgeführt werden kann. Professionelle Labeling Tools heben sich dagegen dadurch ab, dass sie beispielsweise die Administration der Labeldaten erleichtern oder semi-automatische Zeichen-Funktionen anbieten. Für das Labeling im Rahmen dieser Arbeit wurde das Tool ScaLabel verwendet. ScaLabel ist eine vielseitige und skalierbare Labeling-Plattform, die sowohl 2D- als auch 3D-Daten unterstützt. Der Quellcode des Open-Source-Webtools ist auf Github verfügbar (SCALABEL 2022). Das Tool ist so konzipiert, dass die Anwendung auf einem Server ausgeführt wird, der auch die Daten hostet. Eine lokale Datenhaltung ist aber auch möglich, wenn Server und Client auf demselben Rechner laufen. Das User Interface ist browserbasiert und bietet eine Vielzahl von Zeichen- und Editierfunktionen. Die Label-Ergebnisse können in Form von JSON-Dateien konform den Datenformaten bekannter Benchmarkdatensätze, wie z. B. Kitty, exportiert werden. Abbildung 112 zeigt das User Interface von ScaLabel auf einem Linux Mint 17 Betriebssystem.

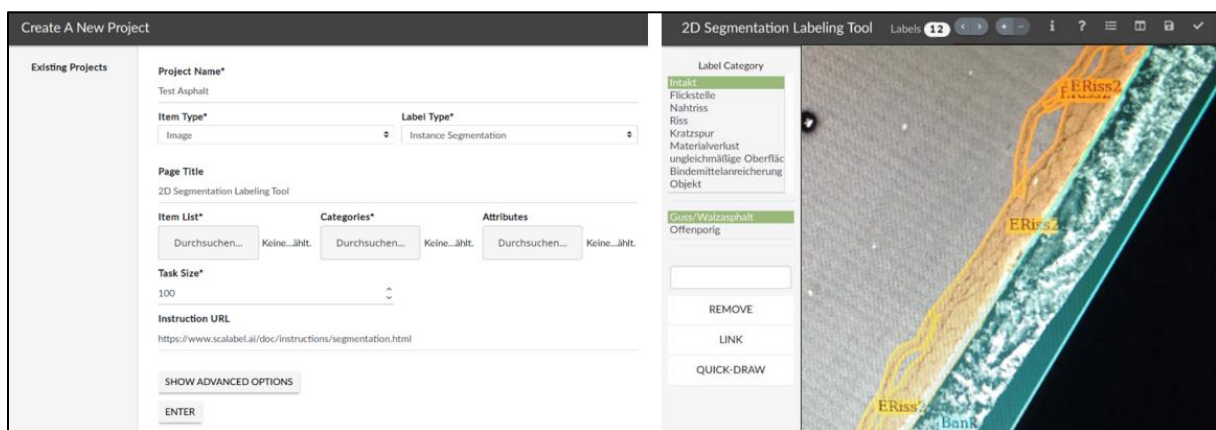


Abbildung 112: User Interface der Software ScaLabel. Links ist der Dialog zum Anlegen eines Label-Projektes dargestellt. Rechts ist das Hauptfenster während des eigentlichen Labelings zu sehen. Die Auswahl der zu labelnden Klasse wird über eine Liste realisiert und dann können ebenenweise Polygone auf das zugrundeliegende Bild gezeichnet werden.

Jupyter Notebook

Für die Erzeugung, Dokumentation und zum Ausführen von Python Codes und Skripten wurden Jupyter Notebooks verwendet. Dabei handelt es sich um eine Open-Source-Webanwendung, mit der im Allgemeinen Dokumente erstellt und geteilt werden können. Entwickelt

wird Jupyter Notebook von dem gemeinnützigen Project Jupyter, das 2014 aus dem IPython Projekt hervorgegangen ist. Der Quellcode ist auf Github unter einer modifizierten BSD-Lizenz veröffentlicht (PROJECT JUPYTER 2022).

Abbildung 113 zeigt die Oberfläche eines Jupyter Notebooks. Ein Notebook wird als separater Reiter im Browser geöffnet und kann als ipynb-Datei gespeichert werden. Die Grundstruktur eines Jupyter Notebooks bilden die sogenannten Zellen, von denen beliebig viele untereinander eingefügt werden. Jede Zelle kann erklärenden Text, Gleichungen, Grafiken oder ausführbaren Code beinhalten. Oben rechts in Abbildung 113 wird der zu verwendende Kernel eingestellt. In diesem Fall wird ein Python 3 Kernel verwendet. So kann der Python-Code, der in einer Zelle steht, korrekt interpretiert und ausgeführt werden. Die Ergebnisse werden dann direkt unter der betreffenden Zelle angezeigt. Dies können einfache Konsolenausgaben, aber auch komplexe Diagramme oder interaktive Nutzeroberflächen sein. Im Beispiel in Abbildung 113 sorgt der Code in der dargestellten Zelle dafür, dass ein Pythonskript mit bestimmten Parametern ausgeführt wird.

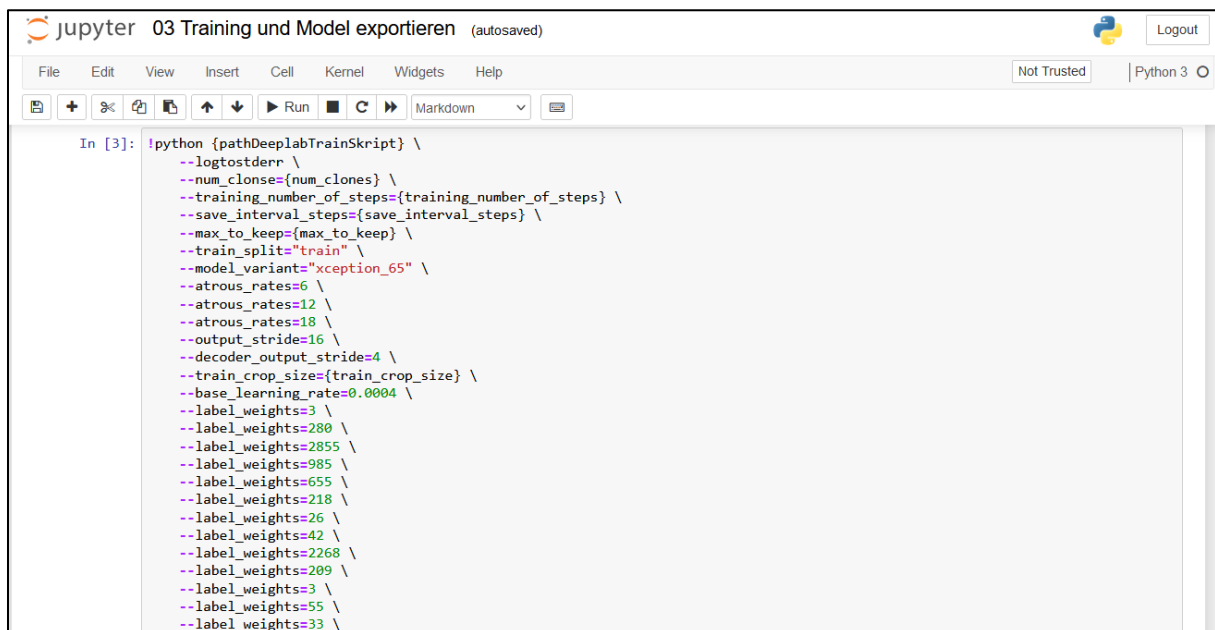


Abbildung 113: Beispiel für ein Jupyter Notebook. Die ausschnittsweise dargestellte Zelle startet das Skript zum Training eines DeepLabv3+ Modells mit diversen Parametern.

TensorFlow

TensorFlow ist eine Open-Source-Plattform für Anwendungen des maschinellen Lernens, deren Quellcode auf Github veröffentlicht ist (TENSORFLOW 2022). Als Framework bietet TensorFlow umfassende Werkzeuge, Bibliotheken und Community-Ressourcen und ist daher in wissenschaftlichen und kommerziellen Anwendungen weit verbreitet. Mit TensorFlow kön-

nen Machine Learning gestützte Anwendungen erstellt, angewendet und vertrieben werden. Ursprünglich wurde das Framework von Forschern und Ingenieuren des Google Brain Teams entwickelt, um maschinelles Lernen und tiefe neuronale Netzwerke zu erforschen. Heute ist der Quellcode unter Apache-2.0 lizenziert. Die Funktionen von TensorFlow können sowohl über Python als auch über C++ verwendet werden. Der Schwerpunkt liegt auf dem Training und der Anwendung (Inferenz) tiefer künstlicher neuronaler Netze. Die funktionalen Grundlagen des Frameworks sind bei ABADI et al. (2016) beschrieben. TensorFlow verwendet ein besonderes Datenflussmodell (Dataflow Graphs), um Berechnungen, Zustände und zustandsändernde Operationen über mehrere Recheneinheiten innerhalb eines Computers (CPU's, GPU's, TPU's) oder mehrere Computer in einem Cluster effizient darzustellen. Eine TensorFlow-Berechnung kann durch einen gerichteten Graphen beschrieben werden, der aus einer Menge von Knoten zusammengesetzt ist, von denen einige ihre Zustände beibehalten oder aktualisieren können. Auch ermöglichen manche Knoten Verzweigungs- und Schleifenkontrollstrukturen innerhalb des Graphen. Jeder Knoten kann keinen oder mehrere In- bzw. Outputs haben und repräsentiert die Instanz einer Operation, wie z. B. einer Matrixmultiplikation oder einen ReLU-Block. Als Tensoren bezeichnet man die Werte, die entlang der Knoten-Kanten durch den Graphen fließen. Sie sind Arrays beliebiger Dimensionalität. Das Erstellen und Anwenden von Berechnungsgraphen ist in TensorFlow als sogenannte Sitzungen (Sessions) realisiert: Eine Session wird immer mit einem leeren Graphen initialisiert. Der Graph kann dann um Knoten und Kanten erweitert und schließlich mit einer run-Methode ausgeführt werden. In der Regel wird ein Graph innerhalb einer Session einmal eingerichtet und dann der gesamte Graph bzw. Teile des Graphen millionenfach ausgeführt. TensorFlow bietet mit TensorBoard ein hilfreiches Visualisierungswerkzeug (siehe Abbildung 114). Es unterstützt durch geeignete Visualisierungen der abstrakten Prozesse innerhalb eines Graphen beim Debugging und der Optimierung von Deep-Learning-Experimenten und ermöglicht, das Verhalten von ML-Modellen besser zu interpretieren. Beispielsweise kann der Verlauf von Metriken, wie z. B. Loss und Accuracy, über den Trainingszyklus hinweg verfolgt werden, was hilft, potentielles Overfitting zu identifizieren. Auch lassen sich Modellarchitekturen grafisch darstellen. Fehler können dadurch gerade bei komplexeren Architekturen schneller auffallen. Im Rahmen der vorliegenden Arbeit wird TensorFlow zum Training und zur Inferenz von Modellen verwendet.

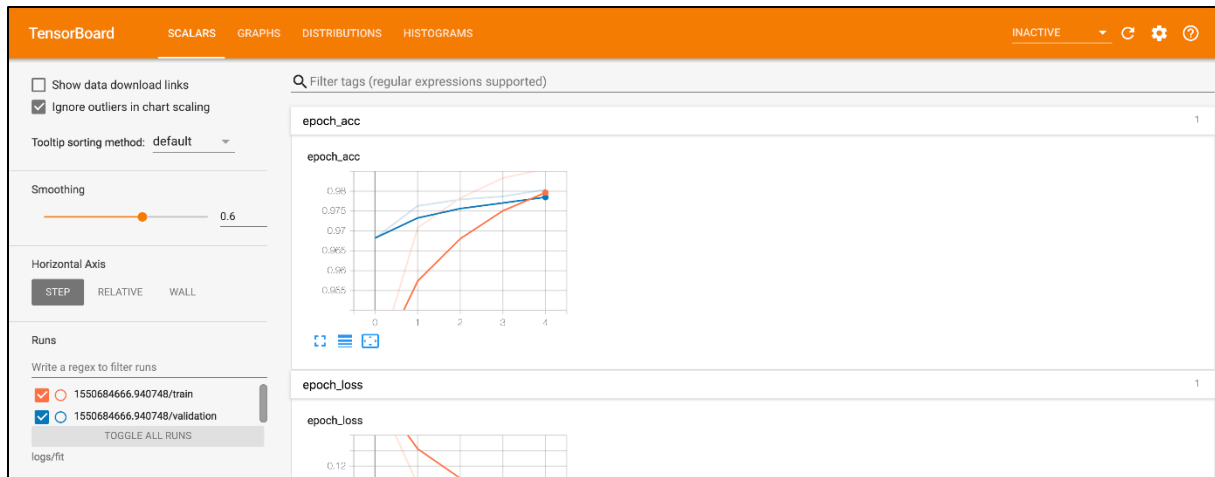


Abbildung 114: User Interface des Visualisierungswerkzeugs TensorBoard (TENSORBOARD 2022).

4.1.5 Erzeugung regelwerkskonformer TP3-Daten

Für die Erfassung von Oberflächenschäden gemäß Teilprojekt TP3 des Regelwerks ZTV ZEB-Stb (FGSV 2006) existiert weder eine proprietäre Standardsoftware noch eine freie und quelloffene Lösung. Jeder Dienstleister nutzt eine auf das jeweilige Erfassungssystem abgestimmte interne Lösung. Folgende wesentliche Funktionen müssen umgesetzt sein, um eine anforderungsgerechte und effiziente Auswertung im Sinne des Regelwerkes durchzuführen: Es muss eine synchronisierte Visualisierung von Oberflächenbildern und Front-Kamerabildern erfolgen. Auf Basis der Oberflächenbilder muss ein interaktives Auswerteraster dargestellt werden. Die Interaktionsmöglichkeiten umfassen einerseits, das Verändern der Lage des Rasters entsprechend des Bildinhalts (Anpassen an Fahrstreifengrenzen) und andererseits das Editieren des Auswerterasters im Hinblick auf Oberflächenschäden (Markieren der von Schäden betroffenen Rasterzellen). Schließlich muss die Auswertesoftware über eine Schnittstelle verfügen, um einen Ergebnisexport gemäß des Austauschformats zu gewährleisten. Abbildung 115 zeigt einen typischen Auswertearbeitsplatz.

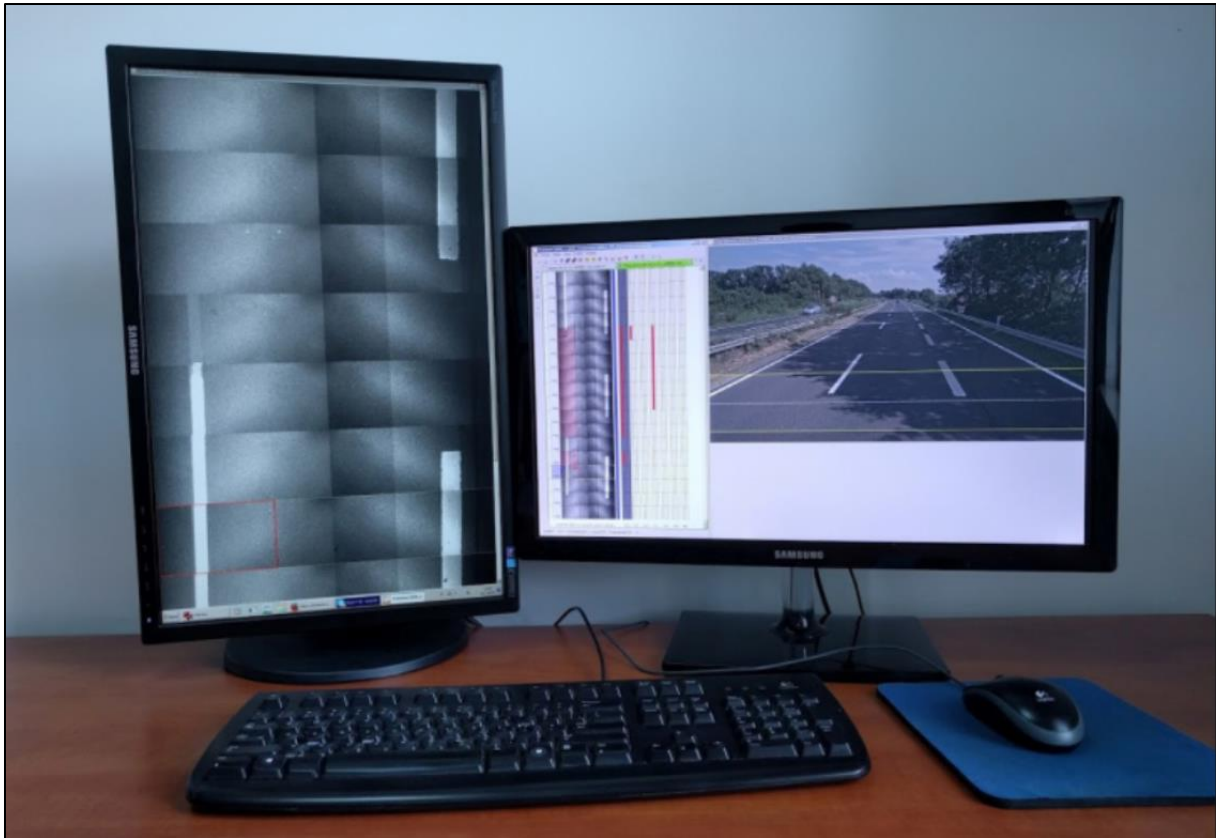


Abbildung 115: Typischer Auswertearbeitsplatz mit der zugehörigen Software der Firma Heller Ingenieurgesellschaft mbH (SKAKUJ et al. 2019).

In Abbildung 116 ist die Softwarelösung der Firma Lehmann + Partner GmbH abgebildet. Bei dem verwendeten Tool handelt es sich um eine Browser-Anwendung, die Bilddaten von einem zentralen Server lädt, ein Auswerteraster über die Bilddaten legt, eine manuelle Eingabe von Schäden in das Auswerteraster erlaubt und die Ergebnisse in einer PostgreSQL-Datenbank speichert. Von dort aus können sie gemäß der Formatanforderung als sogenannte XML-Rohdaten exportiert werden (HELLER INGENIEURGESELLSCHAFT MBH 2018). Diese proprietäre Software wird von der Firma Lehmann + Partner GmbH ausschließlich firmenintern zur Datenbearbeitung im Rahmen von ZEB-Kampagnen im Bundesfernstraßennetz eingesetzt. Im Rahmen der Untersuchungen wurde die Software verwendet, um eine regelwerkskonforme Referenzauswertung für die Evaluierungsstrecke zu erzeugen und die Leistungsfähigkeit der Deep Learning basierten Auswertung von PPS-Oberflächenmodellen in Bezug auf die Anforderungen der Zielanwendung zu evaluieren.

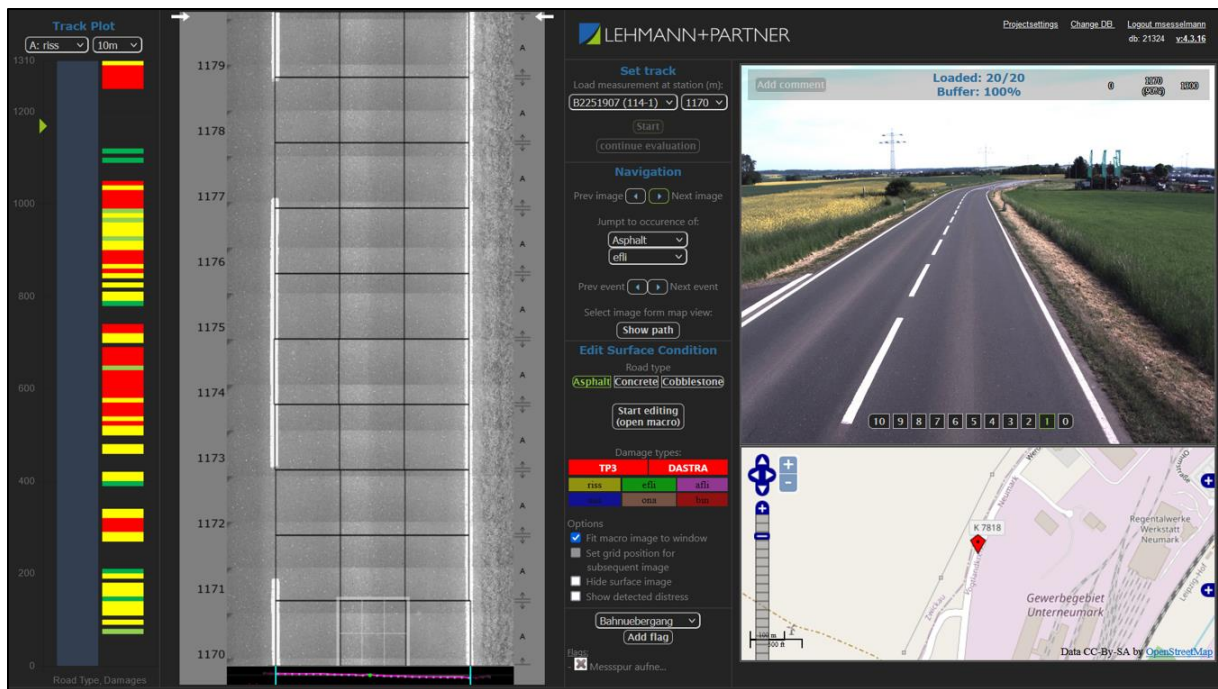


Abbildung 116: User Interface der Software TP3-Websolution der Firma Lehmann + Partner GmbH zur Erfassung von Oberflächenschäden gemäß ZTV ZEB-Stb (FGSV 2006).

4.2 Deterministische räumliche Interpolationsverfahren

In Kapitel 2.3.1.2 wurden räumliche Interpolationsverfahren thematisch eingeführt. Seit der zweiten Hälfte des 20. Jahrhunderts sind eine Vielzahl von räumlichen Interpolationsmethoden entwickelt worden. In ihrer Studie vergleichen LI & HEAP (2011) mehr als 30 Verfahren. Allein diese Fülle von möglichen Interpolatoren lässt erahnen, dass es keine universelle Interpolationsmethode geben kann, die für alle Fragestellungen optimale Ergebnisse liefert. Es stellt sich daher bei der Wahl eines Interpolators die Frage danach, „welches Verfahren dem betreffenden Forschungsgegenstand, den Kenntnissen über die Wirkungszusammenhänge und den Grunddaten angemessen ist“ (RASE 2021, S. 57).

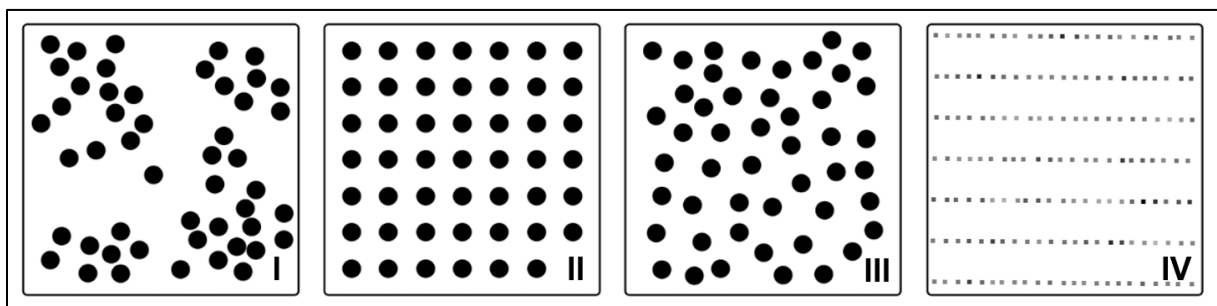


Abbildung 117: Räumliche Verteilungsmuster von Messwerten in der XY-Ebene (nach KEMPF 2020). Gruppierte (I), gleichmäßige (II) und zufällige Messpunktverteilung (III) im Vergleich zur Verteilung im PPS-Laserscan (IV).

Eine wesentliche Rolle für die Wahl des Interpolators und die Qualität des Interpolationsergebnisses spielt die Dichte und räumliche Verteilung der verfügbaren Beobachtungen (RASE 2021). Grundsätzlich werden drei Hauptgruppen von Verteilungsmustern unterschieden: gleichmäßig, zufällig und gruppiert (siehe Abbildung 117). Die in der vorliegenden Arbeit verwendeten Laserscans befinden sich in diesem Spektrum im Übergangsbereich zwischen den Kategorien gruppiert und gleichmäßig, da es zwar entlang der einzelnen Scanprofile zu einer lokalen Punkthäufung kommt, diese aber in regelmäßigen, von der Befahrungsgeschwindigkeit abhängigen, Abständen aufeinander folgen. Der zu interpolierende Raum zwischen den Messpunkten ist in Relation zur hohen Punktdichte und zur Auflösung des Zielrasters vergleichsweise gering. Das Datenmaterial stellt also im Vergleich zu anderen Anwendungsdomänen, wo keine derartige Messpunktdichte verfügbar ist, keine erhöhten Anforderungen an die prädiktive Leistungsfähigkeit der Interpolationsmethode. Auch LI & HEAP (2014) weisen darauf hin, dass keine großen Unterschiede zwischen den Ergebnissen unterschiedlicher Interpolatoren zu erwarten sind, wenn die Dichte der Messdaten hoch ist. Ist sie dagegen niedrig, gewinnt die Wahl eines geeigneten Interpolationsverfahrens an Bedeutung. Im Allgemeinen eignen sich bei gleichmäßig verteilten und dichten Daten Spline-Interpolatoren oder Verfahren mit Inverser Distanzgewichtung (WATSON & PHILIP 1985; LI & HEAP 2014). Bei unregelmäßigen und stark gruppierten Verteilungen sind beide Verfahrenstypen in ihrer Leistungsfähigkeit eingeschränkt. Eine starke Gruppiertheit (engl.: Clustering) beeinflusst die Ergebnisqualität sowohl von deterministischen als auch geostatistischen Verfahren negativ. Für Messwerte, die unregelmäßig verteilt sind, liefert die Ordinary-Kriging-Interpolation brauchbare Ergebnisse, sofern zwischen den Daten-Clustern keine gravierenden Unterschiede bezüglich der Datendichte herrschen (LI & HEAP 2014).

Aus der Gruppe der deterministischen Interpolationsverfahren werden im Rahmen dieser Arbeit fünf Verfahren ausgewählt, erläutert und ihre Leistungsfähigkeit für die gegebene Problemstellung vergleichend analysiert, die in verschiedenen Publikationen mit unterschiedlichen geowissenschaftlichen Ausrichtungen bereits vielversprechende Ergebnisse lieferten (BATER & COOPS 2009; DUMITRU, PLOPEANU & BADEA 2013; GARNERO & GODONE 2013; MONTEALEGRE, LAMELAS & RIVA 2015; SZYPUŁA 2016; BHUNIA, SHIT & MAITI 2018; CĂTEANU & CIUBOTARU 2020). So werden beispielsweise bei CĂTEANU & CIUBOTARU (2020) acht deterministische und ein geostatistisches Verfahren verglichen, um deren Potential im Rahmen der Bodenoberflächeninterpolation unter dichtem Waldbestand aus Airborne-LiDAR-Daten zu untersuchen. Die brauchbarsten Resultate im Hinblick auf die vertikale Genauigkeit lieferten hierbei die Thin Plate Spline, die Natural Neighbors und die Delaunay-Triangulation.

Die Inverse-Distance-Weighted-Methode liegt hinsichtlich ihrer Interpolationsgüte im Mittelfeld der untersuchten Algorithmen. Auch wenn sich in den genannten Studien in Abhängigkeit zur jeweiligen Anwendungsdomäne und dem verwendeten Datenmaterial teils unterschiedliche Verfahren als am besten geeignet erwiesen, konnte eine Auswahl von potentiell geeigneten Methoden getroffen werden, die auf ihre Leistungsfähigkeit im Rahmen der Problemstellung dieser Arbeit hin untersucht werden. Hierfür wurden nicht nur exakte, sondern auch approximierende Verfahren betrachtet (siehe Tabelle 11). Die Verwendung von approximierenden Interpolatoren ist dann probat, wenn davon ausgegangen werden kann, dass die Messpunktdichte groß genug ist, um das Ausmaß der für die Analyse erforderlichen lokalen Oberflächenvariationen zu erfassen (SZYPUŁA 2016). In den nachfolgenden Unterkapiteln werden daher die Delaunay-Triangulation mit linearer Wertinterpolation, die Natural Neighbor Interpolation, die Inverse Distance Weighted Interpolation, die Modified Quadric Sheppard Interpolation und die Thin Plate Spline Interpolation erläutert.

Tabelle 11: Überblick über die untersuchten Interpolationsverfahren und ihre Eigenschaften (nach DUMITRU, PLOPEANU & BADEA 2013; LI & HEAP 2014).

Methoden	Typ	Eigenschaften
Triangulation mit linearer Interpolation	exakt	▪ basiert auf Delaunay-Triangulation ▪ erfordert mittlere bis große Anzahl von Daten für akzeptable Ergebnisse
Natural Neighbor	exakt	▪ basiert auf Gewichtung benachbarter Datenpunkte nach dem Konzept der Voronoi-Polygone ▪ eignet sich für große Datensätze ▪ extrapoliert nicht
Inverse Distance Weighted	exakt	▪ intuitives Funktionsprinzip (räumliche Autokorrelation) ▪ einfache Berechnung ▪ alle Werte des Zielrasters werden aus den Datenpunkten interpoliert ▪ Kontrolle des Glättungsgrades über Exponent ▪ extrapoliert nicht
Modified Quadratic Shepard	approximierend (exakt nur, wenn Glättungsfaktor implementiert)	▪ Kleinste-Quadrate-Schätzung einer Quadrik um jeden Datenpunkt ▪ Teil des Verfahrens nutzt inverse Abstandsgewichtung ▪ Extrapoliert
Thin Plate Spline	approximierend (exakt nur, wenn Regularisierungswert $\lambda = 0$)	▪ geeignet für schnelle Interpolation aus mäßig detaillierten Daten ▪ minimale gekrümmte Oberfläche durch die Datenpunkte ▪ abrupte Oberflächenänderung an Datenpunkten möglich ▪ je höher der Regularisierungswert, desto glatter wird die Oberfläche

4.2.1 Delaunay-Triangulation

Im Grundlagenkapitel 2.3.1.1 wurde ausgeführt, dass Oberflächen verschiedenartig repräsentiert werden können, beispielsweise durch Konturlinien, regelmäßige Rasterzellen oder aber auch als Netzwerk unregelmäßiger Dreiecksflächen (TSAI 1993). Eine Möglichkeit, ein Netz aus unregelmäßigen, zusammenhängenden und sich nicht überlappenden Dreiecksflächen auf Basis von Punktdaten zu erzeugen, ist die Delaunay-Triangulation. Sie kann folgendermaßen definiert werden (LARA-ALVAREZ, ROJAS & BAYRO-CORROCHANO 2014): Unter der Annahme, dass eine Menge bekannter Datenpunkte P , mit einer Punktzahl ≥ 3 in der euklidischen Ebene gegeben ist und dabei weder drei Punkte kollinear noch vier Punkte kozirkular sind, beschreibt die Delaunay-Methode eine Triangulation, bei der im Zuge der Dreieckskonstruktion kein Punkt P innerhalb eines umschließenden Kreises des erzeugten Dreiecks liegt. Mit diesem Ansatz maximiert ein Delaunay-Diagramm den kleinsten Winkel aller Dreiecke und ermöglicht somit die Konstruktion von Dreiecken mit relativ kurzen Kanten. In Abbildung 118 ist dargestellt, wie die Datenpunkte so zu Dreiecksflächen vernetzt werden, dass sich im Umkreis eines jeden Dreiecks keine weiteren Datenpunkte mehr befinden.

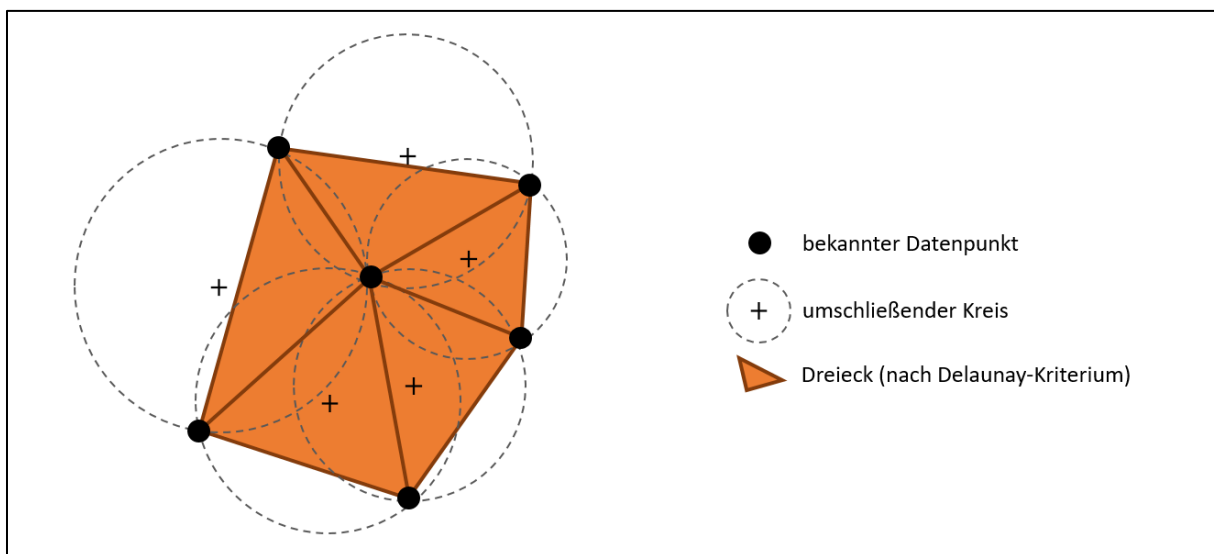


Abbildung 118: Prinzip der Delaunay-Triangulation. Die bekannten Datenpunkte werden so zu Dreiecksflächen verbunden, dass im umschließenden Kreis des Dreiecks keine weiteren Datenpunkte enthalten sind (nach QUINTERO, ESCOBAR & PONTE-NETO 2012).

Um das erzeugte unregelmäßige Dreiecksnetz auf das regelmäßige Gitter eines gerasterten Oberflächenmodells zu übertragen, werden zunächst die Knoten- oder Pixelmittelpunkte in das Delaunay-Diagramm eingefügt. Anschließend werden deren Werte durch lineare oder kubische Polynominterpolation unter Verwendung der betreffenden Dreieckseckpunkte berechnet.

(LI & HEAP 2008; SAMBRIDGE, BRAUN & MCQUEEN 1995). In Abbildung 119 ist die lineare Interpolation schematisch dargestellt.

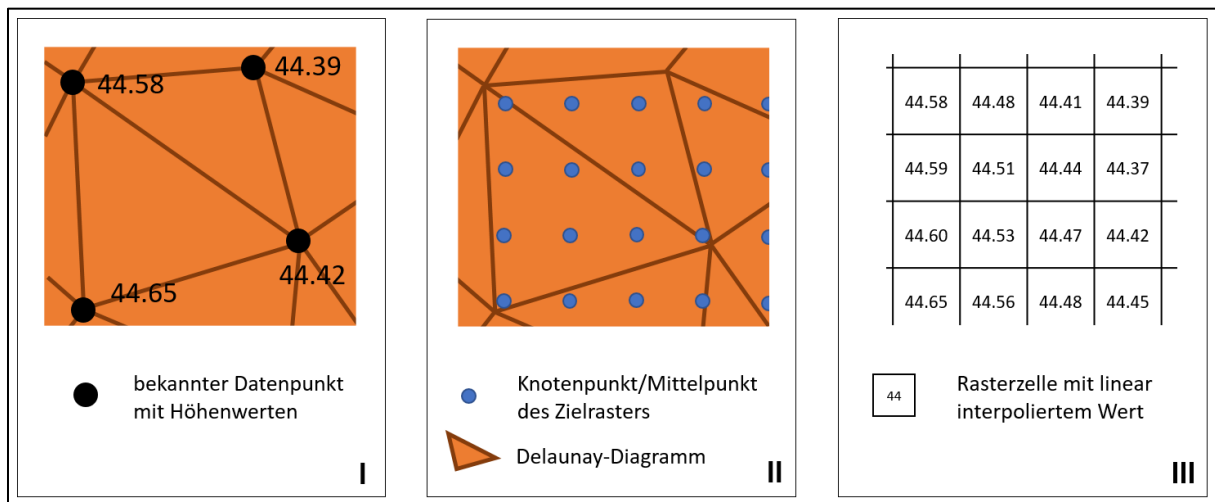


Abbildung 119: Prinzip der Erzeugung eines regelmäßigen Gitteroberflächenmodells aus der unregelmäßigen Delaunay-Dreiecksvermaschung. In das Dreiecksnetz (I) werden die Knotenpunkte des Zielrasters eingefügt (II) und die Wertzuweisung erfolgt über lineare Interpolation (III).

SAMBRIDGE, BRAUN & MCQUEEN (1995) beschreiben den Fall der linearen Interpolation wie folgt: Unter der Annahme bekannter Werte f_i an den gegebenen Knotenpunkten eines Dreiecks (x_i, y_i) ist der interpolierte Wert an einem beliebigen Punkt (x, y) im Inneren des Dreiecks gegeben durch die nachstehende Formel. Dabei steht $\phi_i(x, y)$ für eine zweidimensionale Basisfunktion, die linear zwischen dem Wert 1 am Knoten (x_i, y_i) und 0 am Knoten $(x_j, y_j) (j \neq i)$ variiert.

$$f(x, y) = \sum_{i=1}^3 \phi_i(x, y) f_i \quad (4)$$

Da die Delaunay-Triangulation die nächstgelegenen Nachbarn in einer Punktnachbarschaft miteinander verbindet, wird die lineare Interpolation auf Basis von Delaunay-triangulierten Datenpunkten teilweise auch als lineare Natural Neighbors Interpolation bezeichnet (QUINTERO, ESCOBAR & PONTE-NETO 2012).

Zu den Eigenschaften der linearen Interpolation auf Basis der Delaunay-Triangulation zählt, dass die ursprünglichen Werte an den Datenpunkten exakt wiederhergestellt werden. Die Interpolation arbeitet vollständig lokal: jeder Punkt wird ausschließlich von seinen natürlichen Nachbarknoten beeinflusst. Die Ableitungen der interpolierten Oberfläche sind überall außer an den Referenzpunkten kontinuierlich (SAMBRIDGE, BRAUN & MCQUEEN 1995).

4.2.2 Natural Neighbor

Die Natural-Neighbor-Methode (NN) geht auf die Arbeiten von SIBSON (1981) zurück. Sie kombiniert Aspekte der Nearest-Neighbor-Methode und der Dreiecksvermaschung nach Delaunay (LI & HEAP 2008, 2014). In einem ersten Schritt findet eine Delaunay-Triangulation der Daten zu einem unregelmäßigen Dreiecksnetz (TIN) statt. Auf Basis dieser Vermaschung werden initiale Voronoi-Polygone berechnet, deren Eckpunkte dort entstehen, wo sich die Mittelsenkrechten eines Dreiecks schneiden. Dieser Schnittpunkt kann auch außerhalb der Dreiecksfläche liegen. Ein Voronoi-Polygon definiert somit den Bereich, der einen Datenpunkt von seinem nächsten Nachbarn abgrenzt. Anschließend wird jeder zu interpolierende Punkt in die Vermaschung aufgenommen und das Voronoi-Diagramm sequentiell neu berechnet (TSIDAEV 2016). Anders als bei der Nearest-Neighbor-Methode jedoch, wo der nächstgelegene Nachbarn den Wert eines zu interpolierenden Datenpunkts determiniert, hängt der Wert, der dem interpolierten Punkt zugeordnet wird, davon ab, wie viel von der Fläche der ursprünglichen Diagrammelemente durch das Voronoi-Polygon des neu eingefügten Punkts eingenommen wird. Deshalb wird diese Methode auch als Area-Stealing-Interpolation bezeichnet. Für jeden Nachbarnpunkt des zu interpolierenden Punkts wird das Verhältnis der Überlappungsfläche der Nachbarpunkt-Polygonanteile zur gesamten Voronoi-Polygonfläche des zu interpolierenden Punkts berechnet. Dieses Verhältnis repräsentiert schließlich das Gewicht jedes Nachbarpunkts für die Interpolation des unbekannten Datenpunkts (LI & HEAP 2008).

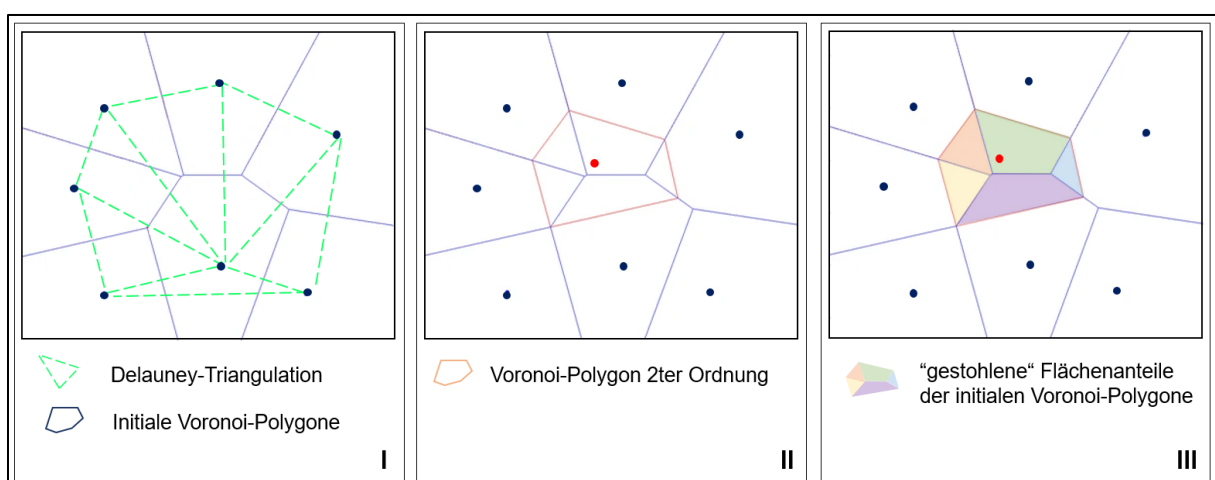


Abbildung 120: Schematische Darstellung des Ablaufs der Sibson Natural Neighbors Interpolationsmethode (nach BEUTEL, MØLHAVE & AGARWAL 2010; KOTULAK et al. 2017). Basierend auf einem Voronoi-Diagramm werden die umliegenden natürlichen Nachbarn identifiziert (I) und unter Einbezug des zu interpolierenden Punkts (rot) ein zweites Voronoi-Diagramm berechnet. Die Verschneidung der beiden Voronoi-Diagramme ergibt unterschiedlich große Flächen (III), die wiederum die Gewichtung der Nachbarn bei der Wertinterpolation bestimmen.

In Abbildung 120 ist das Prinzip der Natural Neighbors Methode schematisch dargestellt. Die blauen Linien stellen das Voronoi-Diagramm dar, welches auf Basis der blauen Eingabedatenpunkte (blau) berechnet wurde. Für den zu interpolierenden Punkt (rot) wird das Voronoi-Diagramm unter Einbeziehung dieses Punkts aktualisiert und das resultierende Voronoi-Polygon um den zu interpolierenden Punkt mit den Flächen des initialen Voronoi-Diagramms verschnitten. Die resultierenden Flächenanteile bestimmen das Gewicht im Zuge der Wertinterpolation für den unbekannten Datenpunkt.

Die nachfolgende Gleichung beschreibt die Berechnung des zu interpolierenden Werts, wobei f_x der Wert für den interpolierten Punkt, M die Anzahl der benachbarten Voronoi-Zellen (natürlichen Nachbarn), f_m der Wert des m -ten Nachbarn ist (KOTULAK et al. 2017).

$$f_x = \sum_{m=1}^M \alpha_m f_m \quad (5)$$

Der Gewichtungskoeffizient α_m der obigen Gleichung wird gemäß der nachstehenden Gleichung berechnet, wobei s_m die Teilfläche der m -ten Zelle und s_0 die gesamte Fläche der neuen Zelle ist.

$$\alpha_m = \frac{s_m}{s_0} \quad (6)$$

Die vorstehend beschriebene Methode wird auch als Sibson-Methode bezeichnet. Daneben existiert ein alternativer Natural-Neighbors-Ansatz. Diese Non-Sibson-Methode basiert ebenfalls auf der Berechnung von Voronoi-Diagrammen. Jedoch werden für die Bestimmung der Interpolationsgewichte keine Flächenanteile genutzt, sondern lineare Größen. Damit adressiert dieser Ansatz beispielsweise das Problem der offenen Voronoi-Polygone, welches bei der Sibson-Methode unter Umständen auftreten kann (KOTULAK et al. 2017). Die Non-Sibson-Methode ist bei BELIKOV et al. (1997) beschrieben. Demnach kann der interpolierte Punkt als lineare Kombination der Nachbarn mithilfe der nachfolgenden Gleichung berechnet werden, bei der f_x der Wert des interpolierten Punkts, M die Anzahl der natürlichen Nachbarn und f_m der Wert des m -ten Nachbarn ist (KOTULAK et al. 2017).

$$f_x = \sum_{m=1}^M \alpha_m f_m \quad (7)$$

Der Gewichtungskoeffizient α_m der obigen Gleichung wird gemäß der nachstehenden Gleichung berechnet, wobei d_m die Länge der Voronoi-Kante ist und h_m für den halben Abstand zwischen dem interpolierten Punkt x und der Voronoi-Kante m steht.

$$\alpha_m = \frac{d_m}{h_m} \quad (8)$$

In Abbildung 121 ist das Prinzip der Non-Sibson-Methode visualisiert. Der Vorteil gegenüber der ursprünglichen Sibson-Methode ist, dass die Berechnung von Gewichten auf Basis von linearen Abständen im Vergleich zur flächenbasierten Gewichtung vereinfacht und dabei effizienter ist (BELIKOV et al. 1997). So erfordert die Non-Sibson-Methode nur eine einmalige Erstellung eines Voronoi-Diagramms.

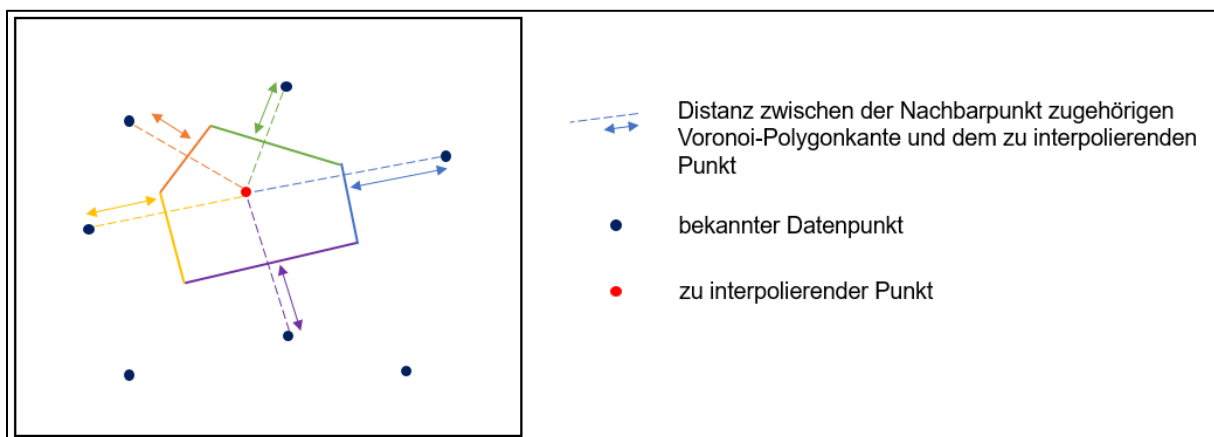


Abbildung 121: Schematische Darstellung des Prinzips der Non-Sibson-Natural-Neighbors-Interpolationsmethode, bei der nur eine einmalige Berechnung eines Voronoi-Diagramms nötig ist (nach KOTULAK et al. 2017).

Im Allgemeinen charakterisieren folgende Eigenschaften die Sibson- und Non-Sibson-Natural-Neighbors-Methode (GARNERO & GODONE 2013): Bei beiden Typen handelt sich um eine lokale Interpolationsmethode, da nur eine Teilmenge von Punkten, die den unbekannten Punkt umgibt, verwendet wird. Ferner ist sie trendunabhängig und sie produziert, sofern nicht bereits in den bekannten Datenpunkten enthalten, keine Spitzen, Löcher, Kämme oder Täler. Die mit der Natural-Neighbors-Methode modellierte Oberfläche verläuft durch die bekannten Datenpunkte und passt sich lokal an die Struktur der Eingabedaten an. Im Unterschied zu anderen räumlichen Interpolationsmethoden sind keine weiteren Parameter nötig, wie beispielsweise die Definition eines Suchradius. Die Natural-Neighbors-Methode kann sowohl mit regelmäßig, als auch unregelmäßig verteilten Daten verwendet werden.

4.2.3 Inverse Distance Weighted

Zu den am weitesten verbreiteten räumlichen Interpolationsmethoden gehört die Inverse-Distance-Weighted-Methode (IDW), weil sie den Einfluss von räumlich verteilten, bekannten Werten auf einen beliebig positionierten, unbekannten Datenpunkt nach dem Konzept der inversen Abstandsgewichtung bewertet (KOTULAK et al. 2017). Dem Algorithmus liegt die Idee der räumlichen Nachbarschaft zugrunde, die besagt, dass der Einfluss von bekannten Punkten auf einen unbekannten Punkt mit zunehmendem Abstand geringer wird. Anders ausgedrückt, beeinflusst ein räumlich nah am zu interpolierenden Punkt gelegener, bekannter Messpunkt den Wert eines zu interpolierenden Punkts stärker, als entferntere Messpunkte (TOBLER 1970). Dieser Einfluss wird im Interpolationsalgorithmus durch Gewichte repräsentiert, die auf Basis der Entfernung bekannter Punkte zum zu interpolierenden Punkt bestimmt werden. Der Wert des zu interpolierenden Punktes berechnet sich aus dem Mittel der umliegenden Punkte, die mit dem Kehrwert der Distanzen gewichtet werden (siehe Abbildung 122). Das Ergebnis der IDW-Interpolation für einen unbekannten Datenpunkt ist also der gewichtete Mittelwert seiner Nachbarn, wobei die Gewichte proportional zu einer Potenzfunktion des zunehmenden Abstands zwischen den Punkten abnehmen (DUMITRU, PLOPEANU & BADEA 2013). Die nachstehende Gleichung beschreibt die IDW-Interpolation, wobei $Z(x_0)$ das Interpolationsergebnis, n die Gesamtzahl der bekannten Datenpunkte, x_i der i -te Datenwert, h_i der Abstand zwischen zu interpolierendem Punkt und bekanntem Datenpunkt 0 sowie β die Gewichtungstärke ist (BHUNIA, SHIT & MAITI 2018).

$$Z(x_0) = \frac{\sum_{i=1}^n \frac{x_i}{h_{i_0}^\beta}}{\sum_{i=1}^n \frac{1}{h_{i_0}^\beta}} \quad (9)$$

Der Potenzparameter β kann eine positive reelle Zahl sein. In der Regel wird die quadratische IDW-Interpolation verwendet (LI & HEAP 2008).

Folgende Eigenschaften charakterisieren die IDW-Interpolation: Die IDW-Methode kann sowohl global als auch lokal implementiert sein. Im Rahmen dieser Arbeit wird der lokale Ansatz betrachtet, bei dem eine bestimmte Anzahl von Punkten den Wert des interpolierten Punkts beeinflussen. Bei gleichmäßig verteilten und dichten Daten liefert die IDW-Methode zuverlässigere Ergebnisse, als bei unregelmäßigen und spärlichen Daten (WATSON & PHILIP 1985). Da es sich bei der IDW um ein gewichtetes Mittel handelt, können keine Werte

produziert werden, die die Extremwerte der Eingaben übersteigen. Sofern nicht bereits in den bekannten Datenpunkten enthalten, werden also auch mit der IDW-Methode keine Spitzen, Löcher, Kämme oder Täler erzeugt.

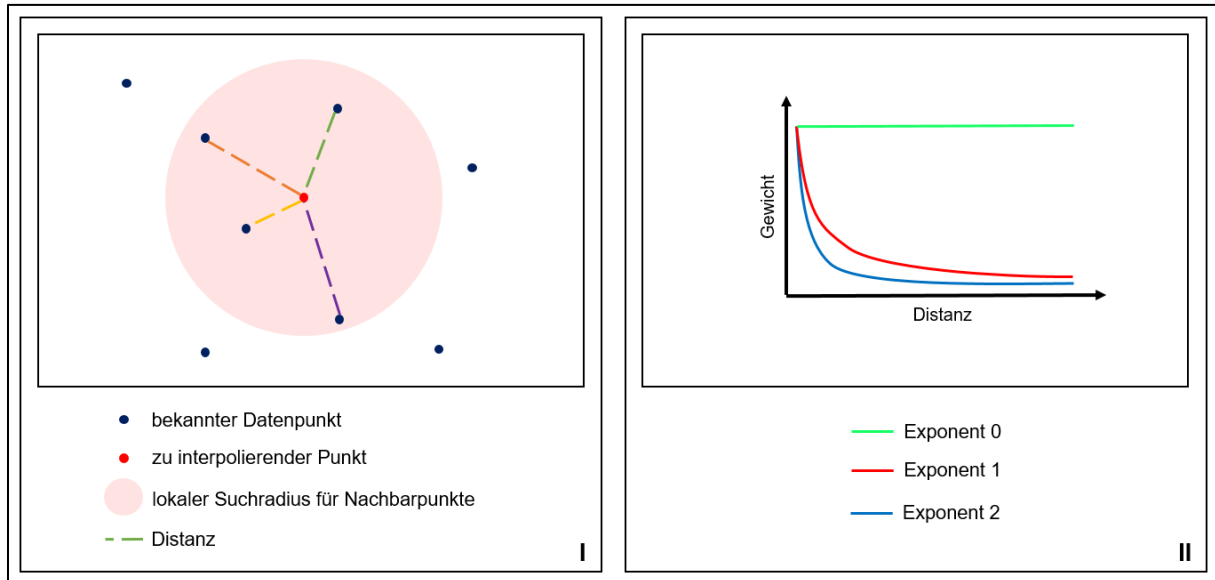


Abbildung 122: Schematische Darstellung des Prinzips der IDW-Interpolationsmethode (I). Im rechten Teil der Abbildung ist der Einfluss des Exponenten schematisch dargestellt (II).

Die IDW-Interpolationsergebnisse können durch einige Parameter gesteuert werden. So beeinflusst beispielsweise der Exponent der Potenzfunktion den Grad der Glättung im Zuge der Oberflächenmodellierung. Ein niedriger Exponent hat glättenden Einfluss auf die Oberflächenmodellierung, da hierdurch die Nachbarpunkte bei der Interpolation ein höheres Gewicht erhalten (LI & HEAP 2008). Durch einen hohen Exponenten erhält die erzeugte Oberfläche potentiell mehr Details, da hierbei Daten nahe am unbekannten Punkt maßgeblichen Einfluss haben. Ein weiterer Parameter ist der Suchradius, mit dessen Hilfe die lokale Punktnachbarschaft definiert wird. Dieser kann fest oder variabel sein. Im Falle eines variablen Suchradius ist die Anzahl relevanter Nachbarpunkte zu definieren. Bei der Wahl des Suchradius sollte beachtet werden, dass mit zunehmender Entfernung von Punkten zueinander nicht mehr von einer räumlichen Korrelation ausgegangen werden kann.

Zu den wesentlichen Eigenschaften der IDW-Interpolation zählen, dass sie ein exakter Interpolator ist und die Werte der generierten Oberflächenmodelle innerhalb der Minima und Maxima der Messdaten liegen. Der Vorteil der IDW-Interpolationsmethode liegt in ihrer wenig komplexen Berechnung. Dadurch können Oberflächen vergleichsweise schnell interpoliert werden. Ein Nachteil ist das in der Literatur als „Bull’s-Eye Effect“ bezeichnete Phänomen (DUMITRU, PLOPEANU & BADEA 2013), welches sich in Form konzentrischer Wertniveaus um

bekannte Datenpunkte äußert (siehe Abbildung 123). Ein solches Ergebnis erzeugt die IDW-Methode oftmals, wenn es um einen zu interpolierenden Punkt herum nur wenige bekannte Datenpunkte gibt, deren Werte zusätzlich stark streuen.

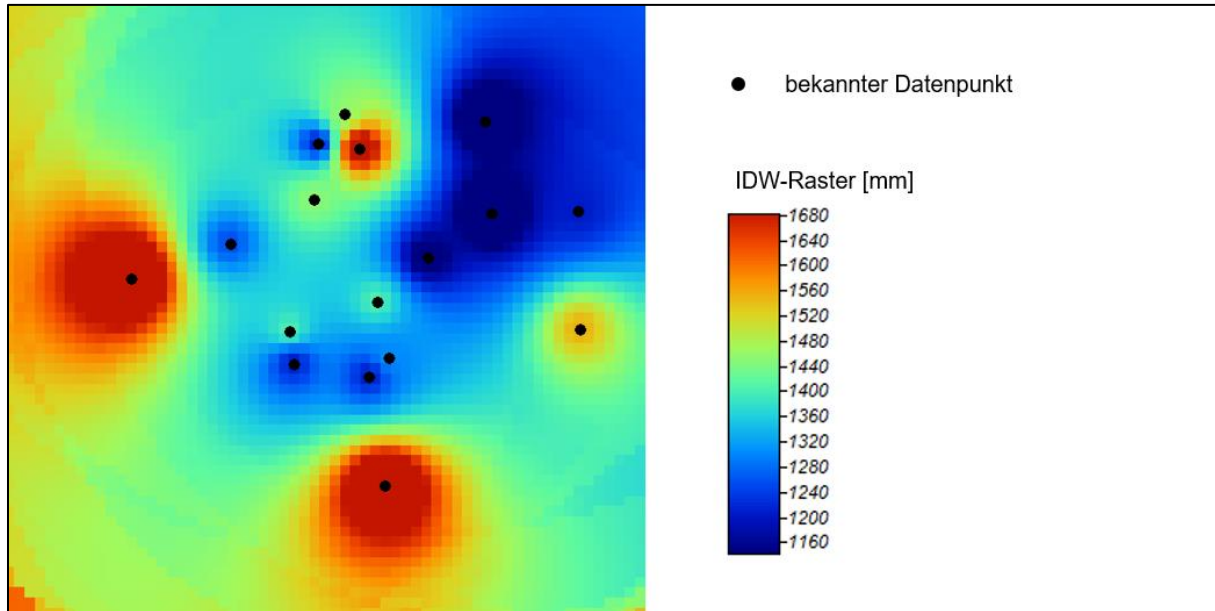


Abbildung 123: Beispiel für den sogenannten "Bull's-Eye Effect", wie er bei der IDW-Interpolation auftreten kann.

4.2.4 Modified Quadratic Shepard

Die Modified-Quadratic-Shepard-Interpolationsmethode (MQS) basiert auf dem IDW-Interpolator. Wie im vorherigen Kapitel beschrieben, ist ein entscheidender Nachteil der IDW-Methode eine Übergewichtung von Datenpunkt-Clustern. Um diesen Effekt zu reduzieren, wurden eine Reihe von Optimierungen der Gewichtungsfunktion vorgestellt, wie beispielsweise die Auswahl naher Punkte mittels gestaffelter Suchradien oder der Einbezug von Information über die räumliche Anordnung von Nachbarpunkten (SHEPARD 1968). Ein Nachteil der Gewichtungsfunktionen des Shepard'schen IDW-Algorithmus ist jedoch, dass sie Datenpunkten, die weit vom zu interpolierenden Punkt entfernt sind, einen vergleichsweise großen Einfluss verleiht. Zur Verbesserung der Genauigkeit der Oberflächenapproximation und Effizienz wurden daher diverse Varianten des ursprünglichen Shepard-Operators entwickelt (DELL'ACCIO & DI TOMMASO 2016). Die von FRANKE & NIELSON (1980) vorgestellte und von RENKA (1988) weiterentwickelte MQS-Variante nutzt für die Oberflächeninterpolation lokale kleinste Quadrate. Sie kann als eine nach dem inversen Abstand gewichtete Methode der kleinsten Quadrate beschrieben werden und unterscheidet sich vom ursprünglichen Shepard-

Ansatz dadurch, dass anstelle einer konstanten Basisfunktion eine quadratische Basisfunktion genutzt wird (DONATO & BELONGIE 2002). Im Zuge der Interpolation mithilfe der MQS-Methode findet eine lokale Polynomanpassung statt. Dabei wird ein Polynom zweiten Grades (quadratische Form) als lokale Annäherung der Funktion genutzt (THACKER et al. 2010). Der Algorithmus umfasst zwei Schritte: zunächst wird eine lokale Anpassung einer quadratischen Fläche um jeden bekannten Datenpunkt berechnet (Least Squares Fit). Anschließend werden die interpolierten Werte anhand einer entfernungsgewichteten Mittelung der zuvor berechneten quadratischen Anpassungen für benachbarte Datenpunkte erstellt. Für n bekannte Nachbardenpunkte um einen zu interpolierenden Datenpunkt N_0 ergibt sich das Interpolationsergebnis aus der nachstehenden Formel, bei der die sogenannte Knotenfunktion Q_k der mit der Methode der kleinsten Quadrate berechnete quadratische Funktionswert ist, der für die Referenzpunkte gewichtet wird (CAKIR 2011).

$$N_0 = \frac{\sum_{k=1}^n W_k Q_k}{\sum_{k=1}^n W_k} \quad (10)$$

Der Gewichtungswert W_k und die Knotenfunktion Q_k sind wiederum durch die nachstehende Gleichung definiert.

$$W_k = \left[\frac{(R_w - d_k)}{R_w d_k} \right]^2 \quad (11)$$

$$Q_k(x, y) = c_{k2}(x - x_i)^2 + c_{k3}(x - x_i)(y - y_i) + c_{k4}(y - y_i)^2 + c_{k5}(x - x_i) + c_{k6}(y - y_i) + f_k$$

Die Koeffizienten $\vec{c} = (c_{k2}, c_{k3}, \dots, c_{k6})$ der Knotenfunktion Q_k werden nach der Methode der kleinsten Quadrate wie folgt bestimmt.

$$\min_{\vec{c}} \sum_{\substack{i=1 \\ k \neq i}}^n W_i [f_k + c_{k2}(x_i - x_k)^2 + \dots + c_{k6}(y_i - y_k) - f_i]^2 \quad (12)$$

Dabei ist W_i die Gewichtung im Zuge der obigen Ermittlung der Koeffizienten für die Knotenfunktion Q_k .

$$W_i = \left[\frac{(R_q - d_i)}{R_q d_i} \right]^2 \quad (13)$$

Bevor die Gewichtungswerte W_k und W_i sowie die Knotenfunktion Q_k berechnet werden können, müssen die entsprechend erforderlichen Kreisradien R_w und R_q bestimmt werden. In den nachstehenden Formeln steht n_q für die Anzahl der Datenpunkte, die für die Berechnung der quadratischen Funktion verwendet werden und n_w gibt die Anzahl der Datenpunkte an, die für die Berechnung der Gewichtungswerte verwendet werden. D bezeichnet den maximalen Abstand zwischen zwei Punkten im Datensatz und n die Gesamtzahl der Punkte im Datensatz. n_q und n_w sind die vom Benutzer festgelegten Parameter.

$$\begin{aligned} R_q &= \frac{D}{2} \sqrt{\frac{n_q}{n}} \\ R_w &= \frac{D}{2} \sqrt{\frac{n_w}{n}} \end{aligned} \quad (14)$$

Der Quadratic-Neighbors-Parameter n_q legt demnach im ersten Berechnungsschritt die Größe der lokalen Nachbarschaft und damit die Nachbarpunkte fest, in die die quadratische Fläche eingepasst wird. Der Weighting-Neighbors-Parameter n_w definiert die Anzahl der lokalen Nachbarpunkte im zweiten Berechnungsschritt. Bei RENKA (1988) wird das Parametersetting $n_q = 13$ und $n_w = 19$ für den zweidimensionalen Anwendungsfall vorgeschlagen und diese Standard-Parametrisierung erläutert.

Die MQS-Methode kann entweder ein exakter oder ein glättender Interpolator sein und zeichnet sich im Allgemeinen durch ihre Effizienz und Genauigkeit aus (MUSTAFA, SHAH & ASIM 2008). Die Werte der generierten Oberflächenmodelle müssen nicht innerhalb der Minima und Maxima der Messdaten liegen (siehe Abbildung 124).

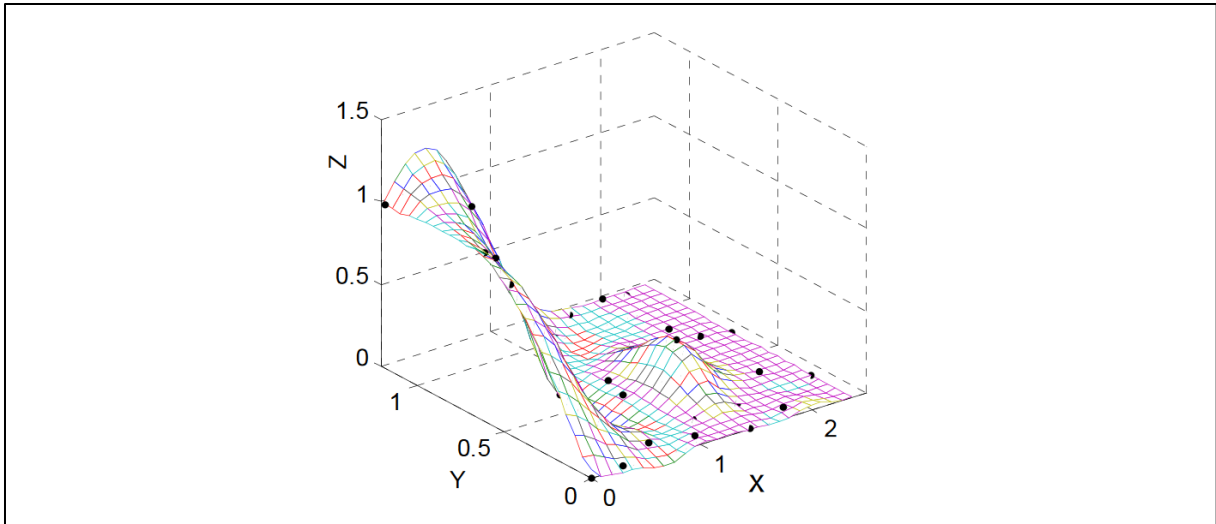


Abbildung 124: Beispiel für eine durch MQS-Interpolation erzeugte Oberfläche durch bekannte Datenpunkte (schwarze Punkte), die Höhenwerte zwischen 0 und 1 haben. Die Grafik zeigt, dass die extremen Z-Werte unter- und überschritten werden können (MUSTAFA, SHAH & ASIM 2008).

4.2.5 Thin Plate Spline

Das Thin-Plate-Spline-Verfahren (TPS) ist eine Variante aus der Gruppe der Spline-Interpolationsverfahren. Als Splines bezeichnet man glatte, harmonische mathematische Kurven dritten Grades, wobei es sich bei Thin Plate Splines um eine 2D-Verallgemeinerung der kubischen Spline-Funktion handelt (DONATO & BELONGIE 2002). Der Name Thin Plate Splines bezieht sich auf dünne, flexible Metallflächen, die mithilfe einer minimalen Biegeenergie so verformt werden, dass sie durch vorgegebene Zwangspunkte verlaufen (MITASOVA & MITAS 1993).

Für den zweidimensionalen Anwendungsfall werden zur Berechnung von Splines oftmals radiale Basisfunktionen (RBF) als zugrundeliegende mathematische Funktion genutzt. RBF basierte Interpolationen können über die nachstehende Formel beschrieben werden (ROUSSOS & BAXTER 2005). Dabei steht i für die reellen Koeffizienten, x_i für die Punkte im 2D-Raum, $\|\cdot\|$ für die euklidische Norm und φ für die Basisfunktion.

$$s(y) = \sum_{i=1}^N \lambda_i \varphi(\|y - x_i\|) \quad (15)$$

Als eine von mehreren möglichen Basisfunktionen kann die TPS-Funktion dienen, welche durch nachstehende Formel definiert werden kann. Innerhalb der radialen Basisfunktion steht r für die euklidische Distanz zwischen bekanntem und approximiertem Punkt x_j und y . Unter der Bedingung, dass die mittlere quadratische Krümmung minimal wird, wird für jeden der N Datenpunkte je eine radiale Basisfunktion berechnet.

$$\varphi(r) = r^2 \log r \quad (16)$$

Im Kern handelt es sich bei der TPS-Methode um ein Minimierungsproblem, bei dem die Fehlerquadratsumme der angepassten Funktion minimiert wird und die Funktion dabei eine gewisse Glattheitsanforderung erfüllen muss. In der nachstehenden Formel steht J für eine Straffunktion, welche eine zu starke Anpassung der Spline-Funktion f an die Datenpunkte durch Glättung ausgleicht und deren Einfluss durch den Regularisierungsparameter λ gesteuert wird (ROSSITER 2016). Ein Regularisierungswert von $\lambda = 0$ bedeutet, dass keine Glättung stattfindet und die interpolierte Fläche exakt durch die Datenpunkte verläuft. Je höher λ ist, desto ungenauer ist die resultierende Anpassung. In Abbildung 125 ist der Einfluss schematisch dargestellt.

$$\min_f \sum_{i=1}^N \{y_i - f(x_i)\}^2 + \lambda J[f] \quad (17)$$

Der namensgebenden Analogie des Interpolationsverfahrens folgend, minimiert das TPS-Verfahren also die Biegeenergie, mit der eine dünne flexible Metallplatte gezwungen wäre, durch die Datenpunkte zu verlaufen. Durch den zweiten Teil der vorstehenden Formel kann dieser Zwang relativiert werden, in dem er zugunsten einer glatteren Oberfläche zulässt, dass die Interpolationsfunktion lediglich in der Nähe der Datenpunkte bleibt. Eine detaillierte Beschreibung der mathematischen Grundlagen sind bei MITASOVA & MITAS (1993), DONATO & BELONGIE (2002), WOOD (2003) und KELLER & BORKOWSKI (2019) zu finden.

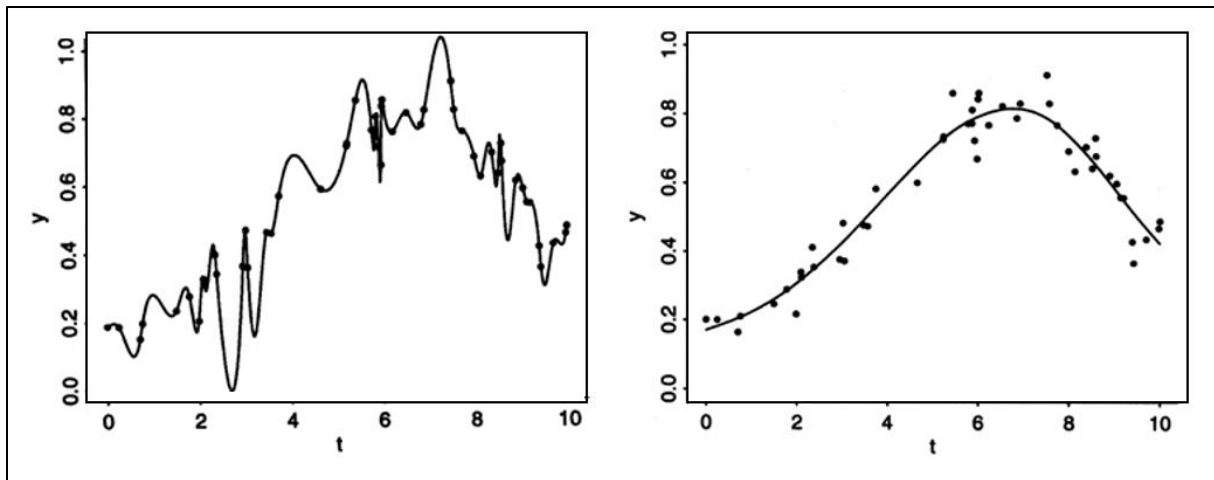


Abbildung 125: Einfluss des Regularisierungsparameters. Ein niedriger Wert ergibt die linke Kurve, die durch die Datenpunkte verläuft. Ein hoher Wert erlaubt die Abweichung der rechten Kurve von den Datenpunkten (BENINI & PONZA 2010).

Das TPS-Interpolationsverfahren erzeugt eine minimal gekrümmte, exakt durch die Datenpunkte verlaufende Oberfläche, die kontinuierlich und differenzierbar ist. Anders als beispielsweise bei der IDW-Methode ist es bei TPS möglich, dass die durch die Datenpunkte vorgegebenen lokalen Minima und Maxima unter- bzw. überschritten werden. Der Regularisierungsparameter ermöglicht es, Messfehler oder andersartige Artefakte in den Datenpunkten zu behandeln, jedoch kann ein zu groß gewählter Regularisierungsparameter auch zu unrealistisch glatt anmutenden Oberflächen führen (LI & HEAP 2008).

4.3 Geomorphometrie

Wie in Kapitel 2.3.2 beschrieben, ist die Geomorphometrie die Wissenschaft der topographischen Quantifizierung. Ihr operativer Schwerpunkt ist die Extraktion von Landoberflächenparametern und Objekten aus digitalen Höhenmodellen (PIKE, EVANS & HENGL 2009). Eine Landoberfläche besteht HENGL & EVANS (2009) zufolge aus einer endlichen Anzahl von spezifischen Formelementen, wie beispielsweise Senken, Erhebungen, Kammlinien, Tiefenlinien, Pässen und Bruchlinien. Diese gilt es im Rahmen geomorphometrischer Anwendungen mit geeigneten Algorithmen aus digitalen Höhenmodellen zu extrahieren (MACMILLAN & SHARY 2009).

Für den Bereich der Straßenzustandserfassung stellt sich grundlegend eine sehr ähnliche Aufgabe, mit dem Unterschied, dass anstelle von Formelementen der Landoberfläche Schäden im Fokus des Interesses stehen. Die in Kapitel 2.1.3.2 beschriebenen Schäden können unter dem Begriff Oberflächendiskontinuitäten subsummiert werden – seien es nun Unstetigkeiten in der

Höhe (z. B. Risse, Ausbrüche, offene Nähte, Ausmagerungen) oder Unstetigkeiten in der Reflexionsintensität infolge von Materialunterschieden (z. B. Flickstellen, Bindemittelaustritt). Zwischen manchen Formelementen des Georeliefs und Oberflächenschäden einer Straße besteht eine skalenabhängige, morphologische Analogie. Wie Abbildung 126 zeigt, kann ein Ausbruch in einer Fahrbahndecke morphologisch einer Doline entsprechen, ein klaffender Riss einem Paläo-Gerinnebett oder ein abgeplatzter Fahrbahnrand einem Kliff. Deshalb werden Fahrbahnschäden im Rahmen dieser Arbeit als Geoobjekte interpretiert, die mithilfe geomorphometrischer Algorithmen aus digitalen Straßenoberflächenmodellen extrahiert werden, um zunächst mittels manueller Kartierung Trainingsdaten für einen Deep-Learning-Ansatz zu generieren und später als Eingangsdaten für eine automatische Analyse mittels eines semantischen Segmentierungsnetzwerks zu fungieren. Aus diesem Grund werden in den folgenden Unterkapiteln typische Algorithmen der Geomorphometrie zur Berechnung von Landoberflächenparametern im Sinne von PIKE, EVANS & HENGL (2009) vorgestellt, mit deren Hilfe klassischerweise geogene sowie anthropogene Hohlformen im Georelief kartiert werden. Aufgrund der vorstehend beschriebenen Analogie sind diese Verfahren potentiell in der Lage, auch Schäden aus Straßenoberflächen zu extrahieren.

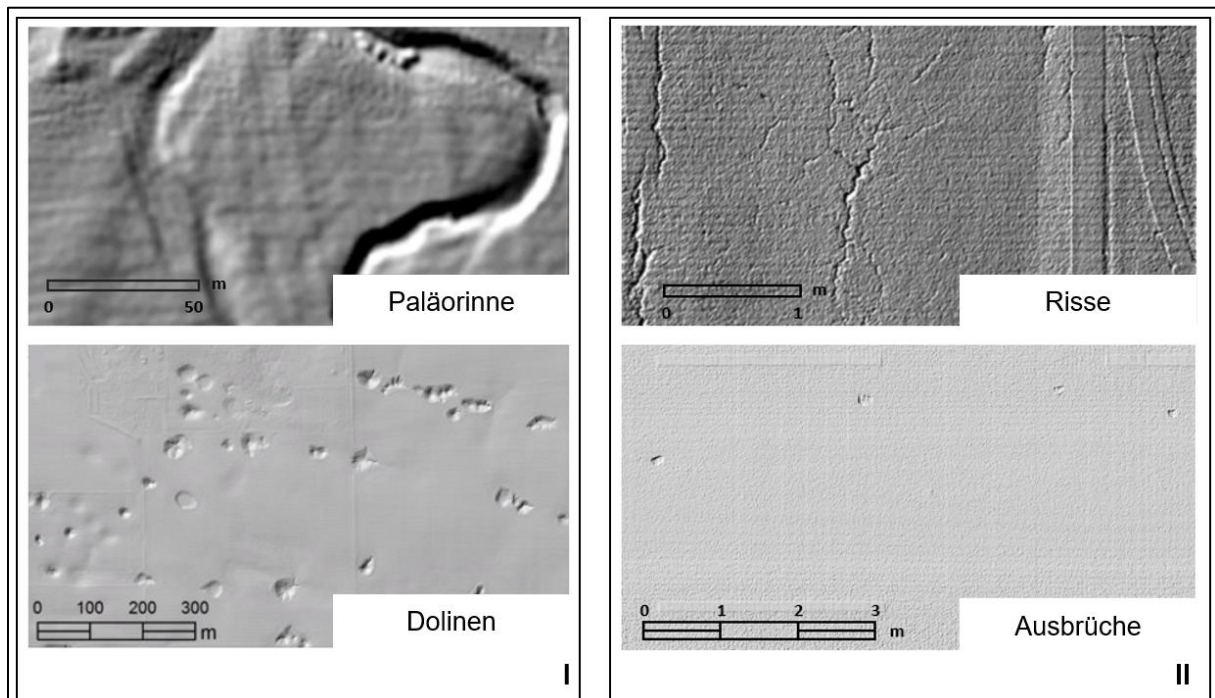


Abbildung 126: Analogie zwischen Formelementen des Georeliefs (I) und Straßenschäden (II) (Teilabbildung I unten nach WU, DENG & CHEN 2016).

Zur Verdeutlichung der skalenabhängigen Formanalogie und der Übertragbarkeit von primär zur Analyse von Formen des Großreliefs entwickelten Algorithmen der Geomorphometrie auf

Formen des Nano- und Pikoreliefs einer Straßenoberfläche, werden die in den folgenden Unterkapiteln beschriebenen Verfahren jeweils an einem DGM (Airborne Laserscan) und einem Straßenoberflächenmodell demonstriert. Bei Ersterem handelt es sich um einen Beispieldatensatz, der von der Bayerischen Vermessungsverwaltung zum Download frei zur Verfügung gestellt wird (BAYERISCHE VERMESSUNGSVERWALTUNG 2022a). In Abbildung 127 ist ein Ausschnitt dieses DGM rechts dargestellt. Mithilfe des Digitalen Orthophotos (Abbildung 127 links), welches ebenfalls als Beispiel-Datensatz heruntergeladen werden kann, lässt sich das DGM interpretieren (BAYERISCHE VERMESSUNGSVERWALTUNG 2022b). Der dargestellte, etwa 400 m lange Profilschnitt durch das DGM zeigt eine etwa 2 m bis 3 m tief in das Gelände eingetieft angelegte Straße.

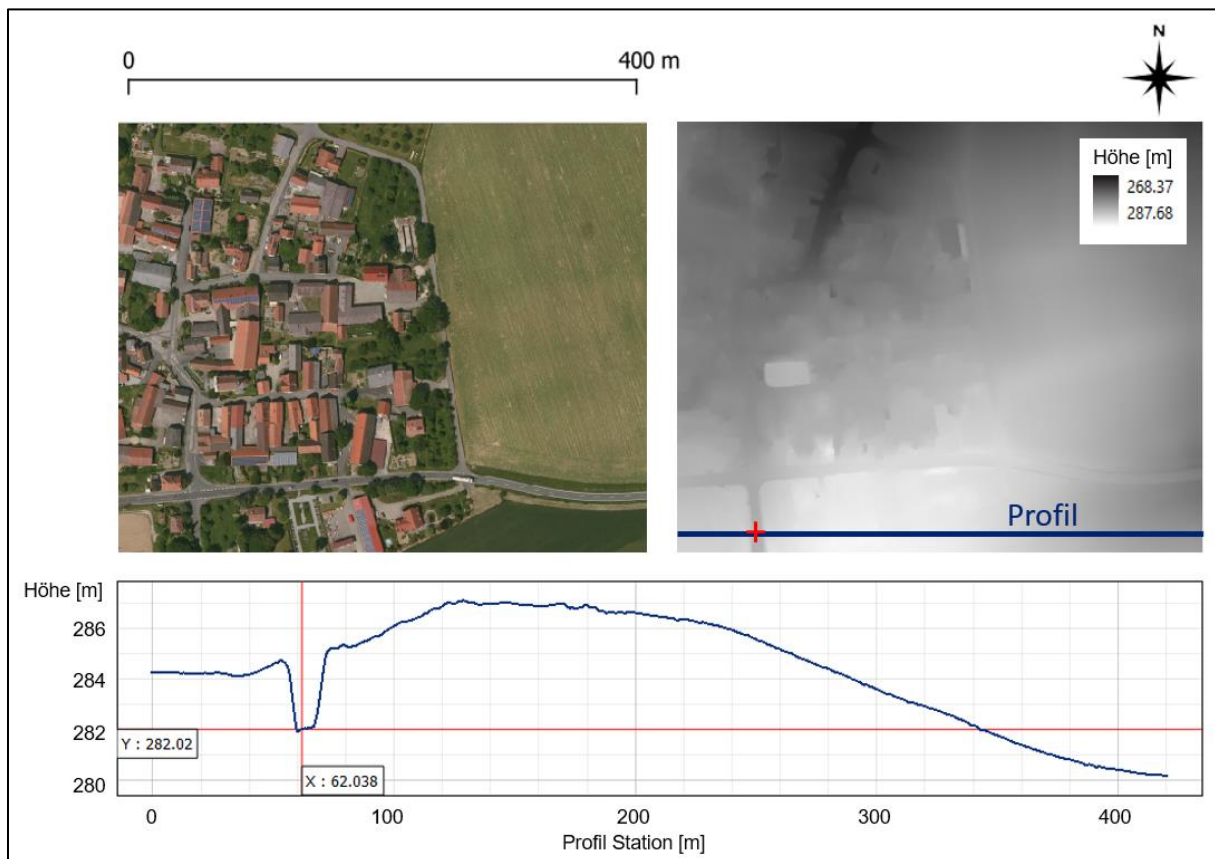


Abbildung 127: Profilschnitt durch ein Digitales Geländemodell. Das rote Fadenkreuz fokussiert eine ins Gelände eingetieft gebaute Straße (Datengrundlage: BAYERISCHE VERMESSUNGSVERWALTUNG 2022a, 2022b).

Morphologisch ähnelt dies dem charakteristischen Querschnitt eines Risses oder Schlaglochs in einer Fahrbahnoberfläche (siehe Abbildung 128). Dort ist links zur besseren visuellen Interpretation des Oberflächenmodells ein hochauflösendes TP3-Oberflächenkamerabild abgebildet, das mit dem Messsystem I.R.I.S 3 aufgenommen wurde. Zu sehen sind mittig am linken Rand ein Schlagloch und am rechten Fahrbahnrand, entlang der gepflasterten Rinne, längs zur

Fahrtrichtung verlaufende Risse. Im rechten Teil der Abbildung 128 ist ein Oberflächenmodell auf Basis des PPS-Laserscans, der ebenfalls mit dem I.R.I.S 3 System erfasst wurde, dargestellt.

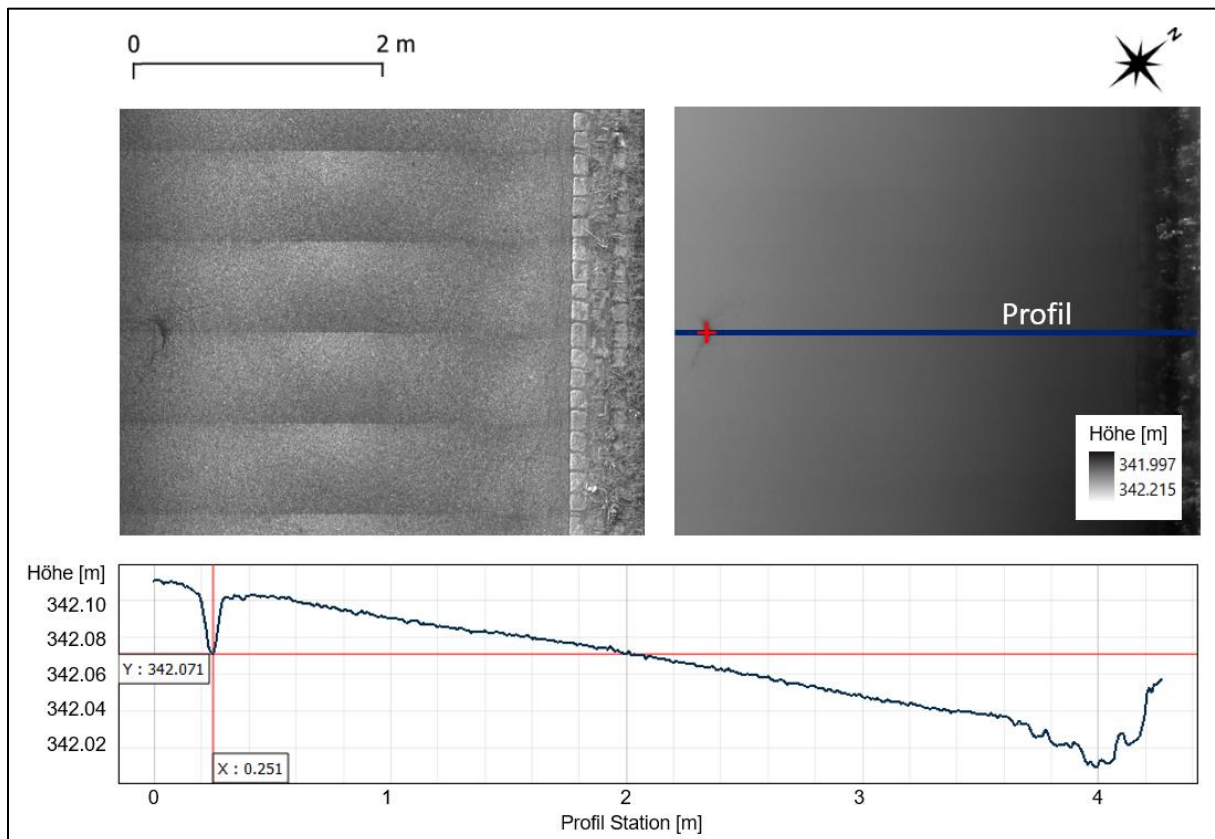


Abbildung 128: Profilschnitt durch ein Oberflächenmodell, das auf Basis von PPS-Laserscans erzeugt wurde. Das rote Fadenkreuz fokussiert einen Ausbruch (Schlagloch) in der Asphaltdecke.

4.3.1 Analytical Hillshading

Die Reliefschummerung ist eine etablierte kartographische Methode, mit deren Hilfe ein plastischer Eindruck von Reliefformen vermittelt werden kann (CLARKE 1995). Schon vor dem Aufkommen der digitalen Kartographie wurde beispielsweise die Böschungsschummerung genutzt, um die Steilheit von Hängen durch sogenannte Schraffen graphisch zu charakterisieren. Die heutige automatische Schräglichtschummerung (engl.: Hillshading oder Shaded Relief) basiert auf digitalen Höhenmodellen und gehört zu den Standardfunktionen moderner GIS-Anwendungen (RASE 2021).

Die mathematischen Grundlagen zur Berechnung von Hillshade-Rastern sind bei HORN (1981) detailliert erklärt. Für jeden Punkt in einem Gitter bzw. für jede Dreiecksfläche in einem TIN wird der dreidimensionale Vektor berechnet, der senkrecht zur Oberfläche steht. Die Oberflächennormale definiert sowohl die Neigung des zugrundeliegenden Oberflächenelements als

auch die Ausrichtung dieser Neigung. Im nächsten Schritt wird eine virtuelle Beleuchtungsquelle definiert, welche die Oberfläche aus einer bestimmten Richtung (engl.: Azimuth) und einer bestimmten Höhe über dem virtuellen Horizont (engl.: Altitude) bescheint. Letzteres wird auch als Zenith bezeichnet und repräsentiert den vertikalen Einfallswinkel des Lichts ausgehend vom Höhenmodell (CLARKE 1995). Auf Grundlage der relativen Ausrichtung des Oberflächenelements zur virtuellen Beleuchtungsquelle wird für jede Zelle eines digitalen Höhenmodells ein Beleuchtungswert zwischen 0 und 255 berechnet. Ein Wert von 255 wäre gegeben, wenn der Normalenvektor direkt auf die Beleuchtungsquelle zeigt. Je mehr der Normalenvektor von der Beleuchtungsquelle abweicht, desto niedriger werden die Beleuchtungswerte. Ein Beleuchtungswert von 0 bedeutet, dass sich dieser Punkt im virtuellen Schatten befindet, seine Oberflächennormale also weg von der virtuellen Beleuchtungsquelle zeigt. In Abbildung 129 ist dies schematisch dargestellt.

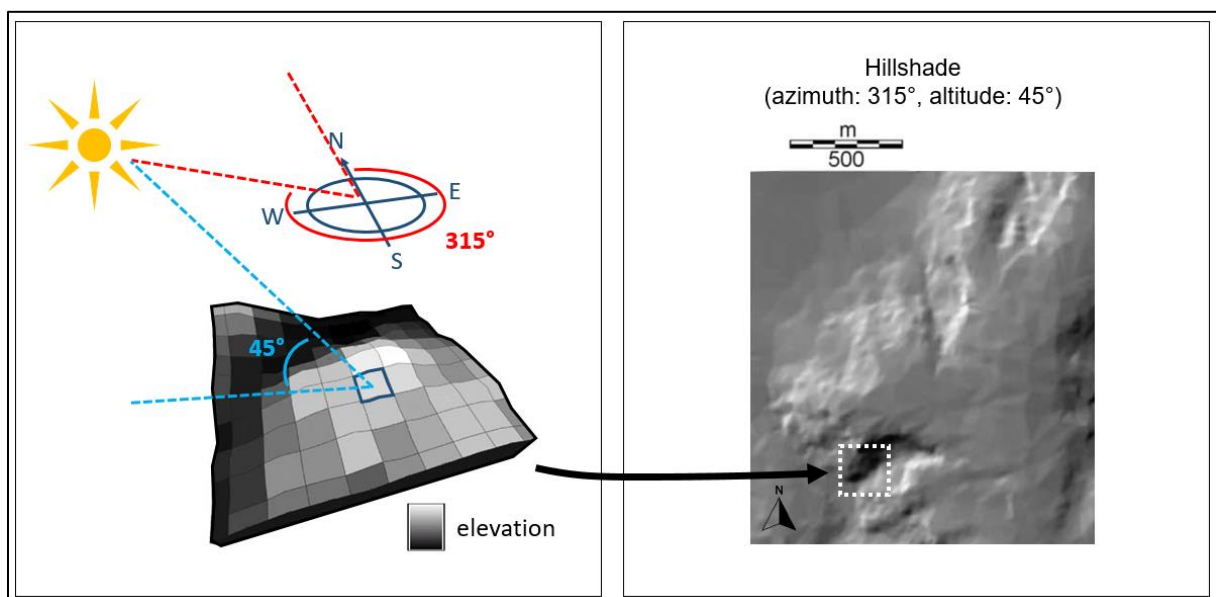


Abbildung 129: Prinzipskizze zur Berechnung eines Hillshades (nach WOOD 2008; VAN DEN EECKHAUT et al. 2005).

Eine Rekonstruktion der ursprünglichen Höhenwerte aus dem Helligkeitswert des Hillshades ist nicht mehr möglich, da sich der Beleuchtungswert aus dem Kosinus der Winkeldifferenz zwischen dem Oberflächennormalenvektor und dem Beleuchtungsvektor berechnet (RASE 2021).

Beispiele für typische Anwendungen des Analytical Hillshading sind die Visualisierung von digitalen Höhenmodellen (PINGEL & CLARKE 2014), die manuelle oder automatische Kartierung von Reliefformen (VAN DEN EECKHAUT et al. 2005; WU, DENG & CHEN 2016) und Solarpotentialanalysen (HONG et al. 2017).

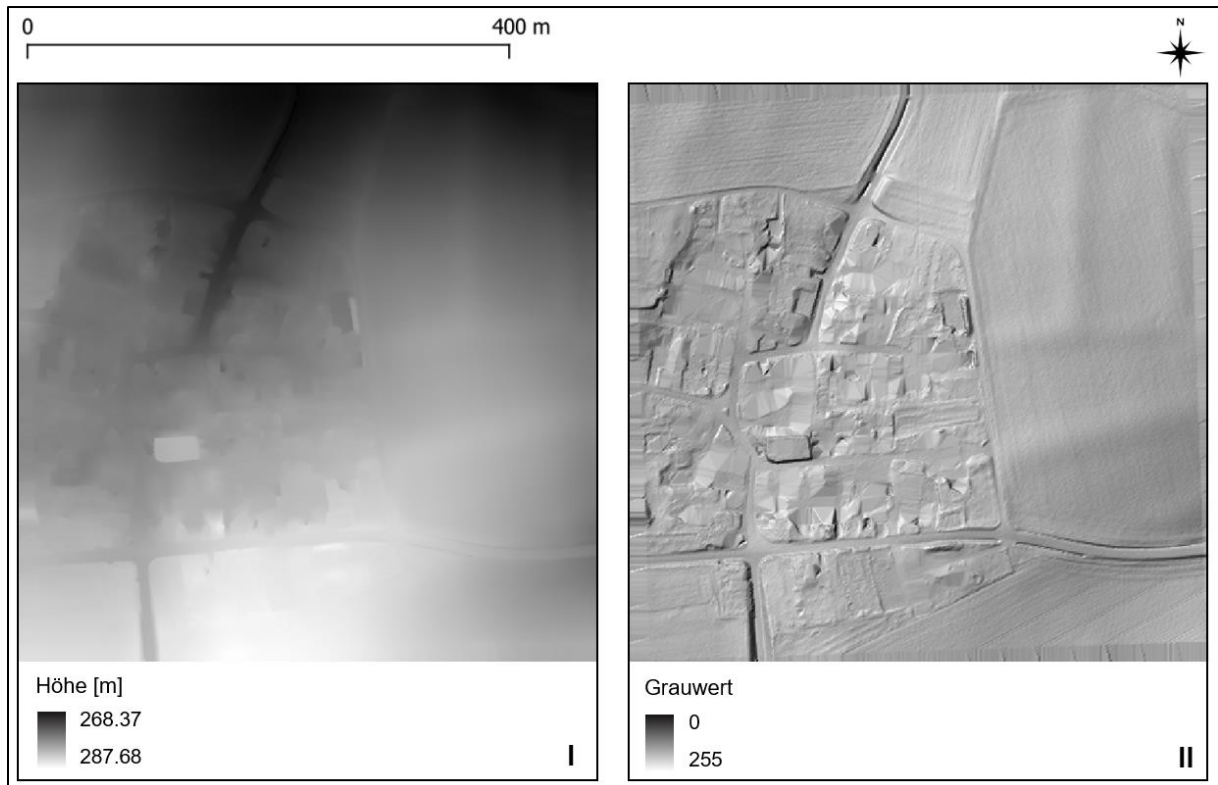


Abbildung 130: Hillshade Raster (II) für ein Airborne-LiDAR-DGM (I) (Datengrundlage: BAYERISCHE VERMESSUNGSVERWALTUNG 2022a).

Wie in Abbildung 130 dargestellt, ist der wesentliche Vorteil von Hillshades, dass sie einen plastischen Eindruck der Geländeoberfläche vermitteln. Feine Strukturen, wie beispielsweise die Ackerfurchen in der oberen linken Ecke, sind im DGM visuell nicht direkt erkennbar. Diese Art der Geländedarstellung lässt sich deshalb so intuitiv interpretieren, weil sie der menschlichen Sehgewohnheit entspricht: Harte Schattenkanten treten eher an steilen Oberflächenunstetigkeiten auf, wogegen weiche Graustufengradienten eher auf leicht bewegte Oberflächenvariationen hindeuten.

Abbildung 131 zeigt rechts ein Hillshade, welches auf Basis eines PPS-Oberflächenmodells (links) berechnet wurde. Die Oberflächenschäden (Ausbrüche links und Risse rechts) wie auch die gepflasterte Rinne (rechter Rand) heben sich visuell deutlich von der Asphaltoberfläche ab. Die Ausrichtung von Rissen relativ zur virtuellen Beleuchtungsquelle ist entscheidend dafür, ob und wie gut ein Riss im Hillshade sichtbar ist.

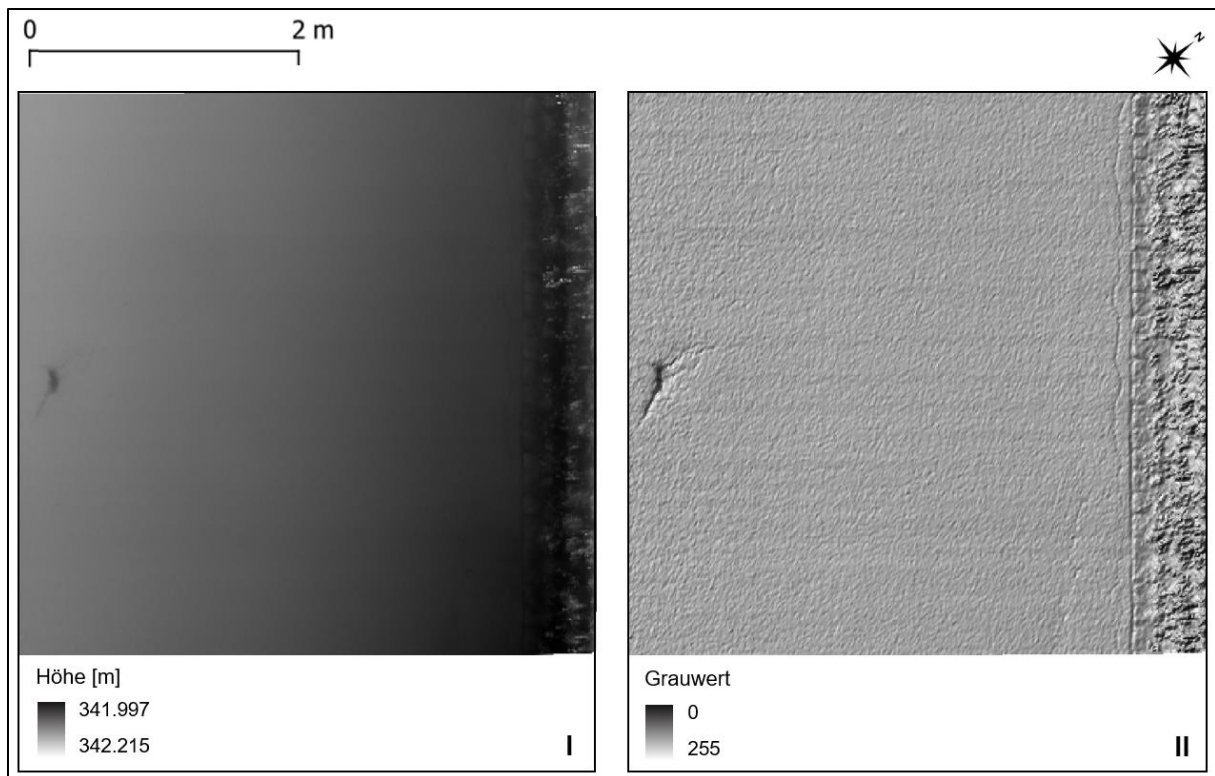


Abbildung 131: Hillshade mit Azimuth 315°, Altitude 45° und zweifacher Überhöhung für eine Straßenoberfläche mit Rissen und Schlaglöchern (II). Teilabbildung I zeigt das zugrundeliegende Oberflächenmodell.

Es existieren diverse Algorithmen und Implementierungen der Schräglichtschummerung (FARMAKIS-SEREBRYAKOVA & HURNI 2020). In der vorliegenden Arbeit wird der Standardalgorithmus genutzt, welcher in SAGA GIS als Tool „Analytical Hillshading“ in der Library „ta_lighting“ implementiert ist. Auch für die beiden Parameter *azimuth* und *declination* (entspricht: Altitude) werden die Standardwerte verwendet (siehe Abbildung 132). Der Wert für die Überhöhung wurde als Kompromiss zwischen zu starker Verstärkung von Oberflächenstruktur und verbesserter Sichtbarkeit von feinen Oberflächenschäden auf den Faktor 2 gesetzt.

```
REM Tool: Analytical Hillshading
saga_cmd -c=1 ta_lighting 0 ^
-ELEVATION %WORK%\SagaTemps\20285_IRIS13_20210402_001_PPS_0000_000080_Z.sgrd ^
-SHADE %WORK%\SagaTemps\20285_IRIS13_20210402_001_PPS_0000_000080_HS.sgrd ^
-METHOD 0 ^
-POSITION 0 ^
-AZIMUTH 315.000000 ^
-DECLINATION 45.000000 ^
-EXAGGERATION 2.000000 ^
-UNIT 0
```

Abbildung 132: Beispiel für den Kommandozeilenaufwurf, der das Analytical Hillshading Tool ausführt.

4.3.2 Topographic Openness

In ihrer Arbeit zeigen YOKOYAMA, SHIRASAWA & PIKE (2002), wie ein Algorithmus aus der digitalen Bildanalyse (Openness) auf ein DGM angewendet werden kann, um Topographie-merkmale zuverlässiger zu visualisieren und zu interpretieren, als es mit den bis dato gängigen Ansätzen möglich war. Als entscheidenden Nachteil des Hillshading-Verfahrens führen die Autoren an, dass dessen gerichtetes Beleuchtungsmodell vor allem dann problematisch ist, wenn Objekte parallel zur Beleuchtungsrichtung ausgerichtet sind. Demgegenüber ist der als Openness betitelte Ansatz in der Lage, dominante Konkavitäten und Konvexitäten einer Oberfläche zu betonen, ohne dabei zu sensitiv auf Oberflächenungenauigkeiten (Rauschen) zu reagieren und ohne dabei auf eine fixierte, virtuelle Lichtquelle angewiesen zu sein.

Das Funktionsprinzip des Topographic-Openness-Algorithmus beruht, ebenso wie das des Analytical-Hillshading-Ansatzes, auf einer simulierten Beleuchtung der Oberfläche, nur dass an die Stelle einer gerichteten Beleuchtung eine diffuse, homogen helle Beleuchtung tritt. Im Openness-Ansatz wird sie durch eine Kugel simuliert, die auf eine Rasterzelle eines Oberflächenmodells zentriert wird (NOACK 2019).

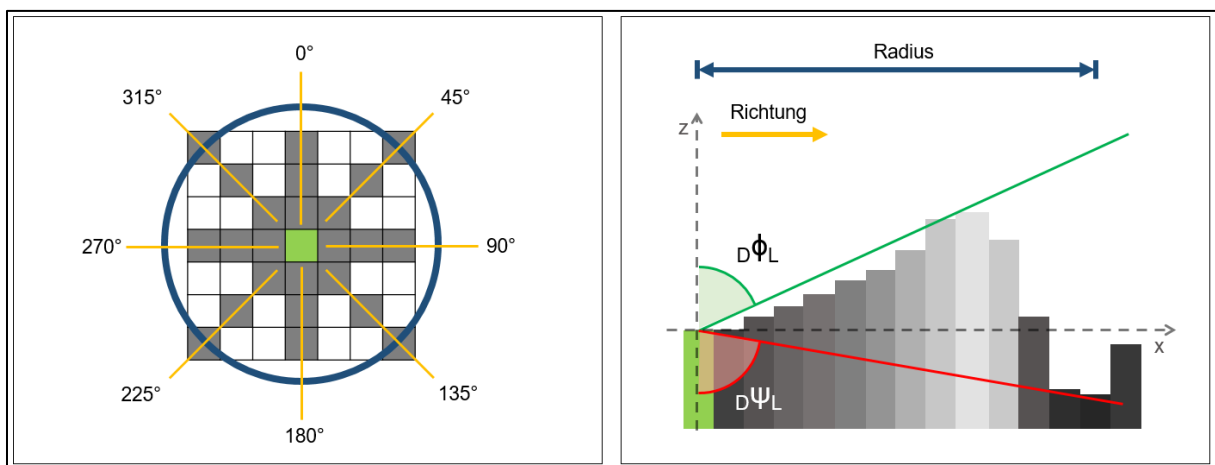


Abbildung 133: Prinzip der Openness-Berechnung. Die Nachbarschaft um eine zentrale Zelle (grün) wird durch acht Richtungen (gelb) und einen Radius (blau) definiert (links). Rechts ist für eine Richtung das Prinzip der Zenit- und Nadirwinkelberechnung abgebildet (nach YOKOYAMA, SHIRASAWA & PIKE 2002).

Für jede Oberflächenzelle wird eine Nachbarschaft analysiert. Sie ist definiert über einen Radius sowie acht von der zentralen Zelle ausgehende Richtungen. In dieser lokalen Nachbarschaft werden Zenitwinkel $D\phi_L$ und Nadirwinkel $D\psi_L$ für jede Richtung berechnet und deren Mittelwert gebildet. Wie in Abbildung 133 rechts dargestellt, wird zwischen positiver und negativer Openness unterschieden. Positive Openness ist demzufolge definiert als der Mittelwert der acht Zenitwinkel $D\phi_L$ für die Richtung D innerhalb des Begrenzungsradius L .

Jeder dieser acht Zenitwinkel wiederum ergibt sich aus dem Winkel, der vom lokalen Horizont und der Verbindungsgeraden zwischen zentraler Zelle und der über dem lokalen Horizont liegenden Nachbarzelle eingeschlossen wird, bei der der Anstellwinkel der Verbindungsgeraden am steilsten ist. Negative Openness hingegen ist definiert als der Mittelwert der acht Nadirwinkel $D\psi_L$ für die Richtung D innerhalb des Begrenzungsradius L . Hier ist pro Richtung der steilste Winkel relevant, der vom lokalen Horizont und der Verbindungsgeraden zwischen zentraler Zelle und der unter dem lokalen Horizont liegenden Nachbarzelle gebildet wird. Negative Landformen, wie z. B. Täler, sind durch niedrige Positive-Openness-Werte gekennzeichnet. Hohe Positive-Openness-Werte treten dagegen an positiven Landformen, wie z. B. Gipfeln, auf.

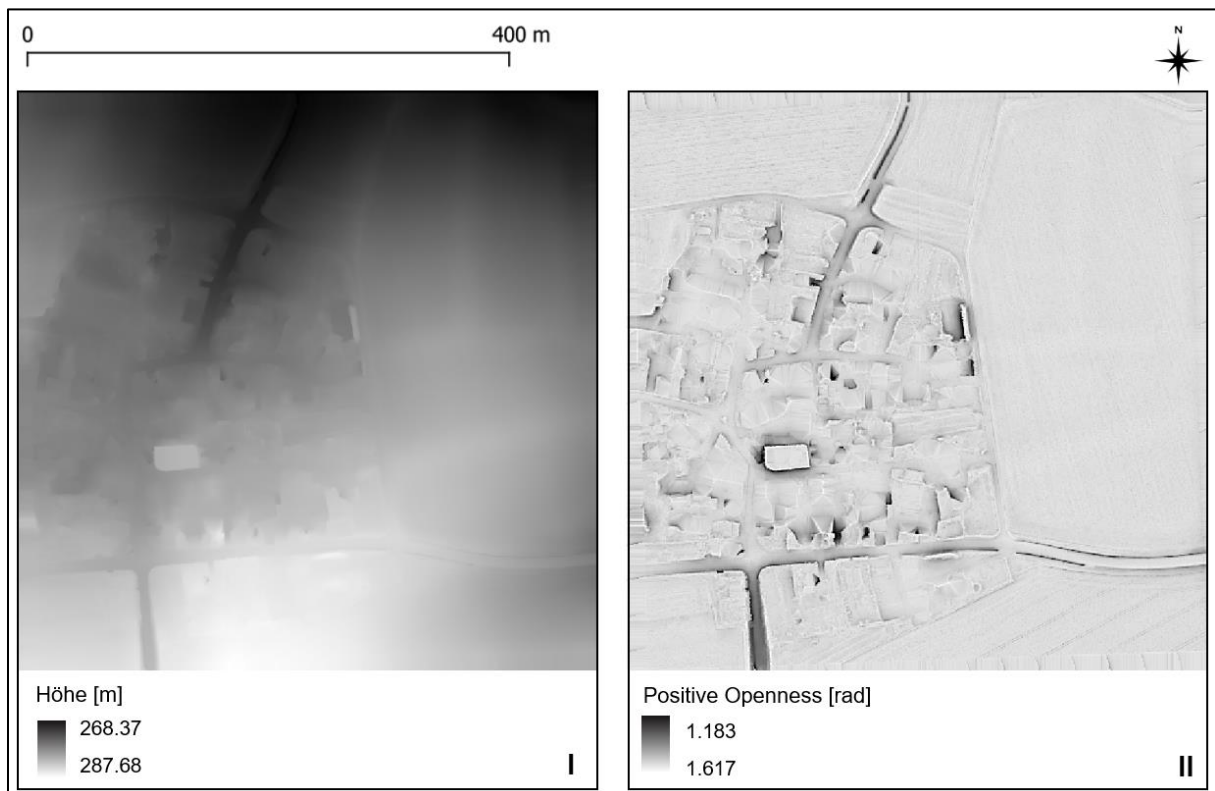


Abbildung 134: Positive Openness bei einem Radius von 10 m (II) für ein Airborne-LiDAR-DGM (I) (Datengrundlage: BAYERISCHE VERMESSUNGSVERWALTUNG 2022a).

Topographic-Openness-Raster haben sich als sinnvolle Grundlage für die computergestützte geomorphologische Kartierung erwiesen (PRIMA et al. 2006). Oft findet der Topographic-Openness-Ansatz auch im Rahmen von archäologischen Fragestellungen Anwendung. DONEUS (2013) nutzt den Topographic-Openness-Algorithmus als eine Visualisierungstechnik, um archäologische Informationen aus DGM zu kartieren. In seiner Publikation zeigt der Autor beispielsweise, wie mithilfe von Positive-Openness-Rastern, die auf Basis von Airborne-

LiDAR-DGM mit 1 m Gitterweite unter Verwendung eines Radius von 7,5 m berechnet wurden, historische Hohlwege und Bombenkrater als dunkle Bildbereiche visualisiert werden können. NOACK (2019) beschreibt, wie der Topographic-Openness-Algorithmus dabei helfen kann, agrarmorphologische Relikte in Form von speziellen Mustern der Ackeroberfläche zu detektieren. Auch sie verwendet ein Airborne-LiDAR-DGM mit 1 m Gitterweite. Auch wenn Topographic-Openness-Werte (Winkel) vergleichsweise schwer zu interpretieren sind, so demonstriert Abbildung 134 doch die Stärken der Topographic Openness als Visualisierungstechnik für Oberflächenstrukturen.

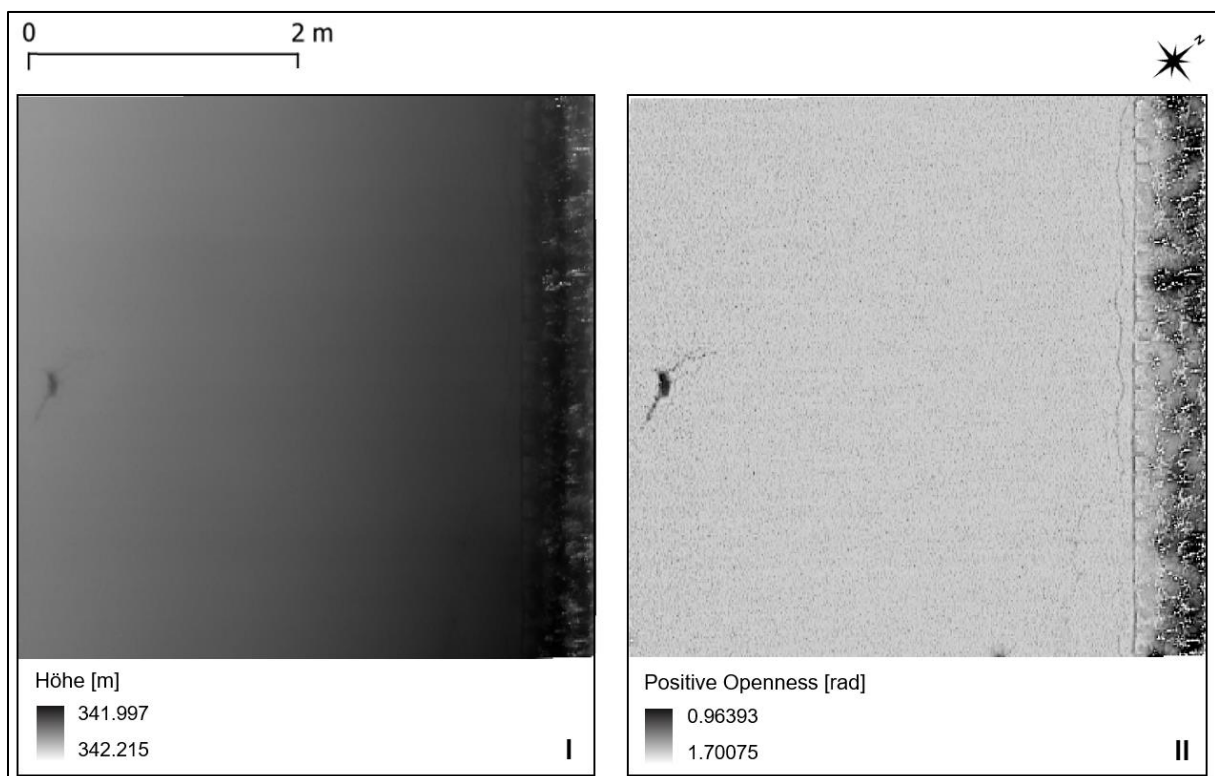


Abbildung 135: Positive Openness mit Radius 0,1 m für eine Straßenoberfläche mit Rissen und Schlaglöchern (II). Teilabbildung I zeigt das zugrundeliegende Oberflächenmodell.

Abbildung 135 zeigt rechts ein Positive-Openness-Raster, welches auf Basis eines PPS-Oberflächenmodells (links) berechnet wurde. Sowohl die Oberflächenschäden als auch die gepflasterte Rinne und der vom Bankett her hereinragende Grasbewuchs heben sich visuell gut von der Asphaltoberfläche ab. Abbildung 136 zeigt den Kommandozeilenaufbau für SAGA GIS, der auf Basis eines Eingabeoberflächenmodells ein Positive-Openness-Raster erzeugt. Das radiale Limit wird auf den Wert 0,1 m eingestellt und die Sektormethode mit acht Richtungen verwendet.


```

REM Tool: Topographic Openness
saga_cmd ta_lighting 5 ^
-DEM %WORK%\SagaTemps\20285_IRIS13_20210402_001_PPS_0000_000080_Z.sgrd ^
-POS %WORK%\SagaTemps\20285_IRIS13_20210402_001_PPS_0000_000080_PO.sgrd ^
-RADIUS 0.1 ^
-METHOD 1 ^
-NDIRS 8

```

Abbildung 136: Beispiel für den Kommandozeilenaufwurf, der das Topographic Openness Tool ausführt.

4.3.3 Black Top Hat Transformation

In der Bildanalyse wird die Black Top Hat Transformation (BTHT) genutzt, um dunkle Bildbereiche von einem variablen Hintergrund zu separieren (SOILLE 2004). Der Algorithmus basiert auf dem Prinzip der morphologischen Filterung, die mit einem sogenannten Strukturelement (SE) arbeitet (JÄHNE 2012). Das SE ist dabei ein Kernel, der über die Eingabematrix bewegt wird (Moving-Window-Prinzip). Beim morphologischen Closing wird mittels des SE zunächst eine Dilatation und im Anschluss eine Erosion ausgeführt. Bei der Dilatation wird der größte und bei der Erosion der kleinste Wert in einer Nachbarschaft für das Ergebnis übernommen. Diese Nachbarschaft wird durch die Größe des SE definiert.

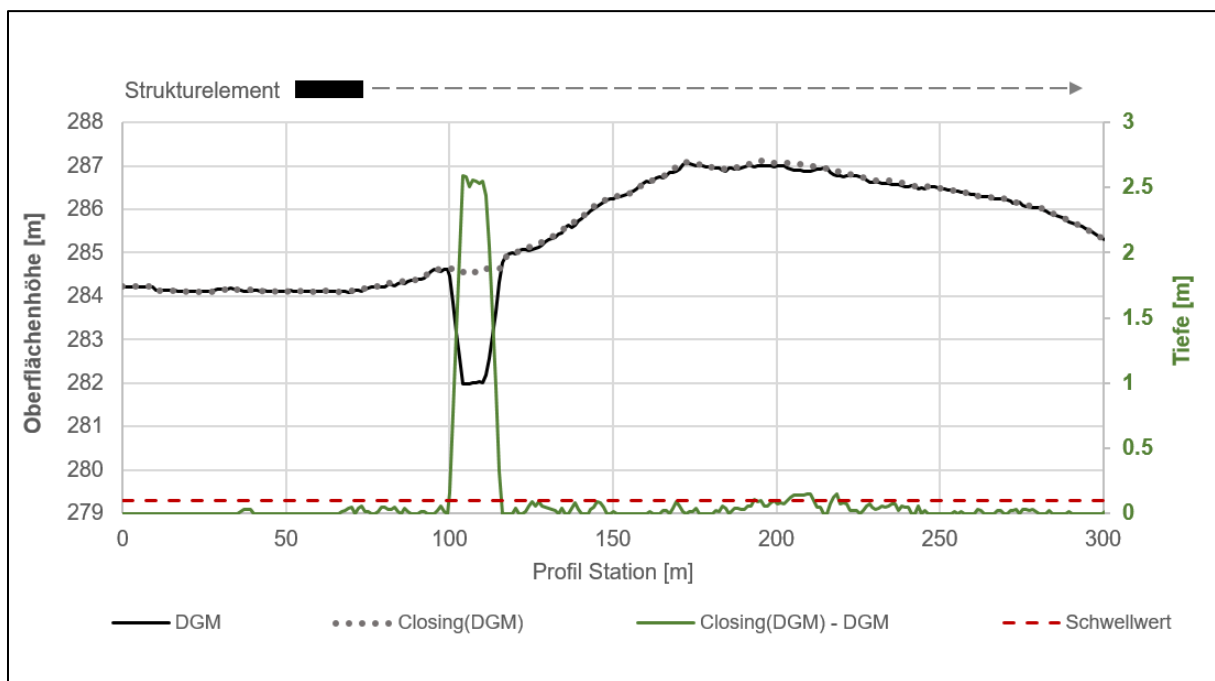


Abbildung 137: Funktionsweise der Black Top Hat Transformation am Beispiel eines 2D-Höhenprofils. Die Tiefe der Hohlform ist auf der sekundären Y-Achse abgetragen (grün).

Angewendet auf ein DGM erlaubt es die BTHT beispielsweise, die Eintiefung von Tälern in gebirgigen Topographien zu approximieren (RODRIGUEZ et al. 2002; LUO et al. 2015). Die White Top Hat Transformation hingegen wird zur Extraktion von Gipfeln angewendet und nutzt anstelle der Closing-Operation ein morphologisches Opening.

Wie in Abbildung 137 anhand eines Profilschnitts dargestellt, wird auf das DGM eine Closing-Operation unter Nutzung eines SE mit definierter Größe angewendet. Im Ergebnis ist die Eintiefung zwischen Profilstation 100 m und 110 m nahezu entfernt. Wenn nun die rasterzellenweise Differenz zwischen originalem DGM und Closing-DGM berechnet wird, entsteht ein Raster, in dem die durch das Closing entfernten Hohlformen und auch kleinere Oberflächenvariationen entsprechend ihrer Tiefe repräsentiert sind. Mithilfe eines Schwellenwertes lässt sich steuern, welche Tiefen letztlich berücksichtigt werden sollen.

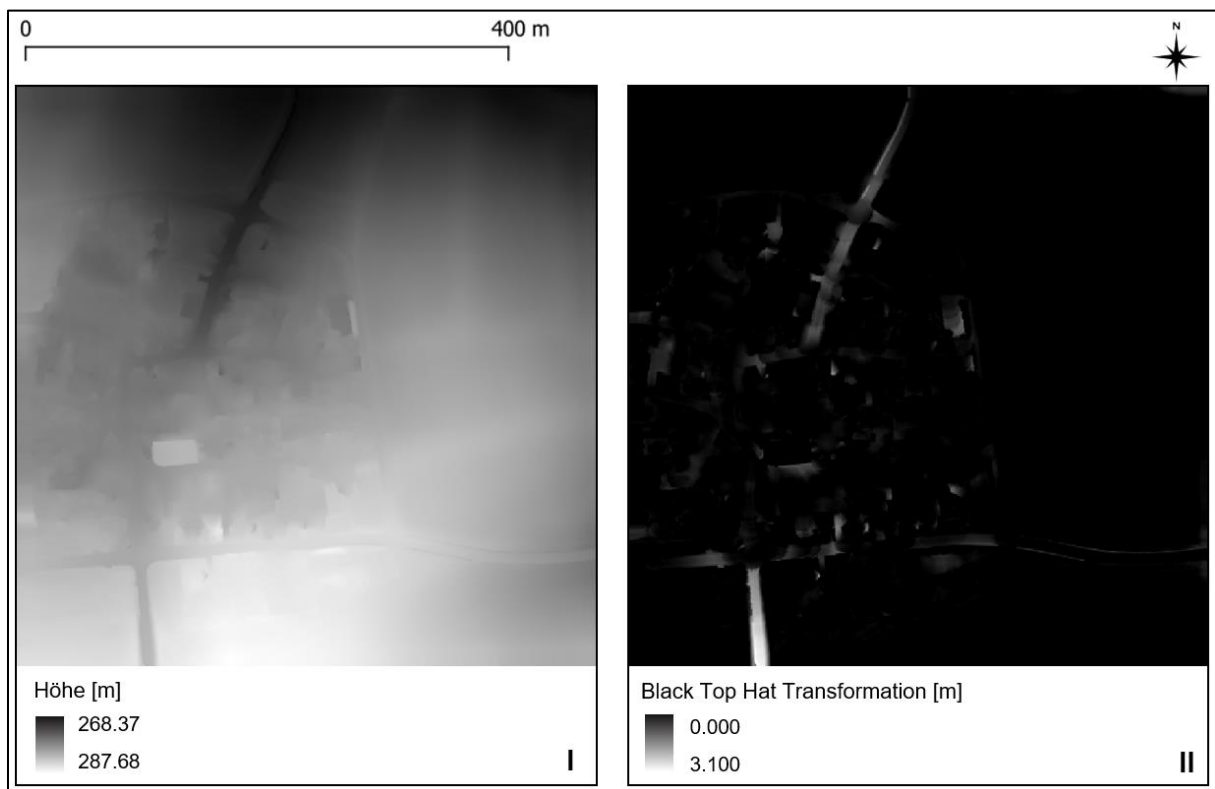


Abbildung 138: BTHT berechnet mit einem Strukturelementradius von 10 m und einem Schwellenwert von 0,1 m (II) für ein Airborne-LiDAR-DGM (I) (Datengrundlage: BAYERISCHE VERMESSUNGSVERWALTUNG 2022a).

Abbildung 138 zeigt, dass nur die negativen Oberflächenformen des DGM im BTHT-Raster erhalten bleiben. Die visualisierten Formen lassen sich scharf an ihren Kanten abgrenzen (keine Ausrundungseffekte an Kanten). Die BTHT-Werte erlauben außerdem nicht nur eine Lokalisierung von Oberflächenvariationen, sondern auch eine Quantifizierung ihrer Tiefe.

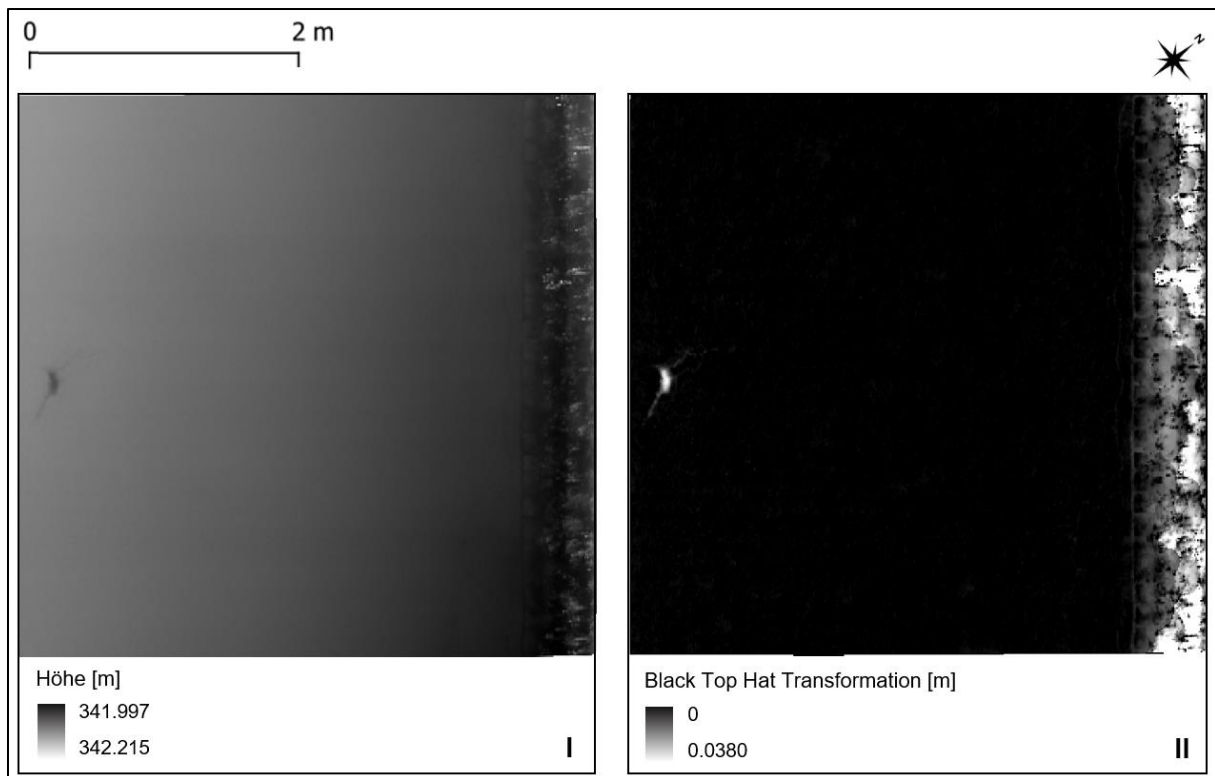


Abbildung 139: BTHT mit Strukturelement-Radius 0,1 m und Schwellenwert 0,0 m für eine Straßenoberfläche mit Rissen und Schlaglöchern (II). Teilabbildung I zeigt das zugrundeliegende Oberflächenmodell.

Abbildung 139 zeigt rechts ein BTHT-Raster, welches auf Basis eines PPS-Oberflächenmodells (links) berechnet wurde. Während der Ausbruch und die davon abstrahlenden Risse deutlich zu erkennen sind, heben sich die feineren Risse rechts entlang der gepflasterten Rinne visuell kaum von der Asphaltoberfläche ab. Abbildung 140 zeigt den Kommandozeilenaufwurf für SAGA GIS, der auf Basis eines Eingabeoberflächenmodells ein BTHT-Raster erzeugt. Die Größe des SE wird in Meter angegeben und auf den Wert 0,1 m eingestellt. Der Schwellenwert wird ebenfalls in Meter angegeben und auf 0,0 m gesetzt.

```
REM Tool: Valley and Ridge Detection (Top Hat Approach)
saga_cmd ta_morphometry 24 ^
-DEM %WORK%\SagaTemps\20285_IRIS13_20210402_001_PPS_0000_000080_Z.sgrd ^
-VALLEY %WORK%\SagaTemps\20285_IRIS13_20210402_001_PPS_0000_000080_TH.sgrd ^
-RADIUS_VALLEY 0.1 ^
-THRESHOLD 0.0
```

Abbildung 140: Beispiel für den Kommandozeilenaufwurf, der das Top Hat Approach Tool ausführt.

4.3.4 Topographic Position Index

Der Topographic Position Index (TPI) ist ein von WEISS (2001) im Rahmen einer ESRI-Anwenderkonferenz vorgestelltes Maß für die relative Differenz zwischen der Höhe einer Rasterzelle in einem DGM und der mittleren Höhe der umgebenden Zellen. Die Größe der lokalen Zellennachbarschaft wird über ein Kernel-Fenster definiert. Die Berechnung des lokalen Mittelwerts im TPI fungiert als Tiefpassfilter, der die lokale Morphologie glättet und lediglich regionale Merkmale der Oberfläche in der Ausgabe erhält. Im Kern stellt der TPI also eine Hochpassfilterung der Höheninformation dar, die in ähnlicher Weise auch in anderen Ansätzen umgesetzt ist. So trennen beispielsweise HILLER & SMITH (2008) in dem von ihnen verwendeten, Residual Relief Separation genannten Ansatz die hochfrequente von der niedrigfrequenten Höheninformation, indem sie auf eine lokale Nachbarschaft einen Median-Filter anwenden und das Ergebnis vom ursprünglichen Höhenmodell subtrahieren. Auch der bei MAJCHER, PLETS & QUINN (2020) verwendete Focal-Statistics-Ansatz, bei dem die Standardabweichung in einer lokalen Nachbarschaft berechnet wird, liefert Ergebnisse, die mit denen des TPI vergleichbar sind.

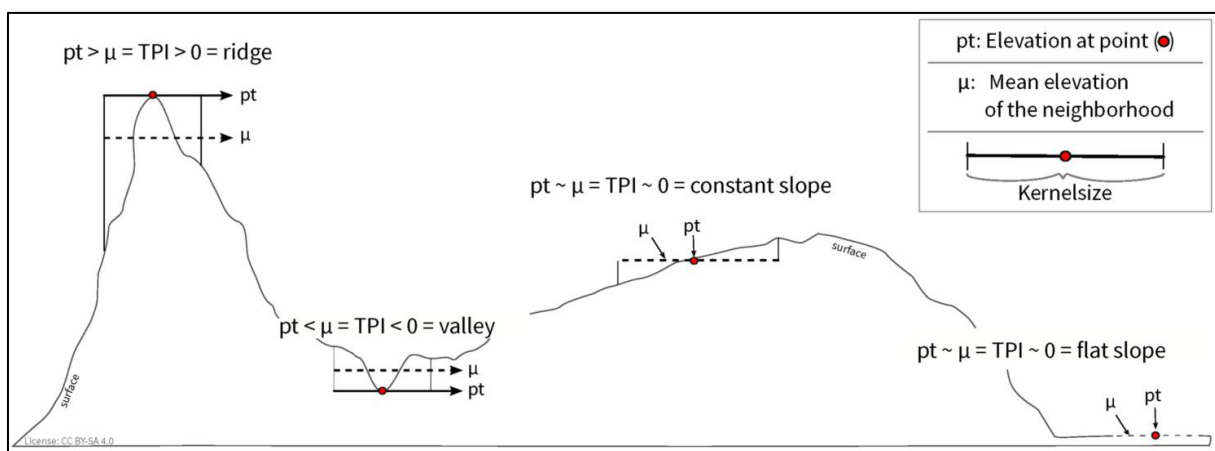


Abbildung 141: Prinzip des Topographic Position Index (KNITTER et al. 2019).

In Abbildung 141 ist das Funktionsprinzip des TPI dargestellt. Während positive TPI-Werte anzeigen, dass sich die betrachtete Oberfläche über ihre lokale Nachbarschaft erhebt, weisen negative TPI-Werte auf tal- oder muldenartige Oberflächenstrukturen hin. Ebenen und konstante Steigungen äußern sich in TPI-Werten um 0.

Das klassische Anwendungsgebiet des TPI ist die sogenannte Landform Classification, bei der der TPI unter Berücksichtigung einer kleinen und einer großen lokalen Nachbarschaft auf ein Höhenmodell angewendet wird und die Ausgaben regelbasiert unter Zuhilfenahme eines Neigungsrasters zehn verschiedenen Zielklassen, wie z. B. Tal, Ebene, Hang oder Kuppe,

zugeordnet werden (WEISS 2001; TAGIL & JENNESS 2008; MOKARRAM & SATHYAMOORTHY 2016). Darüber hinaus findet der TPI oder vergleichbare Hochpassfilterungen unter anderem Verwendung in der fluvialen Geomorphologie, um Flussläufe zu kartieren (NOTEBAERT et al. 2009) oder in der Archäologie, um beispielsweise Siedlungsspuren zu extrahieren (KNITTER et al. 2019).

Abbildung 142 zeigt, wie die TPI-Werte eine Interpretation der Geländemorphologie ermöglichen. Dabei ist von Vorteil, dass klar zwischen Strukturen unterschieden werden kann, die sich oberhalb und unterhalb des mittleren Umgebungsniveaus befinden. Ähnlich wie bei der BTHT kann von den TPI-Werten auf die Tiefe bzw. die Höhe einer Oberflächenvariation geschlossen werden. Im direkten Vergleich mit der BTHT können beim TPI steile Kanten, wie im Bereich der ins Gelände eingeschnittenen Straße unten links, nicht so exakt lokalisiert werden, da hier ein Überschwingen der Werte vom negativen in den positiven Bereich auftritt (Ergebnis der lokalen Mittelung der Höhe). Dies visuell korrekt zu interpretieren erfordert Kenntnis über das Berechnungsprinzip und etwas Übung.

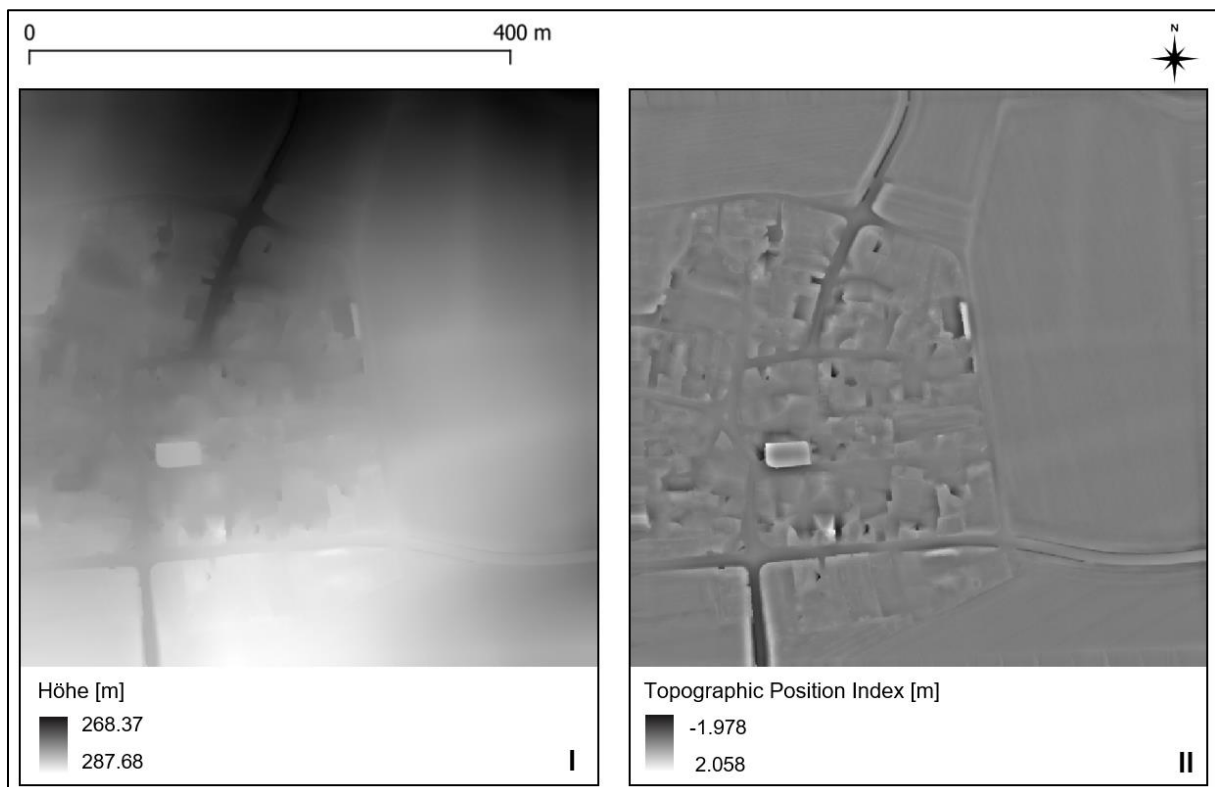


Abbildung 142: TPI berechnet mit einem Radius von 10 m (II) für ein Airborne-LiDAR-DGM (I) (Datengrundlage: BAYERISCHE VERMESSUNGSVERWALTUNG 2022a).

Wie in Abbildung 143 zu sehen ist, können mithilfe des TPI sowohl Risse als auch Ausbrüche anschaulich visualisiert werden. Risse zeichnen sich als dunkle Linien und kleinere Ausbrüche

(Durchmesser $d < 0,2$ m) als dunkle Flächen ab. Größere Ausbrüche sind durch eine unregelmäßige Grenze in Verbindung mit einem orthogonal orientierten Helligkeitsverlauf von dunkel zu hell gekennzeichnet. Objekte, die auf der Oberfläche liegen (z. B. Laub) erscheinen hellgrau bis weiß und leichte Senken dagegen dunkelgrau. Intakte bzw. flache Oberflächenbereiche sind in einem homogenen Grauton dargestellt und lassen sich visuell gut von geschädigten Bereichen oder auch Nicht-Fahrbahnbereichen abgrenzen (z. B. Bankett am rechten Rand).

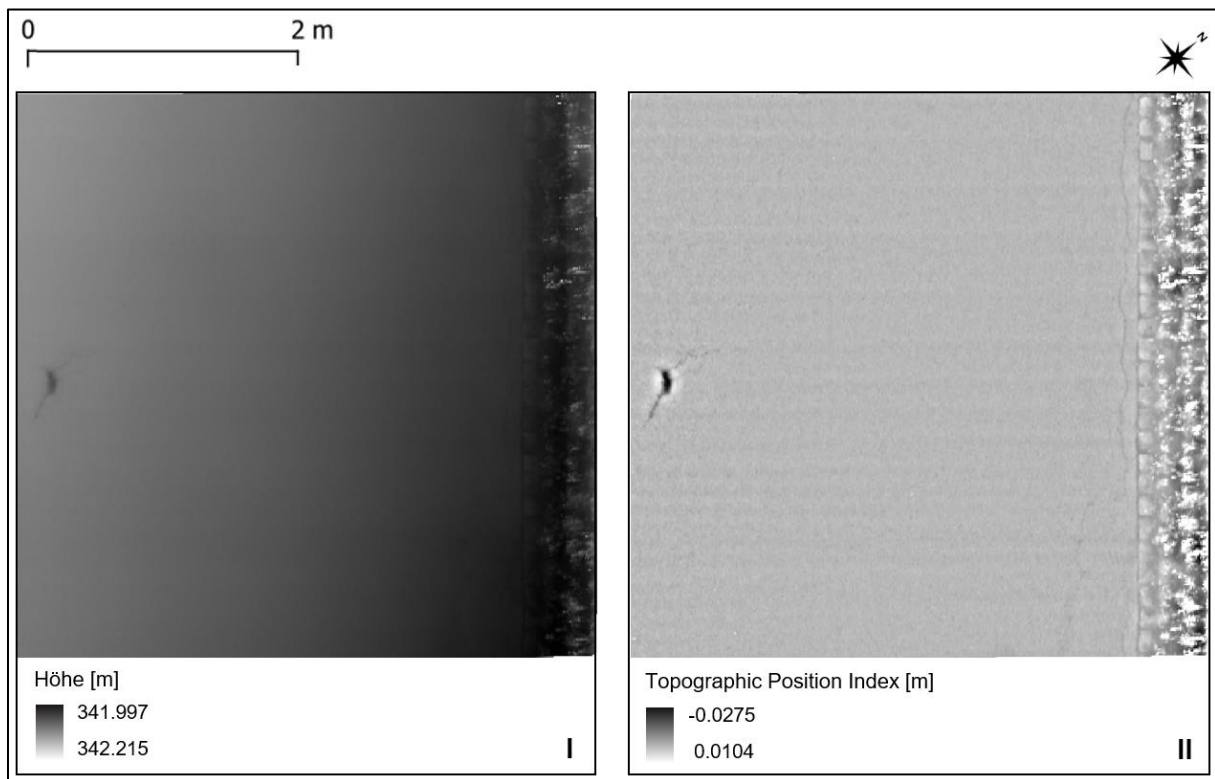


Abbildung 143: TPI mit Radius 0,1 m für eine Straßenoberfläche mit Rissen und Schlaglöchern (II). Teilabbildung I zeigt das zugrundeliegende Oberflächenmodell.

In der vorliegenden Arbeit wird der TPI-Algorithmus genutzt, welcher in SAGA GIS als Tool „Topographic Position Index (TPI)“ in der Library „ta_morphometry“ implementiert ist. Das Tool bietet die Möglichkeit, die lokale Nachbarschaft über einen inneren und äußeren Radius in Metern zu definieren. Ferner kann parametrisiert werden, ob eine Standardisierung der Ausgabe erfolgen und ob eine Distanzgewichtung innerhalb der definierten Nachbarschaft angewendet werden soll. Für die Untersuchungen im Rahmen dieser Arbeit wurde die TPI-Berechnung so parametrisiert, dass die lokale Nachbarschaft durch einen Kreis mit Radius $r = 0,1$ m gegeben ist, keine Standardisierung und keine Distanzgewichtung stattfindet (siehe Abbildung 144).

```

REM Tool: Topographic Position Index (TPI)
saga_cmd ta_morphometry 18 ^
-DEM %WORK%\SagaTemps\20285_IRIS13_20210402_001_PPS_0000_000080_Z.sgrd ^
-TPI %WORK%\SagaTemps\20285_IRIS13_20210402_001_PPS_0000_000080_TPI.sgrd ^
-STANDARD=0 ^
-RADIUS_MIN=0.0 ^
-RADIUS_MAX=0.1

```

Abbildung 144: Beispiel für den Kommandozeilenaufwurf, der das Topographic Position Index Tool ausführt.

In SAGA GIS wird der zu verwendende Kernel nicht in Pixeln definiert, sondern die Kernelgröße und -form ergibt sich aus der Pixelauflösung des Eingaberasters und den eingestellten Radien. In Abbildung 145 sind beispielhaft zwei Kernelformen dargestellt. Im Rahmen der vorliegenden Arbeit findet die rechte Option Anwendung, indem der innere Radius auf 0,0 m und der äußere Radius auf 0,1 m gesetzt wird.

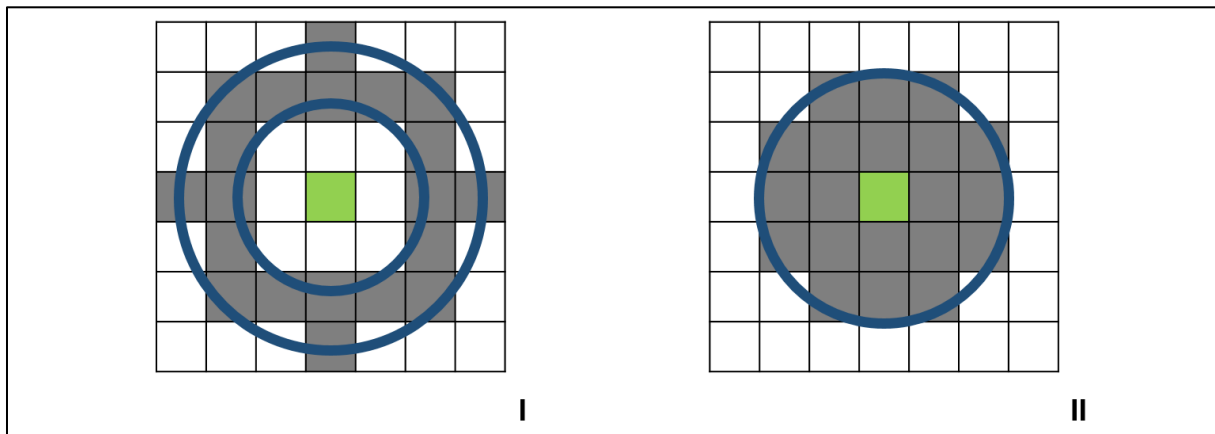


Abbildung 145: Beispiel für Kernel-Formen, die für die TPI-Berechnung verwendet werden können. Teilabbildung I zeigt einen 7x7 Pixel großen Kernel mit einem äußeren Radius von $r = 3$ und einem inneren Radius von $r = 2$. Teilabbildung II zeigt einen 7x7 Pixel großen Kernel mit einem äußeren Radius von $r = 2$ und einem inneren Radius von $r = 0$.

4.4 Das DeepLab System

In den folgenden Abschnitten werden die wesentlichen Architekturmerkmale des Deep-Learning-Ansatzes erklärt, der die semantische Segmentierung von PPS-Oberflächenmodellen im Rahmen der vorliegenden Arbeit realisiert.

Aufbauend auf der Grundidee des im Kapitel 2.4.2.3 beschriebenen Fully Convolutional Networks wurde von Entwicklern des Google Research Departments ein semantisches Segmentierungssystem namens DeepLab entwickelt und in mehreren Optimierungsstufen publiziert (CHEN et al. 2018a, 2018b). Der Quellcode der TensorFlow-Implementierung von DeepLab ist

über Github öffentlich zugänglich und zur Nutzung unter Apache License 2.0 lizenziert (CHEN et al. 2021).

Die wesentlichen Entwicklungsschritte und Innovationen von DeepLab sind in der ersten Version die sogenannte Atrous Convolution zur dichten Merkmalsextraktion und Sichtfeldvergrößerung sowie die Nutzung von Fully Connected Conditional Random Fields zur Verbesserung der Lokalisierungsgenauigkeit durch die Berücksichtigung von Kontextinformationen. In der zweiten Architekturversion wurde das Konzept der Atrous Convolution im Atrous Spatial Pyramid Pooling Modul so weiterentwickelt, dass im Zuge der Bildsegmentierung Kontextinformationen auf mehreren Betrachtungsskalen effizient berücksichtigt werden können. In der dritten DeepLab-Version ersetzen spezielle Encoder- und Decoder-Module das Fully-Connected-Conditional-Random-Fields-Modul der Vorgängerarchitekturen und ermöglichen die Optimierung der Segmentierungsergebnisse im Sinne schärferer Objektgrenzen. Nachfolgend werden diese wesentlichen Schritte zur Entwicklung von DeepLabv3+ im Detail erläutert.

4.4.1 DeepLabv1 und DeepLabv2

Wie vorstehend beschrieben, griffen CHEN et al. (2018a) bestehende Ansätze zur Bildsegmentierung auf und erweiterten sie, um eine wesentliche Herausforderung der semantischen Segmentierung mit tiefen künstlichen neuronalen Faltungsnetzen zu adressieren: die Wiederherstellung detaillierter lokaler Strukturen im Bild, die aufgrund der schrittweisen Verkleinerung der Feature Maps im CNN verloren gehen. Der innovative Kern der ersten DeepLab-Generation liegt in der Kopplung der Erkennungsleistung tiefer neuronaler Faltungsnetze mit der Lokalisierungsgenauigkeit von Fully Connected Conditional Random Fields. Damit war es den Autoren möglich, im Kontext der semantischen Bildsegmentierung eine Wiederherstellung von Objektgrenzen auf einem bis dato unerreichten Detailniveau und hinsichtlich der Rechenzeit wesentlich effizienter zu realisieren.

Wie am Beispiel von AlexNet, FCN und SegNet beschrieben (siehe Kapitel 2.4.2), reduziert das wiederholte Kombinieren von Max-Pooling und Striding in den aufeinanderfolgenden Schichten eines tiefen künstlichen neuronalen Faltungsnetzes die räumliche Auflösung der am Ende des Netzes befindlichen Feature Maps (auch: Score Map). Um diesem Umstand entgegenzuwirken und die Netzwerkausgabe auf die räumliche Auflösung des Eingabebildes rückabzubilden, existieren verschiedene Konzepte. Eines davon ist die transponierte Faltung (engl.: Transposed Convolution), welche es ermöglicht, den Upsampling-Prozess mit trainieren zu

können. Dieses vereinfachend auch als Deconvolution bezeichnete Konzept ist jedoch vergleichsweise speicher- und rechenzeitintensiv. DeepLab wendet mit der sogenannten dilatierten Faltung (engl.: Dilated Convolution) für das Upsampling einen alternativen Ansatz an, der eine Erweiterung des rezeptiven Feldes ohne Verlust an Auflösung oder Abdeckung realisiert. Da bei diesem Konzept die Anzahl von Parametern nicht verändert wird, entsteht auch kein erhöhter Rechenaufwand, was die dilatierte Faltung sehr effizient macht. Die DeepLab-Autoren bezeichnen ihre Implementierung der dilatierten Faltung als Atrous Convolution und beschreiben damit im Kern einen Faltungsprozess mit Filtern, die einem Upsampling unterzogen wurden. Das Upsampling der Filter wurde so realisiert, dass zwischen den einzelnen Filterabgriffen eines Kernels (engl.: Filter Taps) Nullen bzw. Lücken eingefügt sind. Mit dem Parameter Dilationsrate kann die Anzahl von Nullen zwischen den Filterabgriffen und somit die Größe des rezeptiven Feldes gesteuert werden. Abbildung 146 demonstriert schematisch die Merkmalsextraktion mit herkömmlichem Upsampling (siehe Abbildung 62) im Vergleich zur Merkmalsextraktion mittels Upsampling der Filter.

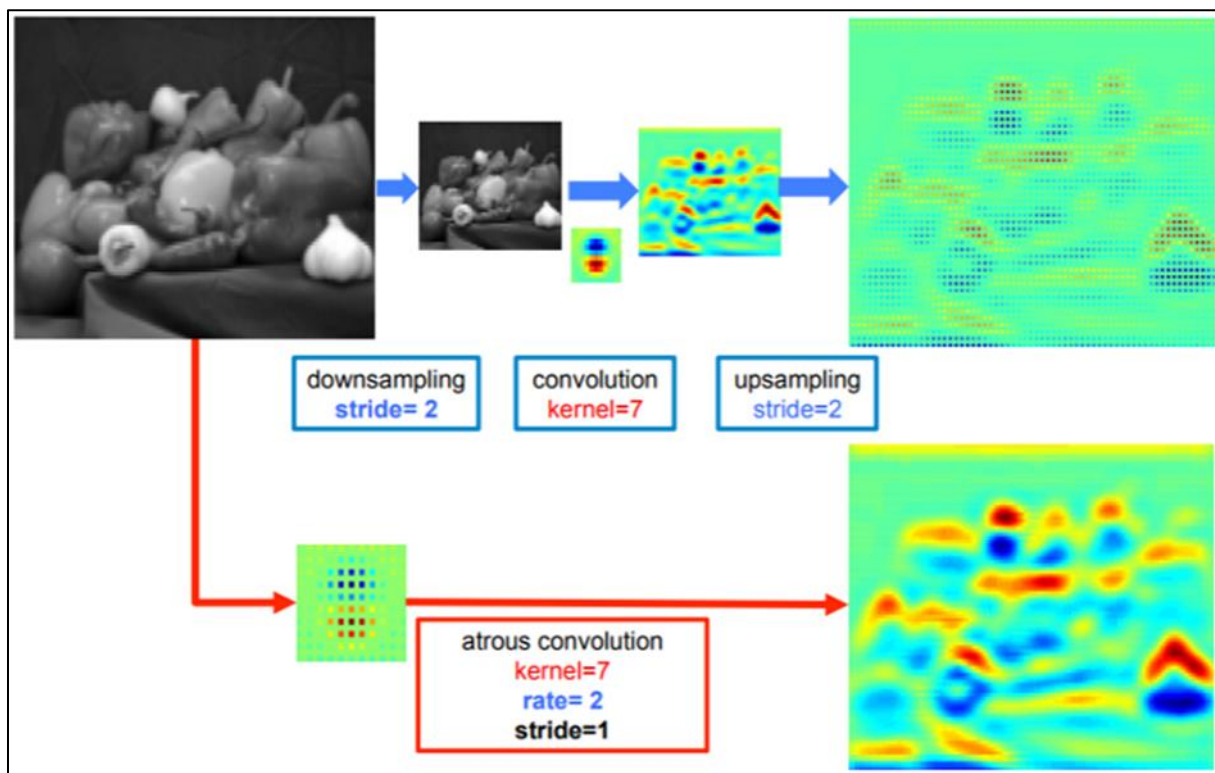


Abbildung 146: Illustration der Atrous Convolution. Obere Reihe: Merkmalsextraktion (sparse) mit Standardfaltung auf einer niedrig aufgelösten Input Feature Map. Untere Reihe: Merkmalsextraktion (dense) mittels Atrous Convolution mit Rate $r = 2$, angewandt auf eine hochauflöste Input Feature Map (CHEN et al. 2018a).

Um auch feine Details exakt erfassen zu können, verwenden die Autoren in der ersten DeepLab-Generation ein Fully Connected Conditional Random Field bzw. ein Dense Conditional

Random Field (CRF). CRFs zählen zu den statistischen Modellierungsmethoden, mit deren Hilfe bei der Prädiktion für einen bestimmten Bildpixel auch dessen Nachbarschaftsbeziehungen mit einbezogen werden können (LAFFERTY, MCCALLUM & PEREIRA 2001). So kann beispielsweise durch das Koppeln benachbarter Knoten des CRFs die Übereinstimmung der semantischen Annotationen zwischen räumlich nahe gelegenen Pixeln maximiert werden. Klassischerweise werden solche CRF-Terme zur Glättung verrauschter Segmentierungsergebnisse genutzt. Darüber hinaus ist auch die Modellierung komplexerer kontextueller Beziehungen möglich. Hierfür modelliert das CRF eine Wahrscheinlichkeitsverteilung basierend auf den Eingabepixeln und ihren kontextuellen Beziehungen. So kann es beispielsweise unwahrscheinlich sein, dass Pixel mit sehr unterschiedlichen Farbwerten derselben Ausgabeklasse angehören. In einem gesonderten Trainingsprozess kann das CRF beispielgetrieben trainiert werden, diese den Trainingsdaten inhärente Wahrscheinlichkeitsverteilung zu erlernen und für eine konturscharfe Abgrenzung von Objekten im Zuge der Bildsegmentierung zu nutzen. DeepLabv1 nutzt ein von KRÄHENBÜHL & KOLTUN (2011) vorgestelltes Fully Connected CRF.

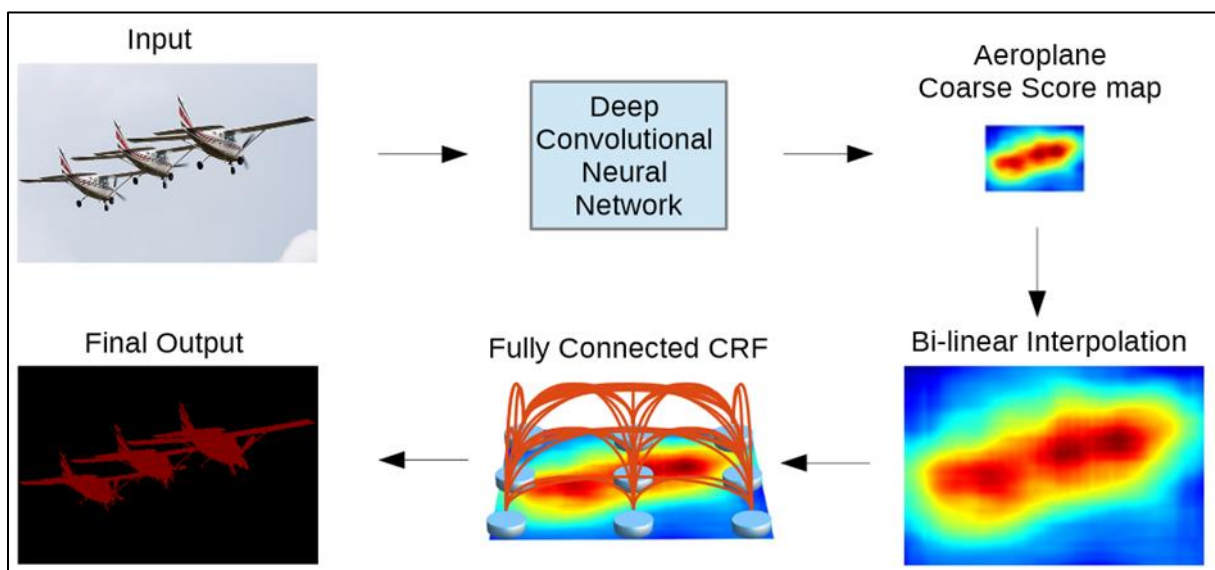


Abbildung 147: Funktionsprinzip eines DeepLabv1-Modells. Die grobe Score Map aus dem Deep Convolutional Neural Network wird durch bilineare Interpolation hochskaliert und ein Fully Connected Conditional Random Field verfeinert das Segmentierungsergebnis (nach CHEN et al. 2018a).

In Abbildung 147 ist dargestellt, wie die beiden vorstehend erläuterten Konzepte in der Verarbeitungskette des DeepLabv1-Modells implementiert sind. Zunächst durchläuft ein Eingabebild zur Merkmalsextraktion die Faltungs- und Poolingschichten eines konventionellen CNN (Backbone-Netz). Als Backbone wird eine VGG-16-Architektur genutzt, bei der die letzten Max-Pooling-Schichten jedoch entfernt und die nachfolgenden Faltungs- und Fully-

Connected-Schichten durch zwei Atrous-Convolution-Schichten mit den Dilationsraten $r = 2$ und $r = 4$ ersetzt werden. Dies führt zu Feature Maps, die mit einer höheren Abtastrate berechnet werden. Im Ergebnis entsteht eine Score Map, die zwar im Vergleich zum Eingabebild eine um den Faktor 8 verringerte Auflösung aufweist, die ohne die Verwendung der Atrous Convolution jedoch noch niedriger aufgelöst wäre (Faktor 32). Diese Feature Map wird im zweiten Teil des Netzes mittels bilinearer Interpolation auf die Größe des Eingabebildes zurückskaliert. Zur Verbesserung des Segmentierungsergebnisses wird abschließend ein Fully Connected CRF angewendet.

Mit den in der DeepLabv1-Architektur verwendeten Modulen konnten die Entwickler bereits grundlegende Problemstellungen der semantischen Bildsegmentierung auf einem hohen Leistungsniveau adressieren. Oftmals können die im Bild zu erkennenden Objekte aber nicht nur in verschiedenen Kontexten auftreten, sondern die relevanten Objekte oder Strukturen können darüber hinaus auch in verschiedenen Maßstäben abgebildet sein. So kann es beispielsweise aus der Aufnahmeperspektive des Bildes resultieren, dass Instanzen derselben Objektklasse unterschiedlich groß abgebildet werden. Genauso ist es möglich, dass Ausprägungen ein und derselben Objektklasse in unterschiedlichen Größen auch ohne perspektivische Abbildungseffekte existieren. Dieser Herausforderung widmete sich die Entwicklung der zweiten DeepLab-Generation. In ihrer Publikation präsentieren CHEN et al. (2018a) ein Modul, das sie Atrous Spatial Pyramid Pooling (ASPP) nennen. Die Idee besteht darin, mehrere Atrous Convolutions mit unterschiedlichen Abtastraten auf die jeweilige Input Feature Map anzuwenden und deren Ausgaben miteinander zu verrechnen, um Objekte und Bildkontexte auf mehreren Skalenniveaus erfassen zu können. Abbildung 148 illustriert schematisch, wie Filter mit unterschiedlichen Abtastraten verschiedene effektive Sichtfelder repräsentieren.

Die DeepLabv2-Architektur unterscheidet sich von der ersten Generation im Wesentlichen durch die bevorzugte Verwendung eines ResNet-Backbone-Netzes anstelle des VGG-Netzes und die entsprechende Integration des ASPP-Moduls. Letzteres ersetzt in der DeepLabv2-Architektur die Fully-Connected-Schicht des Backbone-Netzes und ist aus vier Atrous Convolutions mit den Raten $r = 6$, $r = 12$, $r = 18$ und $r = 24$ aufgebaut.

Ferner verwendet DeepLabv2 einen überarbeiteten Multiskalenansatz, indem der Encoder-Teil dreimal parallel über skalierte Bilder mit dem Faktor 1,0, 0,75 und 0,5 ausgeführt wird. Das Maximum der Ausgaben wird dann als Basis für den Decoder-Teil verwendet.

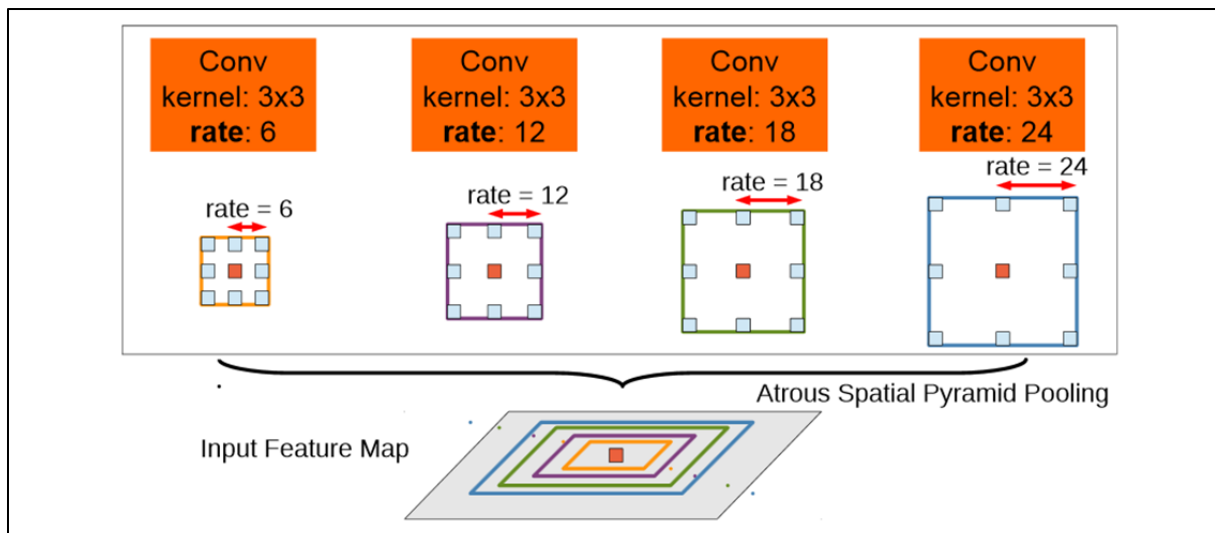


Abbildung 148: Prinzipskizze des Atrous Spatial Pyramid Pooling (ASPP). Um das zentrale Pixel (orange) zu klassifizieren, nutzt ASPP Multiskalenmerkmale, indem es mehrere parallele Filter-Kernels mit unterschiedlichen Raten verwendet. Die Raten definieren die Größe der Lücken zwischen den Filterabgriffen (CHEN et al. 2018a).

4.4.2 DeepLabv3 und DeepLabv3+

Zu DeepLabv3 existiert lediglich eine Pre-Print-Veröffentlichung (CHEN et al. 2017). Darin beschreiben die Autoren, welche Konzepte sie erprobt haben, um die Leistungsfähigkeit der DeepLab-Vorgängergenerationen weiter zu verbessern. Im Unterschied zu den beiden ersten DeepLab-Generationen sind CRF ab Version 3 kein Bestandteil des Decoders mehr. Hier wird bilineares Upsampling genutzt, um Segmentierungsergebnisse entsprechend der ursprünglichen Bildgröße zu erhalten. Ferner wurde das ASPP-Modul um die Möglichkeit der Batch Normalization erweitert, was dem Training robusterer Modelle zuträglich ist.

Mit der Veröffentlichung der Version DeepLabv3+ weisen die Autoren darauf hin, dass die Erfassung von Kontextinformationen, wie sie in DeepLabv3 mit dem ASPP-Modul realisiert ist, bei der Verwendung von sehr tiefen Backbone-Netzen, wie dem ResNet101, sehr rechenaufwändig werden kann (CHEN et al. 2018b). Auch das bilineare Upsampling sei nicht optimal im Sinne einer Erhaltung bzw. Wiederherstellung von Objektgrenzendetails. Unter anderem deshalb wurden für DeepLabv3+ einige Konzepte von anderen Ansätzen adaptiert. So wurden beispielsweise die bisher verwendeten Faltungen durch sogenannte Depthwise Separable Convolutions ersetzt, um die Recheneffizienz zu erhöhen (CHOLLET 2017). Abbildung 149 zeigt in den Teilabbildungen (a) und (b), wie bei der Depthwise Separable Convolution für jeden Eingabekanal ein eigener 3 x 3 Filterkernel angewendet und deren Ausgabe mittels Pointwise Convolution anschließend über die einzelnen Kanäle hinweg kombiniert wird. In der Teilabbildung (c) ist dargestellt, wie dieses Konzept in DeepLabv3+ adaptiert wurde, um die

ursprüngliche Atrous Convolution zu einer Atrous Separable Convolution zu erweitern. Diese findet nun Verwendung sowohl im ASPP-Modul als auch im Decoder-Modul.

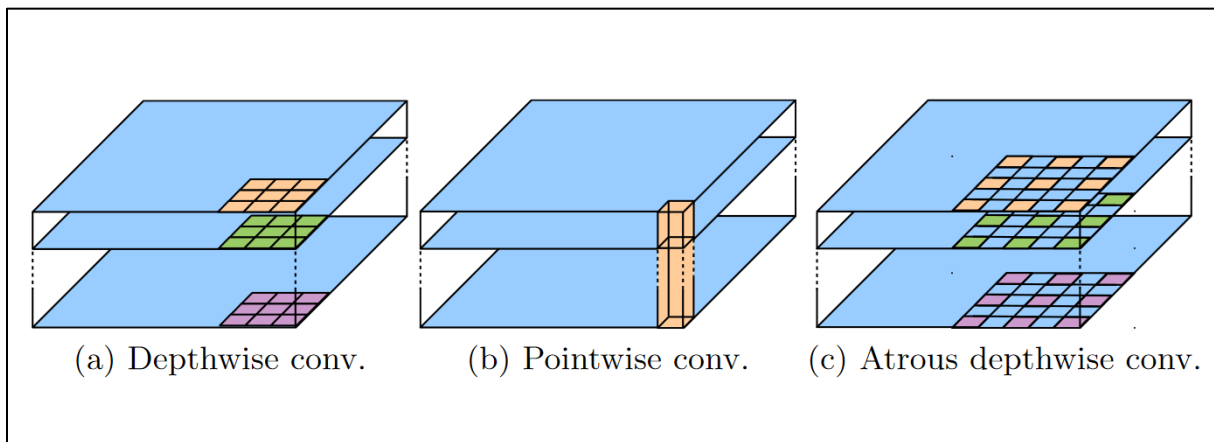


Abbildung 149: Integration der Konzepte Depthwise Separable Convolution und Pointwise Convolution der Xception-Architektur (CHOLLET 2017) in die in DeepLabv3+ verwendete Atrous Separable Convolution in DeepLabv3+ (CHEN et al. 2018b).

Der Decoder-Teil, der in DeepLabv3 noch aus einem bilinearen Upsampling mit dem Faktor 16 bestand, wurde in DeepLabv3+ wie folgt verändert (siehe Abbildung 150): Die Encoder-Ausgaben werden zunächst einem bilinearen Upsampling (Faktor 4) unterzogen und dann mit den entsprechenden Low-Level-Merkmalen aus den ersten Schichten des Backbone-Netzes, welche die gleiche räumliche Auflösung besitzen, verkettet. Dieses Vorgehen, bestimmte Encoder-Schichten mit den entsprechenden Decoder-Schichten zu kombinieren, ist inspiriert von der U-Net-Architektur (RONNEBERGER, FISCHER & BROX 2015). Da die betreffenden Merkmale im Encoder typischerweise bereits mehrere hundert Kanäle umfassen, werden sie mit einer 1×1 Faltung verarbeitet, wodurch sich das Training des Netzwerks deutlich vereinfacht. Um die Ausgaben zu verfeinern, werden im Anschluss an die beschriebene Verkettung verschiedene 3×3 Faltungen angewendet. Danach folgt ein erneutes Upsampling mit dem Faktor 4.

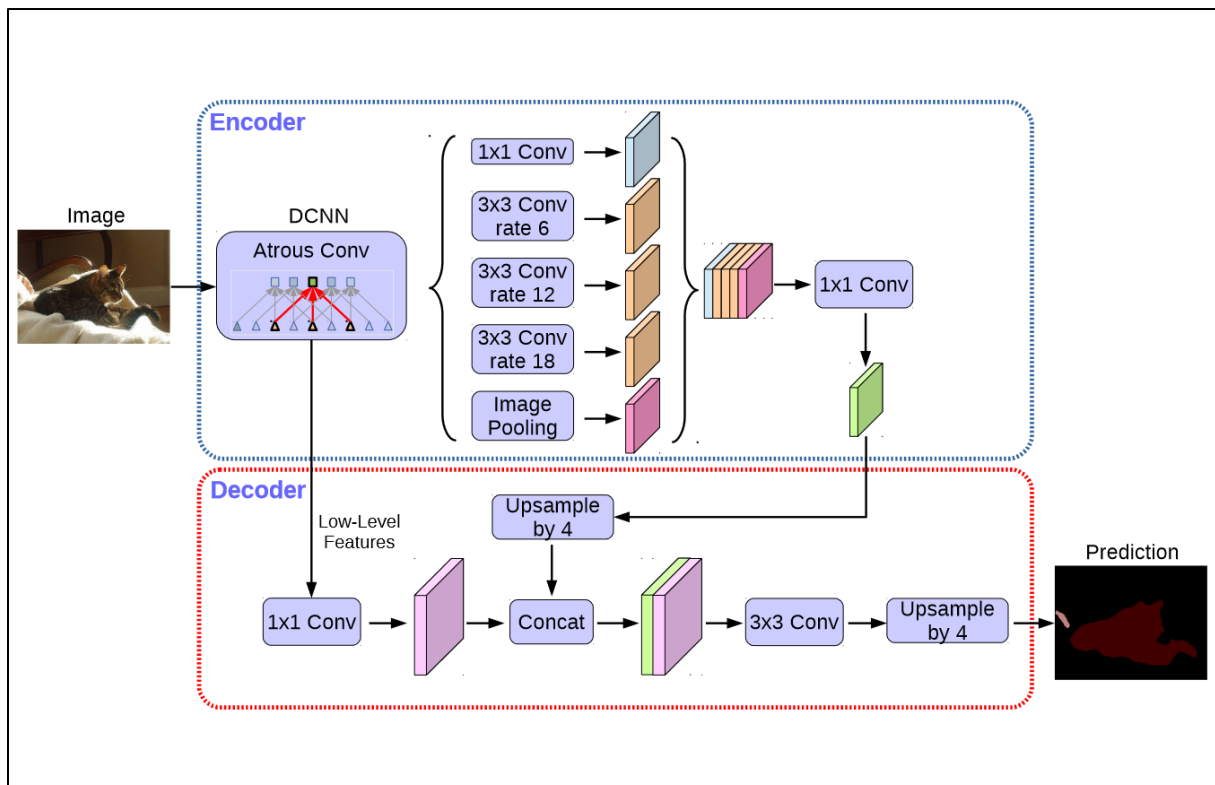


Abbildung 150: Schematische Darstellung der Encoder/Decoder-Struktur von DeepLabv3+ (CHEN et al. 2018b).

Neben den Anpassungen des Decoders wurde auch das bisherige ResNet101 Backbone CNN im Encoder durch ein modifiziertes Xception-Netz ersetzt, wodurch eine Beschleunigung der Rechenzeit erreicht werden konnte. Die Modifikation liegt in erster Linie darin, dass Xception ursprünglich für Fragestellungen der Bildklassifizierung eingesetzt wurde, im DeepLab-System jedoch eine semantische Segmentierung der Eingabedaten produziert. Zu diesem Zweck wurden unter anderem alle Max-Pooling-Operationen durch Atrous-Separable-Convolution-Operationen ersetzt, sowie zusätzliche Batch-Normalization-Operationen im Anschluss an Faltungen ergänzt.

Abbildung 151 zeigt beispielhaft Segmentierungsergebnisse, die mit DeepLabv3+ Modellen erreicht werden können. Der Segmentierungsmaske in der oberen Reihe liegt ein Modell zugrunde, welches auf ImageNet- (DENG et al. 2009) und COCO-Daten (LIN et al. 2014) trainiert wurde. Das Modell, welches die Segmentierung in der unteren Reihe erzeugt hat, war zusätzlich auf den Cityscapes-Datensatz (CORDTS et al. 2016) trainiert.



Abbildung 151: Segmentierungsergebnisse vortrainierter DeepLabv3+ Modelle. Die Teilabbildungen I und III zeigen je ein Eingabebild. Die Segmentierung oben rechts (II) wurde mit einem auf ImageNet- und COCO-Daten vortrainierten Modell erzeugt (nach CHEN et al. 2018b). Die Segmentierung unten rechts (IV) wurde mit einem auf CityScapes-Daten vortrainierten Modell erzeugt.

4.5 Evaluation der Ergebnisse

Nachfolgend wird beschrieben, welche Methoden zur Evaluation der Untersuchungen in Kapitel 5 genutzt werden.

4.5.1 Nutzwertanalyse

Die Nutzwertanalyse ist ein Werkzeug, welches in den Wirtschaftswissenschaften Anwendung findet und für Entscheidungsprozesse sowie das Finden von Alternativen bei komplexen Problemstellungen eingesetzt werden kann (EISENFÜHR, WEBER & LANGER 2010). Der Ablauf einer Nutzwertanalyse lässt sich vereinfacht wie folgt zusammenfassen: Am Anfang einer Nutzwertanalyse steht immer die Formulierung eines Ziels, welches das zu lösende Problem klar umreißt. Anschließend werden Kriterien definiert, die einen Einfluss auf die Entscheidungsfindung haben. Diese werden zunächst einzeln bewertet und anschließend hinsichtlich ihrer Relevanz bzw. vorherrschenden Präferenzen gewichtet. Aus den so resultierenden

Teilnutzwerten kann schließlich der Gesamtnutzwert einer Wahlmöglichkeit berechnet werden. Auch wenn Nutzwertanalysen eine gewisse Objektivität suggerieren, ist zu beachten, dass durch die Auswahl und Gewichtung der Kriterien ein subjektiver Einfluss im Rahmen der Entscheidungsfindung erhalten bleibt (HOFFMEISTER 2000). Das der Nutzwertanalyse zugrundeliegende Konzept wird hier dennoch adaptiert, da mithilfe dieser Vorgehensweise die Entscheidungsfindung für eine räumliche Interpolationsmethode transparent und nachvollziehbar dargelegt werden kann.

Das Ziel der Nutzwertanalyse in dieser Arbeit wurde bereits an anderer Stelle genannt, soll hier aber der Vollständigkeit halber noch einmal formuliert werden: Ziel ist es, eine räumliche Interpolationsmethode zu identifizieren, mit deren Hilfe Oberflächenmodelle aus den Laserscans des Fraunhofer PPS schnell und mit hoher Präzision erzeugt werden können. FRANKE (1982) schlägt verschiedene Eigenschaften vor, die zur Beurteilung der Qualität und Eignung eines Interpolators herangezogen werden können. Dabei handelt es sich zum einen um objektive Kriterien, wie Genauigkeit, Parametersensitivität und Rechenzeit und zum anderen vergleichsweise schwer quantifizierbare Eigenschaften wie visuelle Aspekte oder die Komplexität der Implementierung. Im Hinblick auf die Problemstellung dieser Arbeit wurden zur Bewertung der verglichenen Interpolationsmethoden folgende Kriterien definiert: Da es sich bei der Oberflächenmodellierung im Kontext der avisierten Endanwendung um einen Massendatenprozess handelt, der bereits im Falle weniger Straßennetzkilometer tausendfach ausgeführt werden muss, sollte der Parametrisierungsaufwand gering sein. Der zur Bewertung herangezogene Indikator ist die Anzahl einzustellender Parameter. Aus demselben Grund sollten Oberflächenmodelle möglichst schnell erzeugt werden können. Das verwendete Verfahren sollte eine möglichst geringe Laufzeitkomplexität haben. Der zur Bewertung herangezogene Indikator ist die Laufzeit der Interpolation. Schließlich sollte das Interpolationsverfahren die Messgenauigkeit des Laserscanners erhalten und dementsprechend exakte Oberflächen erzeugen. Der zur Bewertung herangezogene Indikator ist die vertikale Genauigkeit. Hierfür wird mit dem Root Mean Square Error (RMSE) eine Standardmetrik aus dem Bereich der Qualitätsanalyse von digitalen Geländemodellen verwendet (WISE 2011; SZYPUŁA 2016).

4.5.2 Kreuzvalidierung

Um die erzeugten Oberflächenmodelle in Hinblick auf ihre Qualität im Sinne der vertikalen Genauigkeit evaluieren zu können, wird die sogenannte Kreuzvalidierung angewandt. Sie zählt

zu den etablierten Verfahren im Kontext der Qualitätssicherung von Geländemodellierungen (WISE 2011). Im Zuge der Kreuzvalidierung wird die Gesamtheit der verfügbaren Datenpunkte aufgeteilt. Ein Teil der Daten steht tatsächlich für die Oberflächenmodellierung zur Verfügung. Der restliche Teil dient ausschließlich zur Evaluierung der Modellierung, indem die Höhenwerte der bekannten Datenpunkte mit denen der entsprechenden Modellierung verglichen werden. Die verbleibenden Residuen quantifizieren die vertikale Genauigkeit der Modellierung. Eine hohe Güte der Oberflächenmodellierung liegt dann vor, wenn die Differenzen der gemessenen Höhenwerte gering sind, verglichen mit den Werten der Modellierung. In Abbildung 152 ist die zufällige Extraktion von Kontrollpunkten im Zuge der Kreuzvalidierung beispielhaft visualisiert.

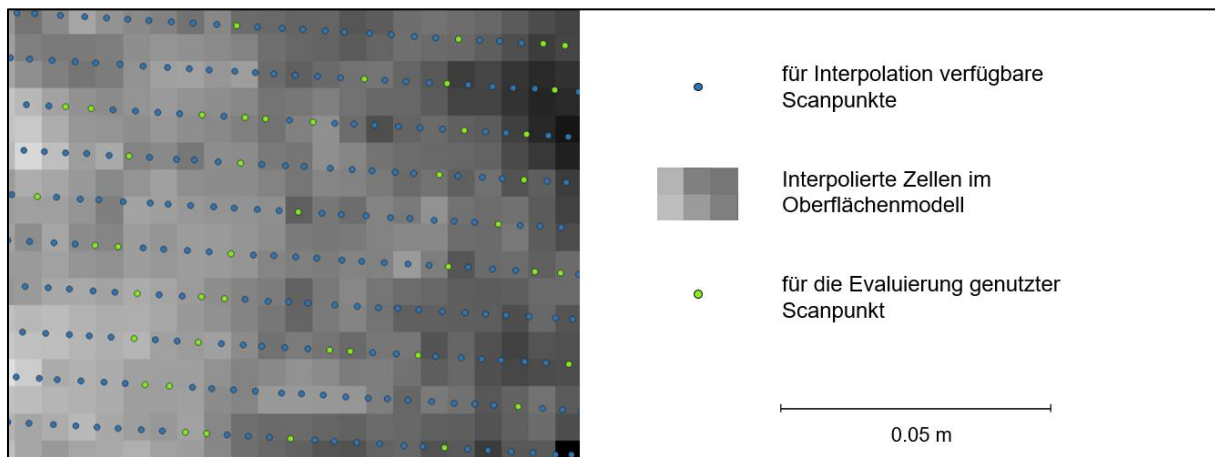


Abbildung 152: Beispiel für die mittels Kreuzvalidierung realisierte Evaluierung der Interpolationsgüte. Im Vordergrund sind die quer zur Fahrtrichtung ausgerichteten Laserscanprofile dargestellt. Die blauen Punkte repräsentieren jene Messpunkte, die zur Interpolation der Oberflächenmodelle verwendet werden. Die grünen Punkte repräsentieren die Punkte, an denen die Interpolationsgüte evaluiert wird. Aus der Datengesamtheit der Messpunkte wurden zufällig 20 % als Kontrollpunkte selektiert. Im Hintergrund befindet sich das am Pixelkoordinatensystem ausgerichtete Raster des interpolierten Oberflächenmodells.

4.5.3 Evaluation von semantischen Segmentierungsmodellen

Um die Güte eines Klassifikationsverfahrens zu messen bzw. verschiedene Verfahren oder Modelle miteinander zu vergleichen, bedarf es einer systematischen Vorgehensweise (FAHRMEIR et al. 2016). Eine Grundvoraussetzung hierfür ist, dass ein Teil der manuell klassifizierten Daten nicht in den Trainingsprozess mit eingeht, sondern ausschließlich für die Evaluierung vorgehalten wird (Hold-Out Splitting). Wie in Kapitel 2.4.1.3 beschrieben, wird die verfügbare Datenbasis in mindestens zwei Teile unterteilt: einen Trainingsdatensatz, dessen Beispieldaten für das Training von Modellen verwendet werden und einen Testdatensatz, dessen Beispieldaten nur zur Evaluierung von trainierten Modellen dienen (CHOLLET 2021). In

der vorliegenden Arbeit wird auf die zusätzliche Separierung eines Validierungsdatensatzes verzichtet. Das Parametersetting bzw. die Modellauswahl erfolgte im Rahmen von Voruntersuchungen am Testdatensatz.

Zur Beurteilung seiner Leistungsfähigkeit wird ein trainiertes Modell auf den Testdatensatz angewendet und die Modellprädiktion mit den manuell hinterlegten Klasseninformationen verglichen. Dabei können prinzipiell vier Fälle eintreten, die sich am Beispiel eines binären Klassifikationsproblems anschaulich erklären lassen, grundsätzlich aber auch bei Mehrklassenproblemen gelten. True Positives (TP) und True Negatives (TN) beschreiben im Fall eines Zweiklassenproblems (Klasse A vs. Klasse B) die Anzahl der Datenpunkte, die vom Modell jeweils korrekt der Klasse A oder B zugeordnet wurden (Treffer). Fehlklassifikationen werden unterschieden in Typ 1 und Typ 2 Fehler: False Positives (FP) sind dabei die Anzahl der Datenpunkte aus Klasse B, die fälschlicherweise vom Modell als Klasse A vorhergesagt wurden. Das Modell erkennt also zu viel als Klasse A (Fehlaktivierung). False Negatives (FN) beschreiben die Anzahl der Datenpunkte aus Klasse A, die fälschlicherweise vom Modell als Klasse B vorhergesagt wurden. Das Modell erkennt also zu wenig als Klasse A (Versäumnis).

		Vorhersage A	Vorhersage B
tatsächlich	A	True Positive	False Negative (Typ 2 Fehler)
	B	False Positive (Typ 1 Fehler)	True Negativ

Abbildung 153: Prinzip der Beurteilung eines Klassifikators mithilfe einer Konfusionsmatrix im Fall einer binären Klassifikation.

Diese beschriebenen Fälle können in einer Konfusionsmatrix für binäre Klassifikationen, aber auch für Mehrklassenprobleme dargestellt werden (siehe Abbildung 153). Sind vom Modell mehrere Klassen zu prädictieren, dann umfasst die Konfusionsmatrix entsprechend der Klassenanzahl mehrere Spalten und Zeilen. Bei einem perfekten Klassifikator wäre nur die Hauptdiagonale von oben links nach unten rechts besetzt, weil dann nur Treffer vorhergesagt würden.

Aus den Einträgen einer Konfusionsmatrix können verschiedene Metriken berechnet werden. So beschreibt die Präzision (engl.: Precision) das Verhältnis von True Positives zur Gesamtzahl aller positiven Entscheidungen. Die Sensitivität (engl.: Recall) hingegen quantifiziert das Verhältnis von True Positives zur Gesamtzahl der tatsächlich positiven Beobachtungen. Ein Überblick über gängige Bewertungsmaße im Bereich der semantischen Bildsegmentierung ist bei CSURKA, LARLUS & PERRONNIN (2013), FERNANDEZ-MORAL et al. (2018) und WANG, WANG & ZHU (2020) zu finden. In der vorliegenden Arbeit werden die Metriken F-Score (F1), Matthews Correlation Coefficient (MCC) und Intersection over Union (IoU) verwendet.

Die Metrik F1 kombiniert Precision und Recall als harmonisches Mittel in einem einzigen Wert, dessen Wertebereich von 0 bis 1 reicht. Je höher der erzielte Wert, desto besser die Klassifikationsleistung. Die Berechnung ist der nachstehenden Formel zu entnehmen (FERNANDEZ-MORAL et al. 2018).

$$F_1 = \frac{2TP}{2TP + FN + FP} \quad (18)$$

Die de facto Standardmetrik im Bereich der Objekterkennung und semantischen Bildsegmentierung ist die IoU (auch: Jaccard-Index). Dieser Index gibt das Verhältnis der Größe der Schnittmenge zur Größe der Vereinigungsmenge von Referenz und Prädiktion an und beschreibt damit die Ähnlichkeit zwischen der Modellausgabe und der Ground Truth. Der Wertebereich der IoU-Metrik reicht von 0 bis 1, wobei 0 für keine Überlappung und 1 für eine perfekt überlappende Segmentierung steht. Sowohl für Zweiklassen- als auch Mehrklassenprobleme wird der IoU für jede Klasse individuell berechnet und abschließend gemittelt als mIoU angegeben. Die Berechnung ist der nachstehenden Formel zu entnehmen (FERNANDEZ-MORAL et al. 2018).

$$IoU = \frac{TP}{TP + FN + FP} \quad (19)$$

Bei Klassifikationsaufgaben, wenn mehr als zwei Klassen unterschieden werden müssen, werden bei der F1- und IoU-Metrik die Typen Micro, Macro und Weighted Averaging unterschieden und am Beispiel von F1 nachstehend kurz erklärt: Beim Macro F1 wird zunächst je Klasse der mittlere Precision- und Recall-Wert berechnet. Jede Klasse hat damit in der F1-Berechnung das selbe Gewicht – ungleiche Klassenbesetzungen werden folglich nicht berücksichtigt. Beim Micro F1 werden erst TP, TN und FP über alle Klassen aufsummiert und

dann in die (18) eingesetzt. Damit entspricht der Micro F1 Score der globalen Genauigkeit (engl.: Accuracy), die ebenfalls sensitiv für eine ungleiche Klassenbesetzung im Datensatz ist. Bei der Weighted-Berechnungsvariante der F1-Metrik, wird zunächst der F1 Score für jede Klasse individuell berechnet und dann der globale Score über den Anteil gewichtet gemittelt, den jede Klasse am Umfang des gesamten Datensatzes hat. Gleiches gilt für die IoU-Metrik. Im Unterschied zu den beiden vorgenannten Bewertungsmaßen berücksichtigt die MCC-Metrik auch TN und damit alle vier möglichen Ergebnisausprägungen, die in der Konfusionsmatrix darstellbar sind. MCC wird im Allgemeinen als balancierte Bewertungsgröße angesehen, die auch dann verwendet werden kann, wenn die Besetzungshäufigkeit der Klassen sehr unterschiedlich ist. Der Wertebereich reicht von -1 bis +1, wobei +1 eine perfekte Klassifikation und 0 eine Zufallsvorhersage anzeigt. Bei einem MCC von -1 ist von einer inversen Vorhersage auszugehen. Die Berechnung ist der nachstehenden Formel zu entnehmen (CHICCO & JURMAN 2020).

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP) \cdot (TP + FN) \cdot (TN + FP) \cdot (TN + FN)}} \quad (20)$$

Wie vorstehend dargelegt, sollten der MCC oder die Weighted-F1- und Weighted-IoU-Werte als globale Metrik verwendet werden, um die Güte eines Klassifikators zu beschreiben, wenn im Testdatensatz eine ungleiche Klassenbesetzung (engl.: Class Imbalance) vorherrscht.

Der Vorteil einer Konfusionsmatrix gegenüber globalen oder klassenweisen Score-Werten ist, dass daraus hervorgeht, welche Klassen mit anderen Klassen verwechselt werden. Im Bereich der semantischen Bildsegmentierung ist es üblich, nur die globalen und klassenweisen Scores in tabellarischer Form aufzuführen. Für die Bewertung der trainierten Modelle in Kapitel 5.5 wird die IoU-Metrik als zentrales Bewertungsmaß verwendet.

4.5.4 Evaluation der automatischen Zustandserfassungsqualität

Wie in Kapitel 2.1.4 beschrieben, wird im Rahmen der ZEB die Erfassungsqualität von Oberflächenschäden über die sogenannte Eigenüberwachung gesichert. Dabei wird eine Erst- und Zweitauswertung desselben Datenmaterials durch zwei verschiedene Personen vorgenommen. Verglichen werden dann die Unterschiede in der Erfassung von Rissen und Flickstellen. Die im Regelwerk ZTV ZEB-StB (FGSV 2006) festgelegten Toleranzen definieren das erforderliche Qualitätsniveau der manuellen Zustandserfassung.

Die messtechnische Erfassung von Oberflächenschäden ist im kommunalen Straßennetz grundlegend an die Vorgehensweise angelehnt, die im Bundesfernstraßennetz Anwendung findet (siehe Kapitel 2.1.2.3 und 2.1.3.2). Deshalb wurde diese Erfassungssystematik für die Evaluierung der Qualität der automatischen Zustandserfassung auf Basis von MLS-Daten übernommen. Alle Sensordaten, die für die Evaluierungsstrecke aufgenommen wurden (siehe Kapitel 3.2.6), müssen deshalb zunächst in ein Rasterbildformat konvertiert werden, das mit einer TP3-Auswertesoftware regelwerkskonform verarbeitbar ist (siehe Kapitel 4.1.5). Standardmäßig werden im TP3-Prozess 10 m lange Bilder verwendet, die unabhängig von der Lageplangeometrie der Straße strikt gerade von unten nach oben im Bildkoordinatensystem verlaufen. Das TP3-Oberflächenkamarasystem von I.R.I.S 3 und das LCMS2-System von I.R.I.S 17 geben ihre Bilddaten per se in dieser Art und Weise aus, sodass hier keine zusätzliche Konvertierung nötig ist. Um jedoch ein entsprechendes Bildformat für die PPS-Laserscans zu erzeugen, muss die Trajektorie des Messfahrzeuges so manipuliert werden, dass der Heading-Winkel stets nach Norden gerichtet ist. So können 3D-Punktwolken und folglich mittels Oberflächenmodellierung auch Bilddaten erzeugt werden, mit denen eine TP3-Auswertesoftware arbeiten kann. Unter der Voraussetzung, dass alle drei Datenströme der Messsysteme (PPS, TP3 und LCMS2) denselben Start- und Endpunkt haben, kann ein und dasselbe Analyseraster einerseits auf Basis unterschiedlicher Auswertungen (manuell vs. automatisch) und andererseits auf Basis unterschiedlicher Sensordaten gefüllt und die Ergebnisse verglichen werden (siehe Abbildung 154).

Im Rahmen der vorliegenden Arbeit wird das Instrument der Eigenüberwachung also primär dazu genutzt, um zu überprüfen, ob die Leistungsfähigkeit einer Deep Learning basierten Zustandserfassung den Anforderungen der systematischen Straßenerhaltung genügen kann. Dazu wird eine manuelle Erstausswertung mit einer automatischen Zweitausswertung auf Basis derselben Sensordaten verglichen. Zusätzlich können dieser Analyse die Ergebnisse des nationalen und internationalen State-of-the-Art in einem einheitlichen Bezugsrahmen gegenübergestellt werden.

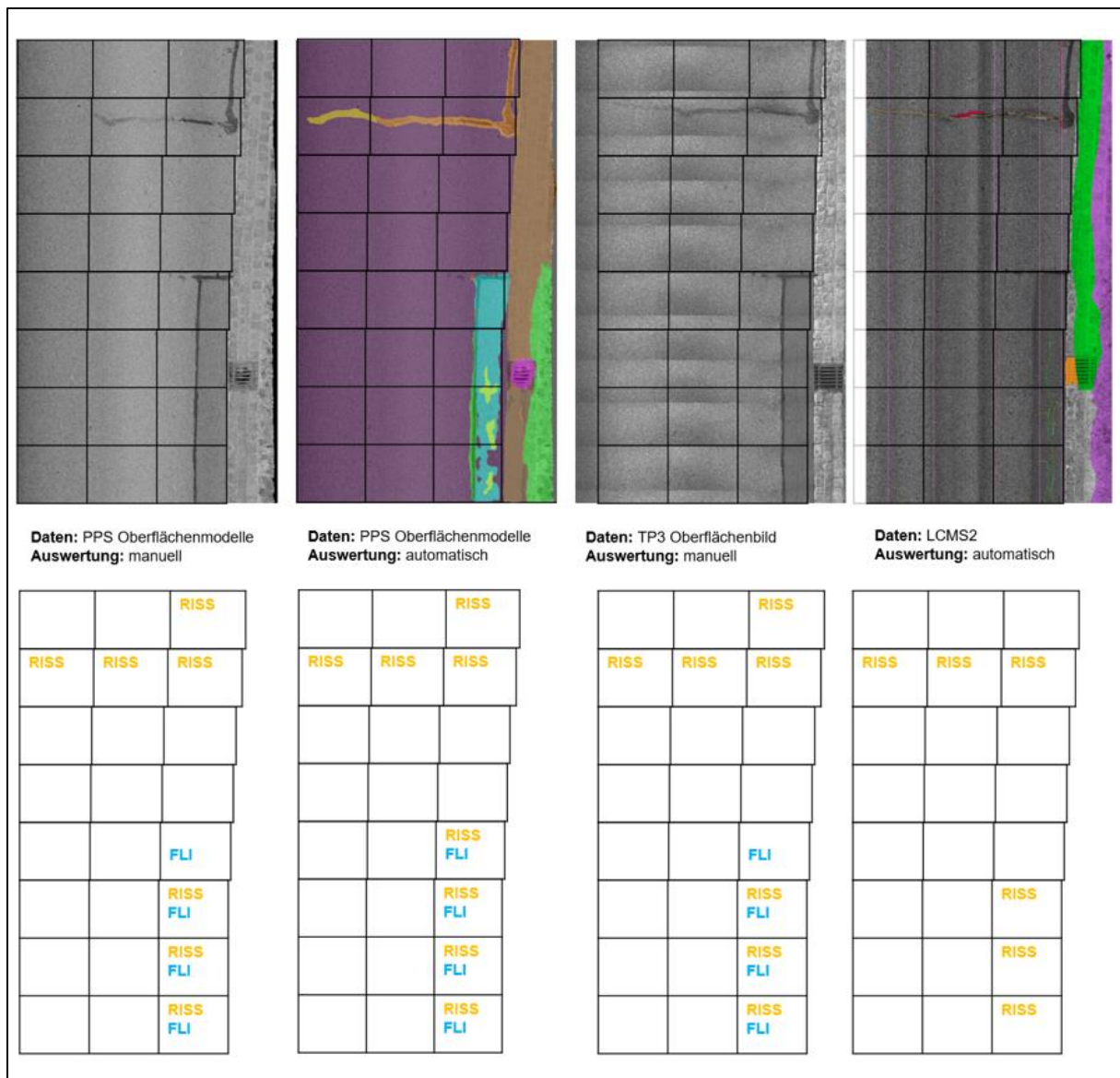


Abbildung 154: Prinzip der vergleichenden Evaluation mithilfe eines gemeinsamen Auswerterasters für die synchronisierten Datenströme der drei Aufnahmesysteme. Das Auswerteraster wurde manuell an die Situation im PPS-Datenstrom angepasst (Fahrstreifengrenzen) und auf die anderen Datenströme 1:1 übertragen. Für die PPS-Daten wurde das Auswerteraster sowohl manuell als auch automatisch gefüllt. Für die TP3-Oberflächenbilder wurde das Auswerteraster nur manuell ausgefüllt. Für die LCMS2-Daten wurde das Auswerteraster nur automatisch gefüllt.

5 Ergebnisse

In diesem Kapitel werden die praktischen Untersuchungen beschrieben und deren Ergebnisse in Bezug auf die zentralen Fragestellungen der Arbeit vorgestellt.

5.1 Absolute Positionierungsgenauigkeit

Die nachfolgend beschriebenen Untersuchungen beziehen sich auf die absolute Positionierungsgenauigkeit. Eine zentrale Fragestellung hierbei ist, wie genau Laserscans, die mit dem in Kapitel 3.1.1 beschriebenen MLS-System I.R.I.S erfasst wurden, verortet werden können. Damit ist auch eine Aussage über die absolute Lagegenauigkeit von aus den Laserscans extrahierten Schäden und Objekten möglich.

5.1.1 Analyse der Abweichung zwischen Passpunkten und Laserscan

Wie in Kapitel 3.2.1 beschrieben, wurden zur Analyse der absoluten Positionierungsgenauigkeit 51 markierte Passpunkte auf acht Kreisstraßen sowohl mittels handgeführtem GNSS-RTK als auch mit dem MLS-System I.R.I.S 13 aufgenommen.

Die Trajektorien der I.R.I.S 13 Messfahrten liegen in einem Gebiet, welches sich über ca. 50 km x 30 km erstreckt und sie umfassen ca. 38 km Länge. Mit der proprietären Software POSpac MMS von Applanix (siehe Kapitel 4.1.1) wurden die Realtime-Trajektorien in einem Post-Processing-Schritt optimiert. Hierbei wurden präzise Ephemeriden integriert und ein Netz aus acht DGNSS-Referenzstationen unter Nutzung des SmartBase-Ansatzes berechnet (siehe Abbildung 155). Die Genauigkeit der Trajektorien wird von der Software POSpac MMS als RMSE-Wert pro Trajektorienpunkt ausgegeben. In Anhang V sind statistische Kennzahlen dieser RMSE-Werte für die einzelnen Messfahrten aufgeführt. Insgesamt liegt die Lagegenauigkeit der Trajektorien in einem Bereich von 2 cm - 13 cm (im Mittel ca. 4 cm). Die Genauigkeit in der Höhe schwankt zwischen 5 cm und 13 cm (im Mittel ca. 7 cm). Unter Verwendung dieser Trajektorien wurden die 3D-Punktwolken des PPS erzeugt. Ihre absolute Positionsgenauigkeit ist maßgeblich determiniert von der Genauigkeit der Trajektorie.

Die terrestrischen Passpunktaufnahmen konnten direkt als 3D-Punktliste mit Metadaten, wie beispielsweise Punktcodes, Uhrzeit oder Qualitätsmaß, aus dem Leica CS 20 Controller exportiert werden. Wie in Anhang II dargestellt, sind diese Referenzmessungen in der Lage

auf 0,4 cm - 2,0 cm (im Mittel ca. 1,1 cm) und in der Höhe auf 0,7 cm - 3,7 cm (im Mittel ca. 1,6 cm) genau.

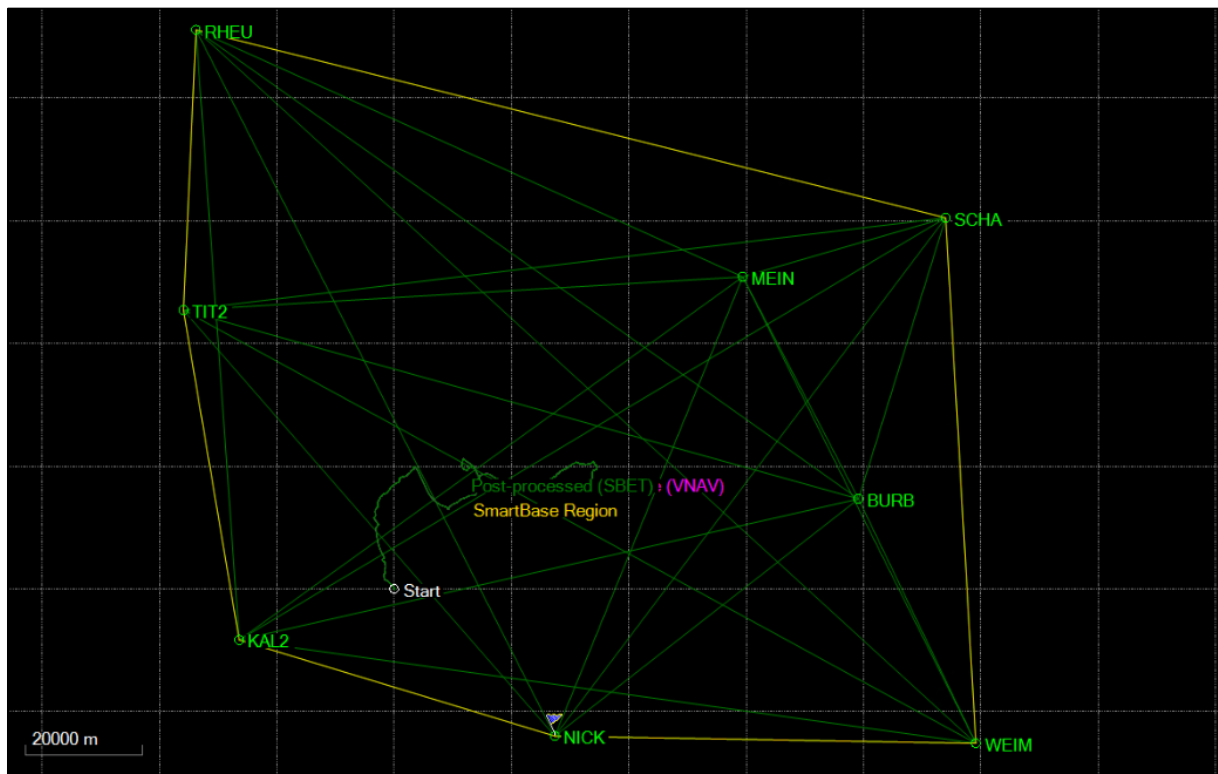


Abbildung 155: Post-Processing der Trajektorien mithilfe eines SmartBase-Netzes (gelb) aus acht DGNSS-Referenzstationen, die die Trajektorie (grün) räumlich einschließen.

Mithilfe der Software CloudCompare (siehe Kapitel 4.1.2) wurden die relevanten 3D-Punktwolken des PPS sowie die terrestrisch aufgenommenen Passpunktkoordinaten dreidimensional analysiert. Über ihre entsprechende Signatur in der Reflexionsintensität wurden die Passpunktmarkierungen identifiziert und aus den Laserscans ausgeschnitten. Um sicherzustellen, dass eine Messung am Mittelpunkt der Passpunktmarkierung auch dann möglich ist, wenn dieser zufällig nicht von einem Lasermesspunkt abgetastet wurde, erfolgte eine Vermaschung der freigestellten Scanpunkte mithilfe einer Delaunay-Triangulation zu TIN-Oberflächen. So konnten die Differenzen entlang der drei Raumachsen ermittelt werden, indem manuell vom Mittelpunkt der Passpunktmarkierung im TIN zum entsprechenden 3D-Passpunkt gemessen wurde (siehe Abbildung 156). Diese Vorgehensweise wurde für alle Passpunkte in gleicher Weise wiederholt.

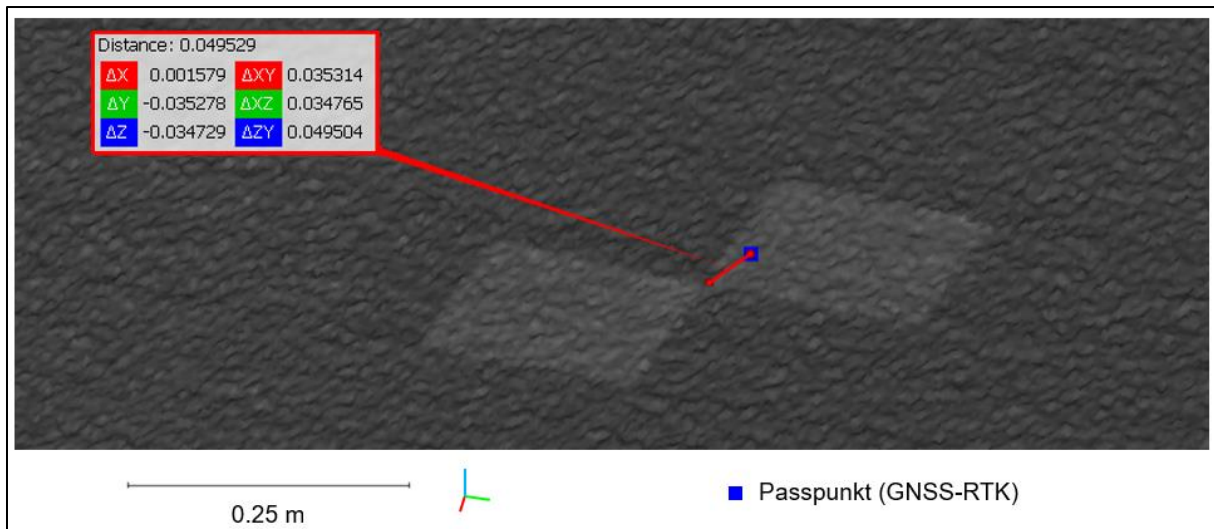


Abbildung 156: Beispiel für die Bestimmung des 3D-Abstands zwischen terrestrisch aufgenommenen Passpunktkoordinaten und den entsprechenden Koordinaten aus dem Laserscan. Die Lage- (ΔXY) und die Höhengenaugkeit (ΔZ) betragen im Beispiel ca. 3,5 cm.

5.1.2 Erkenntnisse zur absoluten Positionierungsgenauigkeit

Wie im vorherigen Abschnitt dargelegt, wurde jede einzelne GNSS-RTK-Passpunktmessung mit ihrem 3D-Abstand zur entsprechenden Passpunktmarkierung im Laserscan attribuiert. Das Ergebnis ist in Form eines Diagramms in Abbildung 157 dargestellt. Dort sind die Beträge der festgestellten Abweichungen aufgeteilt nach ihren X-, Y- und Z-Komponenten für die 51 Passpunkte abgetragen. Die Häufigkeitsverteilungen für die Abweichungen, getrennt nach den drei Raumachsen, sind in Anhang VI zu finden. Die Abweichungen zwischen den GNSS-RTK-Messungen und den PPS-Laserscans schwanken in der Lage in einem Bereich von 1 cm - 14 cm (im Mittel ca. 4 cm). In der Höhe haben die Schwankungen eine Spannweite von 0 cm - 18 cm (im Mittel ca. 5 cm). Es kann davon ausgegangen werden, dass Straßenschäden und Objekte, die im PPS-Laserscan erfasst sind, im Mittel auf weniger als 10 cm genau in Lage und Höhe georeferenziert sind.

Die Bereiche mit geringerer Genauigkeit liegen alle auf Außerortsstraßen und Ortsdurchfahrten, die durch Wälder und teilweise durch enge Mittelgebirgstäler des Westerwalds bzw. des Rothaargebirges verlaufen. Wie in Abbildung 158 für etwa ein Drittel des Messzeitraumes zu sehen ist, waren zeitweise weniger als vier GPS- bzw. GLONASS-Satelliten sichtbar, was bereits bei der Realtime-Trajektorie zu erhöhten RMSE-Werten führt. Nach der Trajektorienoptimierung ist der RMSE in diesen Bereichen zwar immer noch höher als in Bereichen mit optimalen Empfangsbedingungen, aber er ist gegenüber der Realtime-Trajektorie deutlich niedriger.

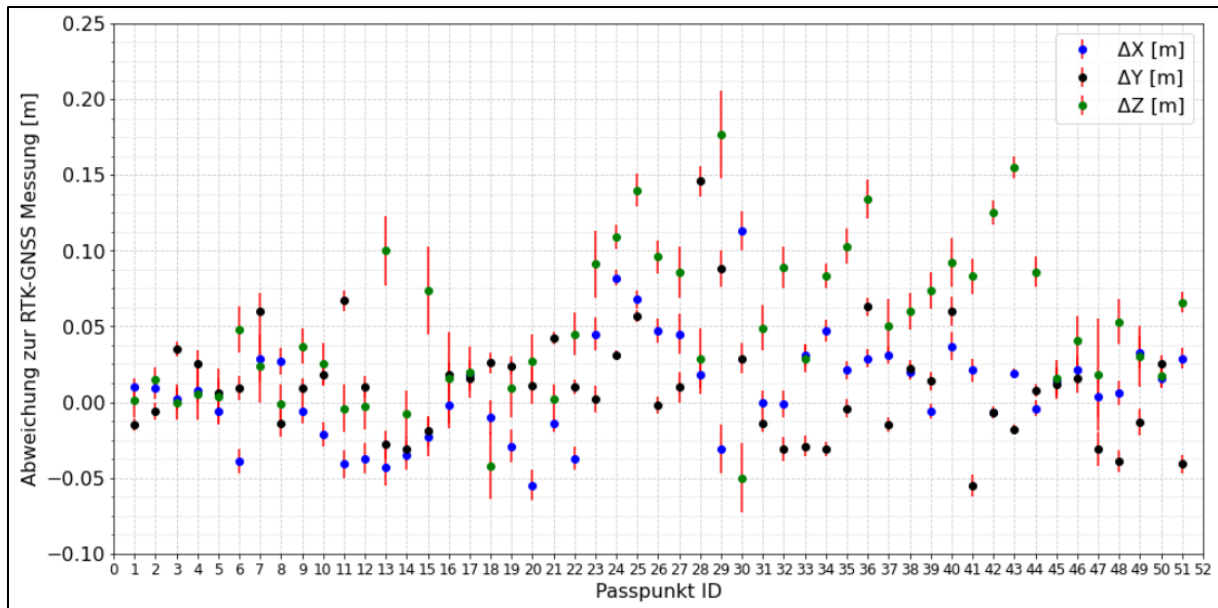


Abbildung 157: Darstellung der Ergebnisse zur Untersuchung der absoluten Positionierungsqualität der Laser-scans. Die roten vertikalen Linien repräsentieren die Qualität der GNSS-RTK-Messungen (Standardabweichung der jeweiligen Raumkomponente).

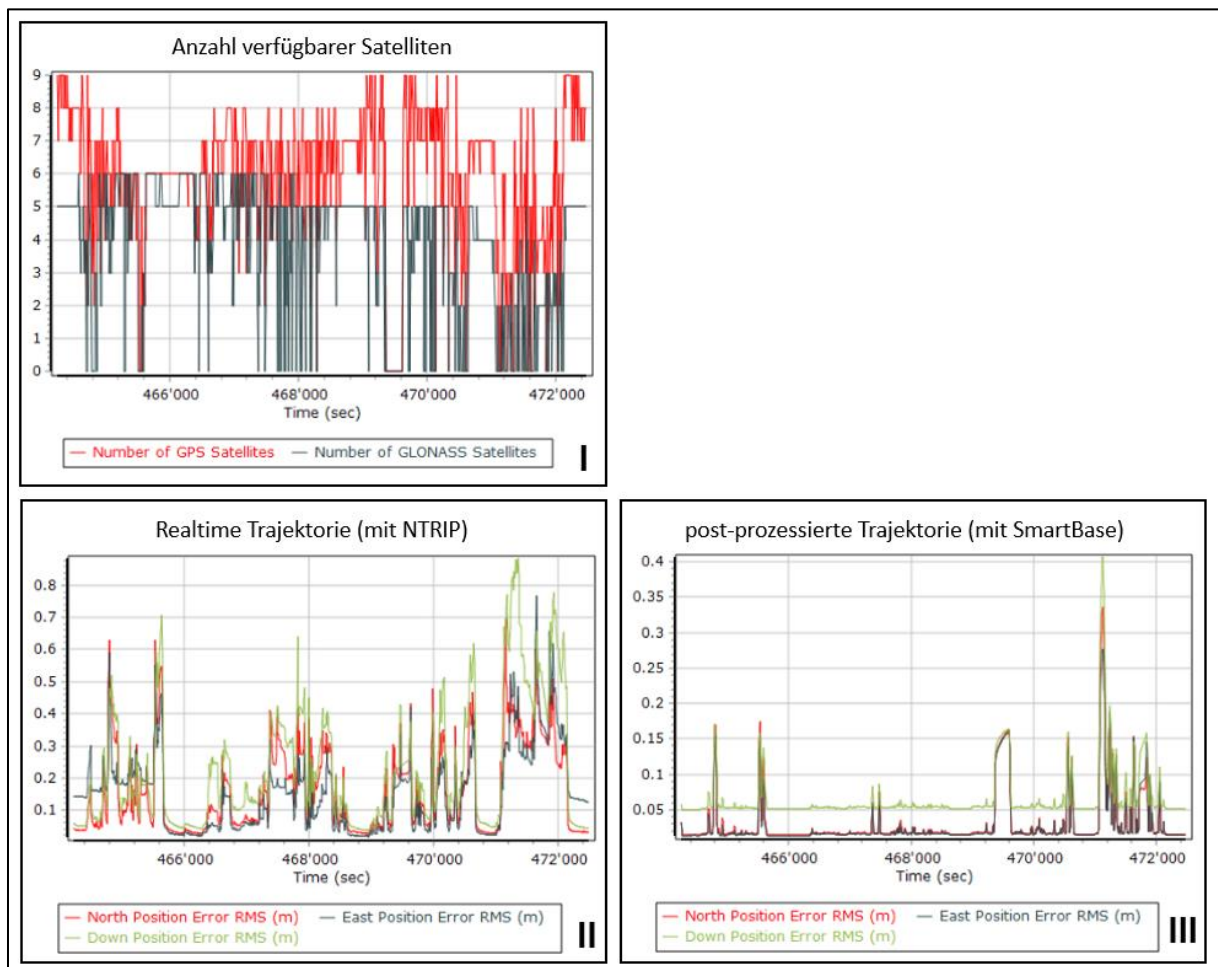


Abbildung 158: Einfluss der reduzierten Anzahl sichtbarer Satelliten (I) auf den Positionierungsfehler der Realtime-Trajektorie (II) und post-prozessierten Trajektorie (III). Die Ursache für die GNSS-Abschattung im dargestellten Zeitfenster ist der Straßenverlauf durch bewaldete, teilweise enge Täler.

5.2 Präzision des Messsystems

Der Begriff „Wiederholgenauigkeit“, der in der Normung durch den Begriff „Präzision“ abgelöst wurde, ist ein Kriterium, um die Qualität eines Messverfahrens zu beurteilen (DIN ISO 3534-1 2009). Die Präzision beschreibt demnach die statistische Abweichung eines Messwertes, die aus wiederholt durchgeführten Messungen resultiert und wird oft synonym mit dem Begriff der inneren Genauigkeit verwendet, welche die interne Streuung eines Messergebnisses bewertet (HENNES 2007). Sie wird in der Regel als Standardabweichung oder RMSE-Wert angegeben.

Im Rahmen dieses Kapitels wird dargelegt, mit welcher Präzision die Geometrie eines Messobjekts abgebildet werden kann, wenn es wiederholt mit dem PPS-Laserscanner in einem dynamischen Szenario von einer beweglichen Messplattform aus erfasst wird.

5.2.1 Analyse des Prüfkörpers

Wie in Kapitel 3.2.2 erläutert, wurde für die Untersuchungen zur Präzision des im Messsystem I.R.I.S 3 integrierten PPS-Laserscanners ein Testkörper verwendet, der auch zur Kalibrierung des Laser Crack Measurement Systems (LCMS) des kanadischen Herstellers Pavemetrics verwendet wird. Dieser Prüfkörper wurde mithilfe der Software CloudCompare aus Geometrieprimitiva dreidimensional modelliert (siehe Kapitel 4.1.2). Hierfür wurden ein flacher Quader als Grundplatte, zwei identische längliche Quader für die vordere und hintere Erhebung sowie ein längliches Dreiecksprisma für die mittlere Erhebung miteinander kombiniert. Konstruktive Details, wie beispielsweise die Flügelmuttern an den Standfüßen oder die seitlich am Prüfkörper befestigten Libellen, wurden nicht modelliert. Abschließend wurden die konstruierten Primitiva als ein sogenanntes Mesh-Objekt zusammengefasst und vom lokalen Koordinatenursprung an die grobe Geoposition der Prüfkörpermessung transformiert (siehe Abbildung 159).

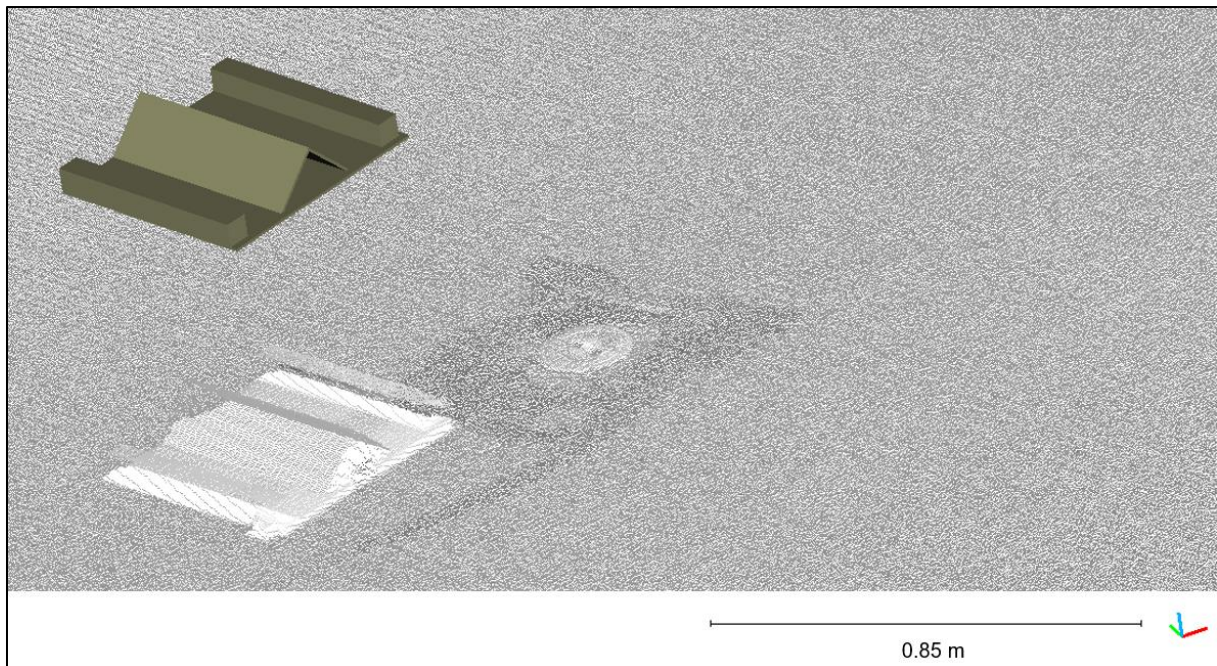


Abbildung 159: Perspektivische Darstellung eines Laserscans der acht Wiederholungsmessungen zusammen mit dem modellierten Prüfkörper, der grob an die Geoposition des gescannten Prüfkörpers transformiert wurde. In der Abbildung befindet sich das 3D-Modell des Prüfkörpers ca. 1 m über dem gescannten realen Prüfkörper.

Mithilfe der Software CloudCompare wurde aus jedem Laserscan der gescannte Prüfkörper als separates Punktkluster freigestellt und eine TIN-Vermaschung mittels Delaunay-Triangulation berechnet. Als minimale Kantenlänge für die Triangulation wurde der Parameter 0,05 m genutzt. Um die Bestimmung der Präzision vom Einfluss der absoluten Positionierungsgenauigkeit zu bereinigen, wurden die TINs der acht Messungen jede für sich auf den konstruierten Prüfkörper registriert, bevor die Abweichung von gescannter zu konstruierter Zielloberfläche bestimmt wurde (siehe Abbildung 160).

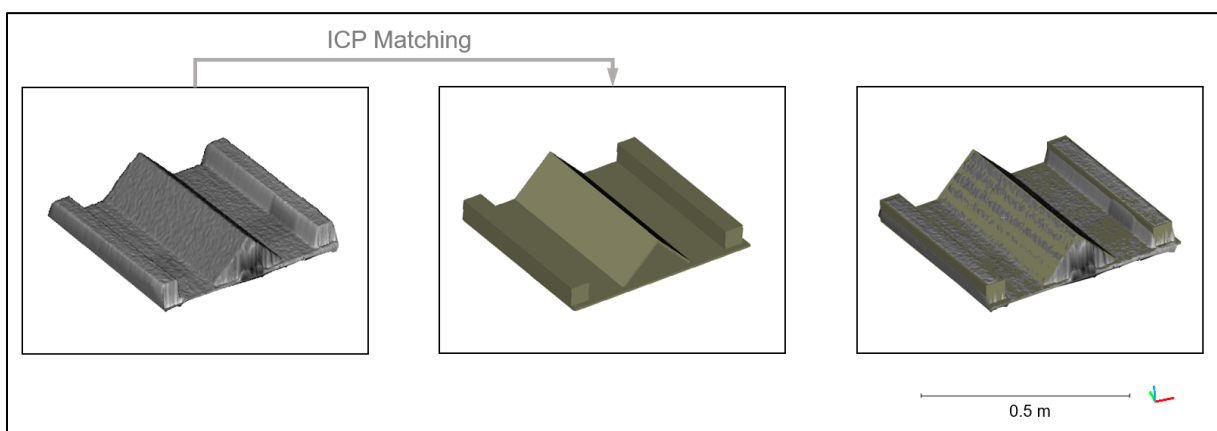


Abbildung 160: Registrierung der vermaschten Scandaten auf das 3D-Modell des Prüfkörpers mittels ICP-Verfahren.

Die Software CloudCompare nutzt für die Registrierung von zwei 3D-Punktwolken oder zwei TIN-Oberflächen P_1 und P_2 , die zumindest teilweise dasselbe Objekt abbilden, jedoch beispielsweise aus unterschiedlichen Richtungen oder zu unterschiedlichen Zeitpunkten erfasst wurden, das Verfahren Iterative Closest Point (ICP) (BESL & MCKAY 1992). Bei der ICP-Methode werden die Translations- und Rotationsparameter, die P_1 und P_2 aufeinander registrieren, durch die Beziehung zwischen in beiden Datensätzen übereinstimmenden Punkten berechnet. Der schrittweise Anpassungsprozess wird solange ausgeführt, bis die Minimierung des Registrierungsfehlers an den Punktpaaren konvergiert bzw. bis eine vorgegebene Konvergenzgenauigkeit erfüllt wird. Wichtig bei der ICP-Implementierung in der Software CloudCompare ist, dass die zu registrierenden Punktwolken vor der Feinregistrierung durch den ICP-Algorithmus bereits grob zueinander transformiert wurden. Da die vom Algorithmus vorgenommene Selektion der initialen Punktpaare Einfluss auf den gesamten Iterationsprozess hat, kann der Registrierungsprozess bei einer ungünstigen Ausgangslage von P_1 und P_2 zueinander in ein nicht optimales lokales Minimum des Registrierungsfehlers geraten (HE et al. 2017).

Für die Registrierung der acht Laserscan-basierten TIN-Oberflächen auf die 3D-Konstruktion des Prüfkörpers wurde als Abbruchkriterium definiert, dass die Verbesserung des RMS-Wertes zwischen zwei aufeinanderfolgenden Iterationen kleiner als 0,00001 wird (default-Wert). Tabelle 12 zeigt die Ergebnisse des Registrierungsprozesses in der zweiten Spalte.

Tabelle 12: Aufgeführt sind die RMSE-Werte nach der ICP-Registrierung der jeweiligen Messdaten auf den konstruierten Prüfkörper. Die RMSE-Werte beziehen sich v. l. n. r. auf den gesamten Prüfkörper, auf die horizontalen Oberflächen des Prüfkörpers und die geneigten Oberflächen des Prüfkörpers. Die mit * gekennzeichneten RMSE-Werte werden CloudCompare-intern nur auf einen Teil der registrierten Daten berechnet.

Messung [Nr.]	RMSE [mm] ICP Registrierung *	RMSE [mm] alle Oberflächen	RMSE [mm] horizontale Oberflächen	RMSE [mm] geneigte Oberflächen
1	1,91	1,28	0,71	0,80
2	2,15	1,30	0,69	0,83
3	1,92	1,28	0,67	0,80
4	1,89	1,31	0,69	0,81
5	1,73	1,21	0,67	0,77
6	1,76	1,19	0,67	0,84
7	1,86	1,23	0,70	0,84
8	1,90	1,22	0,68	0,75

Um abschließend die Präzision zu beschreiben, wurden alle registrierten 3D-Punkte der acht Messungen in einem einzigen 3D-Punktwolkenobjekt zusammengefasst und dann der Abstand

jedes Punktes zum modellierten Prüfkörper berechnet. Daraus wurde ein RMSE berechnet, der die Streuung des Messsystems und somit dessen Präzision quantifiziert.

Da Laserscanning bei der Messung von Kanten messprinzipbedingten Einschränkungen unterliegt (siehe Abbildung 161), werden in Tabelle 12 in den letzten beiden Spalten die RMSE-Werte für die horizontalen und geneigten Oberflächenelemente jeweils ohne die Messungen im Bereich der Prüfkörperkanten ausgewiesen.

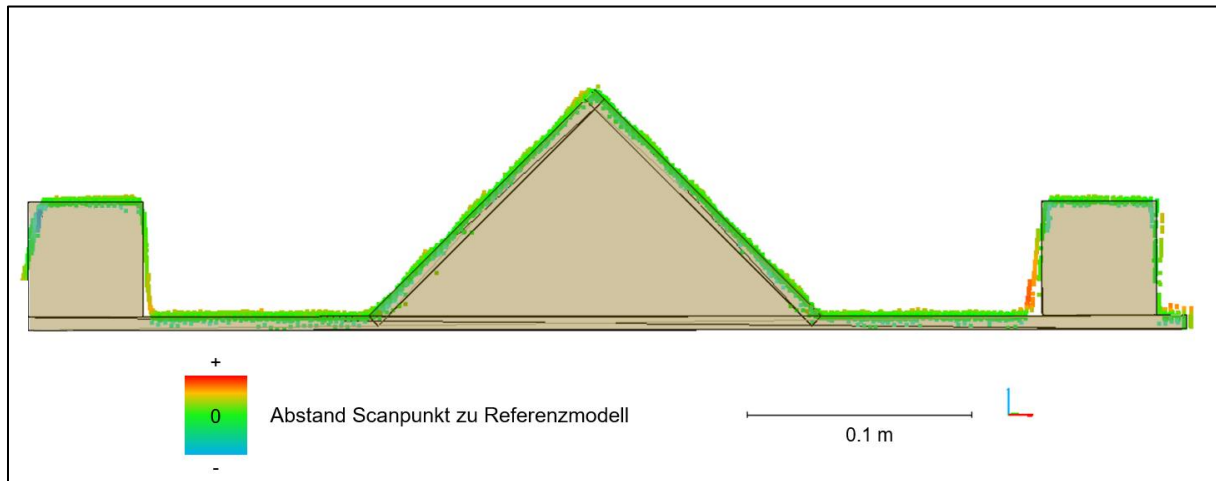


Abbildung 161: Profilansicht durch den modellierten Prüfkörper (braun). Die Farben der ICP-registrierten Scanpunkte kodieren den Abstand zum Referenzmodell. Auffällig sind die erhöhten Abweichungen (rot und blau) an den Kanten der Quader, die sich am Beginn und Ende auf der Grundplatte befinden.

5.2.2 Erkenntnisse zur Präzision des Messsystems

Die Untersuchungen zur Präzision des Messsystems auf Basis einer Achtfachmessung eines Prüfkörpers kommt zu dem Ergebnis, dass die durch das Messsystem verursachte Streuung (RMSE) bei 1,25 mm liegt. Die Aussagekraft von Lasermessungen an Objektkanten ist jedoch wegen der Strahldivergenz und der daraus resultierenden Spotgröße des Lasers am Messobjekt eingeschränkt, da hier Mischsignale entstehen, die zu einer Ausrundung bzw. Abrundung von Kanten und Ecken führen. In Abbildung 161 ist dieser Effekt am Prüfkörper im Längsprofilschnitt deutlich zu sehen. Da ein nicht unerheblicher Teil des Prüfkörpers von diesen Effekten betroffen ist, wurde die Präzision für die horizontalen und geneigten Flächen des Prüfkörpers zusätzlich ohne die Kantenbereiche und ohne die vertikalen Flächen untersucht. Betrachtet man nur die horizontalen und geneigten Oberflächen des Prüfkörpers, beträgt der RMSE 0,68 mm bzw. 0,80 mm. Dementsprechend liegt die durch das Messsystem verursachte Streuung ohne die Kantenbereiche des Prüfkörpers bei 0,75 mm.

Wie bereits in Kapitel 3.1.1.2 beschrieben, sind herstellerseitig keine Angaben zur Strahlaufweitung oder zur Größe des Footprints gemacht. Das festgestellte Ausmaß der Ausrundung deutet aber darauf hin, dass eine Spotgröße von etwa 5 mm senkrecht unter dem Fahrzeug angenommen werden kann (siehe Abbildung 162).

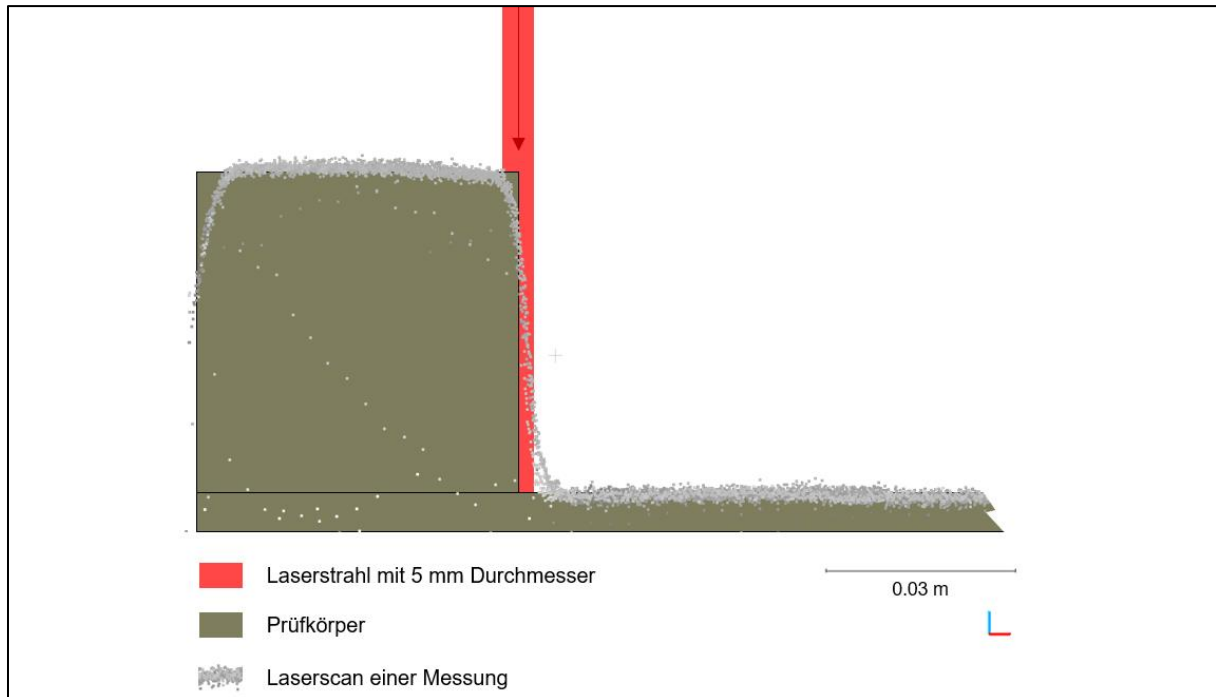


Abbildung 162: Ausrundungseffekt an den Kanten des Prüfkörpers. Idealisiert dargestellt ist außerdem ein Laserstrahl mit 5 mm Durchmesser, der orthogonal auf die Prüfkörperoberfläche auftrifft. Durch das an der Kante entstehende Mischsignal kann die Kantenmorphologie nicht korrekt abgebildet werden. Wenn der Laserstrahl die Kante anschneidet, werden infolge der entstehenden Mischsignale 3D-Punkte erzeugt, die weder auf der Oberseite der Kante noch am Fußpunkt liegen.

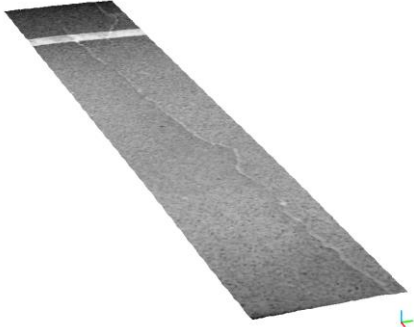
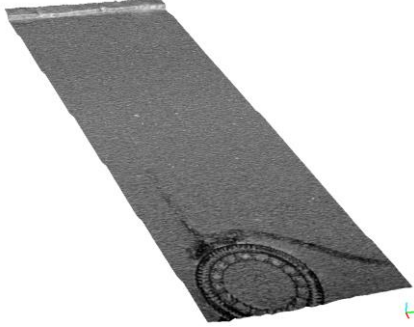
5.3 Oberflächenmodellierung

In den folgenden Abschnitten wird eine Auswahl von deterministischen, räumlichen Interpolationsverfahren vergleichend analysiert. Ziel des Vergleichs ist es, aus den untersuchten Interpolationsverfahren jene Methode auszuwählen, mit deren Hilfe Oberflächenmodelle aus PPS-Laserscans erzeugt werden können, die im Hinblick auf die Extraktion von Oberflächenschäden optimal sind. Hierzu wird in Anlehnung an die sogenannte Nutzwertanalyse eine Entscheidungsmetrik erstellt, welche die einzelnen untersuchten Interpolationsverfahren entlang definierter Kriterien bewertet.

5.3.1 Parametrisierung der Interpolationsverfahren

Als Datengrundlage für den Vergleich von Interpolationsmethoden wurden Ausschnitte aus Oberflächenscans von Straßen genutzt, die als typisch für in Asphaltbauweise ausgeführte Straßen des kommunalen Netzes angesehen werden können. In Kapitel 3.2.3 sind diese im Detail beschrieben. Tabelle 13 fasst wichtige Eigenschaften der Testareale zusammen. Unter Nutzung der Software SAGA GIS (siehe Kapitel 4.1.3) wurden die ausgewählten deterministischen Interpolationsverfahren mit unterschiedlichen Parametersätzen auf die absoluten Höhenwerte der gescannten Testareale angewandt und evaluiert. Die verwendeten Algorithmen zur Erzeugung gerasterter Oberflächenmodelle aus den ungeordneten 3D-Punktwolken des Laserscans sind Teil der SAGA GIS Bibliotheken „Grid – Gridding“ und „Grid – Spline Interpolation“. Die im folgenden aufgeführten Parameter richten sich daher nach der dortigen Implementierung und Nomenklatur der Interpolationsmethoden.

Tabelle 13: Überblick über wichtige Charakteristika der beiden Testareale.

Merkmal	Testareal 1	Testareal 2
Mittlere Punktdichte ($r = 0,03\text{m}$)	79	52
Punktzahl	100876	99502
Bounding-Box ($x[\text{m}]; y[\text{m}]; z[\text{m}]$)	0,85; 4,33; 0,17	1,79; 4,04; 0,13
Besonderheiten	offener Querriss Fahrbahnmarkierung	sanierter Riss Abgesenkter Bordstein Schachtdeckel
Übersichtsbild (Schrägsicht TIN)		

Die Parametrisierung der Oberflächenmodellierung teilt sich in zwei Blöcke: Parameter, die für alle Methoden identisch sind und verfahrensspezifische Parameter. Zu Ersteren zählen die Parameter Suchradius, Mindestpunktzahl und Zielzellengröße. Im Zuge der Oberflächeninterpolation wurden standardmäßig immer alle lokalen Nachbarpunkte in einem

festen Radius von $r = 0,03$ m berücksichtigt. Dieser Radius wurde deshalb gewählt, weil der Abstand zwischen zwei aufeinanderfolgenden Scanprofilen des PPS-Laserscanners bei der maximalen Erfassungsgeschwindigkeit des I.R.I.S-Systems von 80 km/h ca. 0,028 m beträgt. Mit einem Suchradius von $r = 0,03$ m ist somit sichergestellt, dass für die Oberflächenerstellung nicht nur die Messpunkte eines einzigen Profils genutzt werden, sondern immer auch Messpunkte des Vorgänger- bzw. Nachfolgerprofils. Für die Interpolationsverfahren, die nicht nur mit einem festen, sondern zusätzlich auch mit einem variablen Suchradius analysiert wurden, wurde der maximale Suchradius ebenfalls auf $r = 0,03$ m begrenzt, die maximale Anzahl zu nutzender Nachbarpunkte wurde auf 30 eingestellt und als Richtungsparameter wurde die Option „Quadranten“ genutzt. Die Quadrantensuche sorgt für eine gleichmäßigere Selektion relevanter Nachbarpunkte und ist hilfreich, um die räumliche Verteilung der Datenpunkte, die durch die profilweise Erfassung mittels Laserscanner determiniert wird, auszugleichen. Die betreffenden Varianten der IDW- und der TPS-Methode sind in Tabelle 14 mit dem Zusatz *VR* (variabler Radius) gekennzeichnet. Ferner wurde für alle Interpolationsverfahren, egal ob mit festem oder variablem Suchradius berechnet, die Mindestanzahl zu verwendender Punkte auf 1 und die Pixelauflösung des Zielrasters (engl.: *Cellsize*) auf 0,005 m gesetzt. Nachfolgend sind die verfahrensspezifischen Parameter beschrieben, die im Zuge der vergleichenden Analyse variiert wurden.

Die NN-Interpolationsmethode wurde als lineare Variante, sowie als Sibson- und Non-Sibson-Variante evaluiert. SAGA GIS bietet darüber hinaus keine weiteren Parametrisierungsmöglichkeiten. Es ist hierbei anzumerken, dass die lineare Variante im Grunde eine TIN-Interpolation auf Basis einer Delaunay-Triangulation ist. Dies geht aus dem Quellcode hervor, der von SAGA GIS genutzt wird (SAKOV 2021).

Für die IDW-Interpolation wurde lediglich der Exponent variiert. Neben dem Standardwert $\beta = 2$ wurden noch die Exponenten $\beta = 4$ und $\beta = 6$ untersucht. Je höher der Zahlenwert des Exponenten, desto geringer ist der Einfluss der Nachbarpunkte auf das Interpolationsergebnis und desto akzentuierter bzw. detailreicher wirkt die Oberfläche. Mit den genannten Exponenten wurde jeweils auch eine Variante mit variablem Radius berechnet.

SAGA GIS bietet für die MQS-Interpolation die Möglichkeit, die Anzahl der Quadratic Neighbors n_q und die der Weighting Neighbors n_w zu parametrisieren. Zusätzlich zu dem bei RENKA (1988) als optimal vorgeschlagenen Setting ($n_q = 13$, $n_w = 19$) wurden die Parametersätze ($n_q = 5$, $n_w = 15$), ($n_q = 9$, $n_w = 13$), ($n_q = 16$, $n_w = 23$) sowie ($n_q = 23$, $n_w = 29$) untersucht. Die Parametersettings ($n_q = 30$, $n_w = 42$) und ($n_q = 45$, $n_w = 65$) führten zu keinem Ergebnis und werden daher auch nicht weiter aufgeführt.

Für die TPS-Interpolation wurden verschiedene Ausprägungen des Regularisierungsparameters λ getestet. Ausgehend von $\lambda = 0,0$, was einer exakten Interpolation entspricht, wurden noch die Werte $\lambda = 0,001$, $\lambda = 0,01$, $\lambda = 0,1$ und $\lambda = 0,5$ untersucht. Mit genannten Regularisierungsparametern wurde jeweils auch eine Variante mit variablem Radius berechnet.

5.3.2 Vergleich der Interpolationsverfahren

Nachdem die Oberflächenmodelle mit verschiedenen Interpolationsmethoden und Parametersätzen erzeugt wurden, wurden die Höhenwerte der Kontrollpunkte mit den entsprechenden Höhen in den generierten Oberflächenmodellen kreuzvalidiert. Für die Evaluation der ausgewählten räumlichen Interpolationsverfahren wurden 20 % der im Laserscan des jeweiligen Testbereichs enthaltenen Messpunkte als Kontrollpunkte separiert. Die verbleibenden 80 % dienten als Datengrundlage zur Interpolation von Oberflächenmodellen. Aufgrund der hohen Datendichte der verwendeten Laserscans wurde diese Vorgehensweise der zufälligen Aufteilung der Daten nur einfach ausgeführt.

Die räumliche Verschneidung der gerasterten Oberflächenmodelle mit den Kontrollpunkten wurde mit dem Werkzeug „Add Grid Values To Points“ mit der Wertzuweisungsoption „Nearest Neighbor“ umgesetzt. Das Werkzeug ist Teil der SAGA-GIS-Bibliothek „Shapes – Grid Tools“. Schließlich wurde zur Evaluierung der vertikalen Genauigkeit der RMSE gemäß der nachstehenden Formel berechnet, indem man die Differenzen zwischen den tatsächlichen und den berechneten Werten quadriert, die resultierenden Residuen aufaddiert, durch die Gesamtzahl der Werte teilt und schließlich die Quadratwurzel des Ergebnisses berechnet (SZYPUŁA 2016).

$$RMSE_{elev} = \sqrt{\frac{\sum (z_p - z_o)^2}{n}} \quad (21)$$

Hierbei ist z_o ein Höhenwert des interpolierten Modells und z_p bezeichnet die Höhe des entsprechenden Kontrollpunkts. Die Anzahl der Kontrollpunkte wird repräsentiert durch n . Der RMSE basiert auf der Annahme, dass die Fehler im Geländemodell zufällig und normalverteilt sind. Abbildung 163 zeigt, dass diese Annahme für das verwendete Datenmaterial zutreffend ist. Dargestellt ist die Streuung der Laserscannerpunkte auf einer der horizontalen Flächen des LCMS-Prüfkörpers (Kapitel 3.2.2).

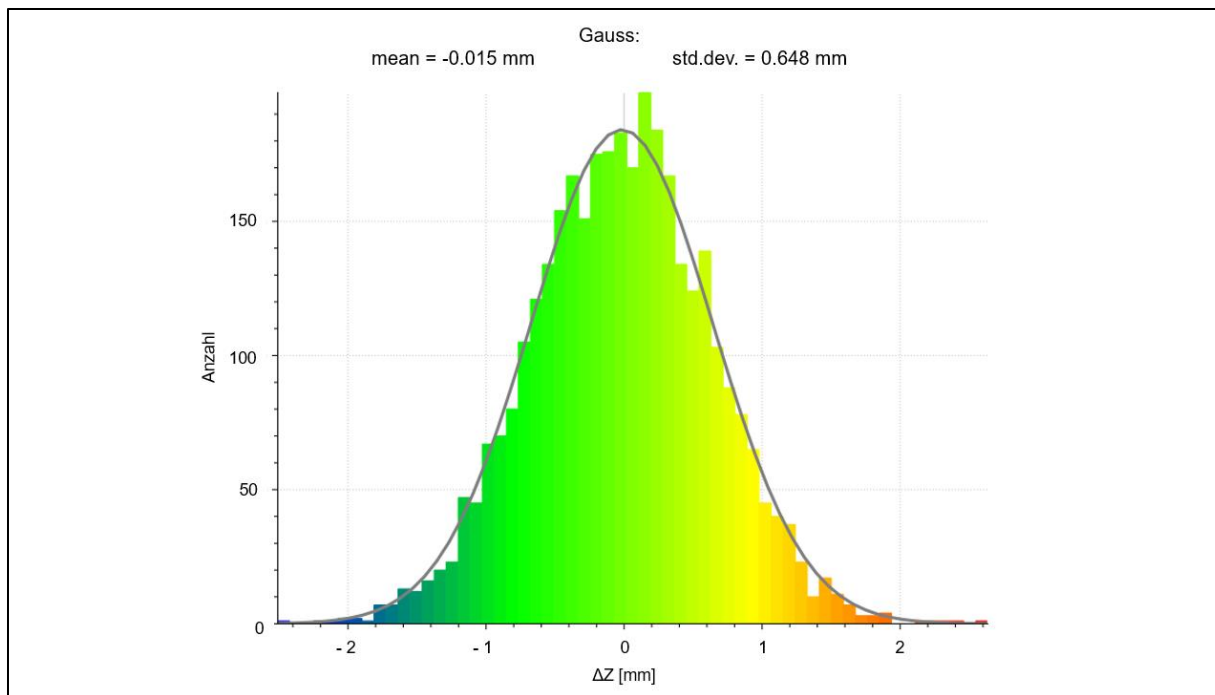


Abbildung 163: Streuung der Laserscanner-Messpunkte auf einem Oberflächenelement von 9 cm x 44 cm Größe, welches durch 3436 Lasermessungen erfasst wurde. In grau ist die Kurve einer Gauß-Verteilung dargestellt.

Zur Quantifizierung der Interpolationslaufzeit, wurde weder Testareal 1 noch Testareal 2 (siehe Tabelle 13), sondern ein Laserscan aus dem in Kapitel 3.2.5 beschriebenen Datensatz genutzt. Er umfasst 747492 3D-Messpunkte auf 10 m Länge und ca. 4,2 m Breite. Diese Dimensionen und die Punktzahl bzw. -dichte sind repräsentativ für Laserscans, die mit dem PPS im kommunalen Straßennetz außerorts erfasst werden. Die zur Oberflächenerzeugung benötigten Rechenzeiten wurden den Log-Informationen von SAGA GIS entnommen.

Die Parameteranzahl gibt die Anzahl der zur Anwendung der Interpolationsmethode in der Software SAGA GIS nötigen Parameter an. Ausgenommen sind hierbei jene Parameter, die für alle Verfahren identisch sind. Im Rahmen der Untersuchung sind das die Parameter Zellauflösung des Zielrasters und Mindestpunktzahl.

Zur Berechnung der Entscheidungsmetrik, auf deren Basis die Auswahl einer geeigneten Interpolationsmethode getroffen werden soll, wurden alle Kriterien auf ihren Maximalwert normiert, das gewichtete Mittel gebildet und dieses mit 100 multipliziert. Um zum Ausdruck zu bringen, dass die Genauigkeit des Interpolationsverfahrens eine höhere Bedeutung hat, als die Laufzeit und die Parameteranzahl, wurden die RMSE-Werte der beiden Testfelder jeweils mit dem Faktor 1, die Parameteranzahl mit dem Faktor 0,5 und die Rechenzeit mit dem Faktor 0,75 gewichtet. Der Wertebereich des resultierenden Scores liegt zwischen 0 und 100, wobei der Wert 100 der maximal erreichbaren Punktzahl für eine Wahlmöglichkeit entspricht.

Die Ergebnisse der vergleichenden Analyse von Interpolationsmethoden zur Erzeugung von Oberflächenmodellen aus PPS-Laserscans sind in Tabelle 14 zusammengefasst. Anhang VII stellt die normierten Wertausprägungen der einzelnen bewerteten Kriterien in Form eines Balkendiagramms dar.

Tabelle 14: Ergebnisse des Vergleichs von verschiedenen räumlichen Interpolationsmethoden und unterschiedlichen Parametersätzen anhand von zwei Testfeldern. Die analysierten Kriterien sind die vertikale Genauigkeit ausgedrückt als RMSE, die Anzahl von verfahrensspezifischen Parametern sowie die Rechenzeit.

Methode (Parameter)	RMSE Testfeld 1 [mm]	RMSE Testfeld 2 [mm]	Parameter [Anzahl]	Rechenzeit [s]	Score [0 bis 100]
NN-Linear	1,01	1,46	0	5	100
NN-Sibson	0,98	1,42	0	10	97
NN-NonSibson	0,96	1,39	0	12	96
IDW ($\beta = 2$)	0,95	1,44	3	7	87
IDW ($\beta = 4$)	1,03	1,49	3	7	87
IDW ($\beta = 6$)	1,09	1,54	3	7	87
IDW-VR ($\beta = 2$)	1,00	1,58	4	7	83
IDW-VR ($\beta = 4$)	1,07	1,51	4	7	83
IDW-VR ($\beta = 6$)	1,13	1,54	4	7	83
MQS ($n_q = 5$, $n_w = 7$)	26,11	40,83	2	7	30
MQS ($n_q = 9$, $n_w = 13$)	2,88	13,96	2	10	77
MQS ($n_q = 13$, $n_w = 19$)	1,38	7,69	2	13	83
MQS ($n_q = 16$, $n_w = 23$)	1,17	6,94	2	15	82
MQS ($n_q = 23$, $n_w = 33$)	1,09	2,83	2	21	83
TPS ($\lambda = 0,0$)	1,07	1,79	3	48	65
TPS ($\lambda = 0,0001$)	1,07	1,74	3	49	65
TPS ($\lambda = 0,001$)	1,06	1,61	3	48	66
TPS ($\lambda = 0,01$)	1,02	1,48	3	48	66
TPS ($\lambda = 0,1$)	0,95	1,37	3	47	66
TPS ($\lambda = 0,5$)	0,92	1,35	3	47	66
TPS-VR ($\lambda = 0,0$)	1,19	1,81	4	20	76
TPS-VR ($\lambda = 0,0001$)	1,19	1,77	4	19	77
TPS-VR ($\lambda = 0,001$)	1,18	1,64	4	19	77
TPS-VR ($\lambda = 0,01$)	1,14	1,54	4	19	77
TPS-VR ($\lambda = 0,1$)	1,04	1,43	4	20	77
TPS-VR ($\lambda = 0,5$)	0,98	1,39	4	19	77

5.3.3 Erkenntnisse zur Oberflächenmodellierung

Im Zuge der vorstehend beschriebenen Analyse konnten Unterschiede zwischen den untersuchten räumlichen Interpolationsverfahren festgestellt werden. Neben unterschiedlicher Parametersensitivität und Rechenkomplexität unterscheiden sich die Ergebnisse auch im Hinblick auf ihre vertikale Genauigkeit und visuelle Aspekte. Folgende Ergebnisse des durchgeführten Vergleichs von Interpolationsmethoden können festgehalten werden.

Die höchste vertikale Genauigkeit bzw. der geringste RMSE wird mit der TPS-Interpolation mit dem Regularisierungsparameter $\lambda = 0,5$ erreicht. Einen nur unwesentlich höheren RMSE hat die IDW-Interpolation mit $\beta = 2$. Das Verfahren mit der schlechtesten vertikalen Genauigkeit ist die MQS-Interpolation mit dem Parametersetting ($n_q = 5$, $n_w = 7$). Während die Parametersettings ($n_q = 13$, $n_w = 19$), ($n_q = 16$, $n_w = 23$) und ($n_q = 23$, $n_w = 33$) im Testgebiet 1 noch RMSE-Werte liefern, die etwa im Bereich der Interpolationsgüte der übrigen Verfahren liegen, so ist das für das Testgebiet 2 nicht der Fall. Hier spielt die geringere Punktdichte eine wesentliche Rolle. Die Spannweite der RMSE-Werte beträgt mit Ausnahme der MQS-Varianten 0,27 mm im Testgebiet 1 und 0,47 mm im Testgebiet 2.

Die drei untersuchten Interpolationsverfahren aus der Gruppe der NN-Interpolation (Linear, Sibson und Non-Sibson) erfordern den geringsten Parametrisierungsaufwand. Es muss lediglich die Auflösung des Zielrasters angegeben werden. Das Verfahren mit dem höchsten Parametrisierungsaufwand ist die TPS-Interpolation mit variablem Radius. Hier müssen zusätzlich zur Zielauflösung die Parameter maximaler Suchradius, maximale Anzahl zu nutzender Nachbarpunkte sowie der Regularisierungsparameter λ eingestellt werden.

Die mit Abstand geringste Laufzeit benötigt die lineare NN-Methode. Schon die nächst langsameren Interpolationsverfahren aus der Gruppe der IDW-Varianten sind um den Faktor 1,4 langsamer. Das langsamste Verfahren ist die TPS-Methode mit $\lambda = 0,0001$. Die Laufzeit ist nahezu um den Faktor 10 langsamer als die schnellste Interpolationsmethode.

Dem Konzept der Nutzwertanalyse folgend kann die lineare NN-Methode mit der maximal erreichbaren Punktzahl als für das definierte Ziel am besten geeignete räumliche Interpolationsmethode angesehen werden, auch wenn sie im Versuch nicht die bestmöglichen vertikalen Genauigkeiten liefert. Der Abstand zum RMSE des genauesten Verfahrens beträgt im Testfeld 1 nur 0,09 mm und im Testfeld 2 nur 0,11 mm. Der große Vorteil gegenüber den anderen untersuchten Verfahren ist die Schnelligkeit und der geringe Parametrisierungsaufwand der linearen NN-Methode. Demzufolge wird sie auch im Rahmen aller weiterführenden Analysen dieser Arbeit genutzt. Das betrifft auch die Interpolation der radiometrischen Information zu Intensitätsbildern. Es kann davon ausgegangen werden, dass

die Interpolation der Reflexionsintensität zu keiner abweichenden Bewertung und Entscheidung hinsichtlich der zu wählenden Interpolationsmethode führen wird, denn wie die absolute Höhe ist die Reflexionsintensität ein Attribut an einer 2D-Position. Wie Abbildung 164 zeigt, ist das Interpolationsergebnis mit der linearen NN-Methode visuell ansprechend: alle relevanten Informationen werden fotorealistisch wiedergegeben. Andere Interpolationsverfahren erzeugen dagegen teils verwaschene bzw. stark verrauschte Ergebnisse oder enthalten sehr kleinräumige Artefakte.

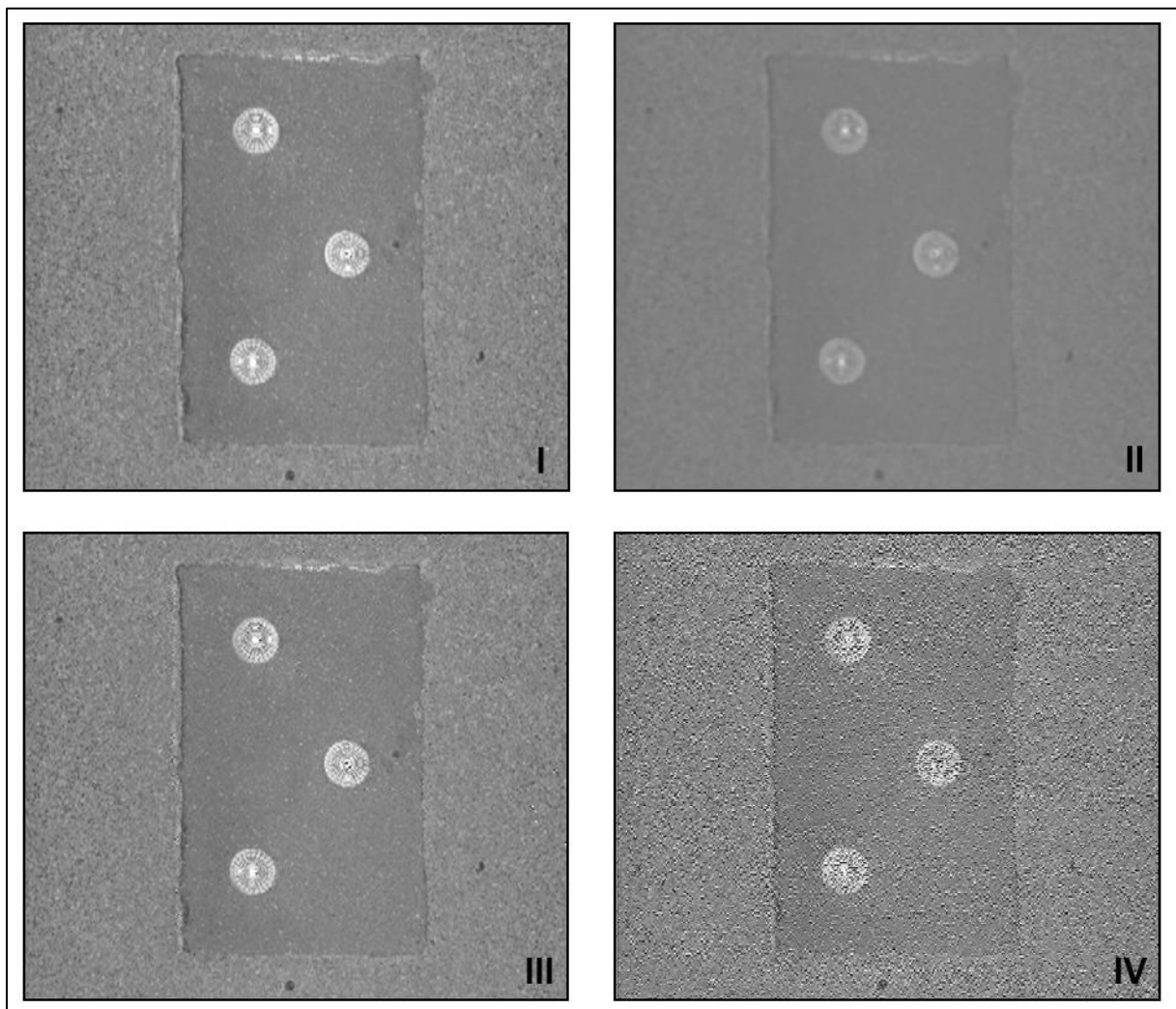


Abbildung 164: Vergleichende Darstellung von Oberflächenmodellen, die auf Basis der Reflexionsintensität mit unterschiedlichen Interpolationsmethoden erzeugt wurden. Verglichen wurden von oben links nach unten rechts die Methoden NN linear (I), TPS mit $\lambda = 0,5$ (II) und MQS mit $n_q = 9$, $n_w = 13$ (III) sowie $n_q = 5$, $n_w = 7$ (IV). Die Situation zeigt eine Flickstelle, in der sich drei Wasserschieber mit einem genormten Kappendurchmesser von jeweils 19 cm befinden.

5.4 Erkennbarkeit von Oberflächendiskontinuitäten

Die in diesem Kapitel beschriebenen Untersuchungen zielen darauf ab, zu analysieren, bis zu welchem Ausmaß Oberflächendiskontinuitäten in Oberflächenmodellen, die auf Basis von PPS-Laserscans generiert wurden, identifiziert werden können. In diesem Zusammenhang steht auch die Frage nach dem Einfluss der Befahrungsgeschwindigkeit auf die Qualität der Oberflächenmodelle und die Frage danach, wo die Grenze der Erkennbarkeit von rissähnlichen Strukturen aus PPS-Oberflächenmodellen liegt.

5.4.1 Analysekonzept

Die Untersuchung der Erkennbarkeit von Oberflächendiskontinuitäten wurde anhand eines Prüfkörpers durchgeführt, der Spalten definierter Breite und Tiefe aufweist. Details zum Prüfkörper und zur Datenaufnahme sind im Kapitel 3.2.4 beschrieben. Der Prüfkörper wurde in verschiedenen Messanordnungen mehrfach gescannt. Dabei wurde der Prüfkörper sowohl längs als auch quer zur Fahrtrichtung ausgerichtet und bei 30 km/h sowie bei 50 km/h aufgenommen. Auf Basis dieser Laserscans wurde je ein Oberflächenmodell unter Verwendung der linearen NN-Methode erzeugt. Die Rasterweite der Oberflächenmodelle beträgt 5 mm.

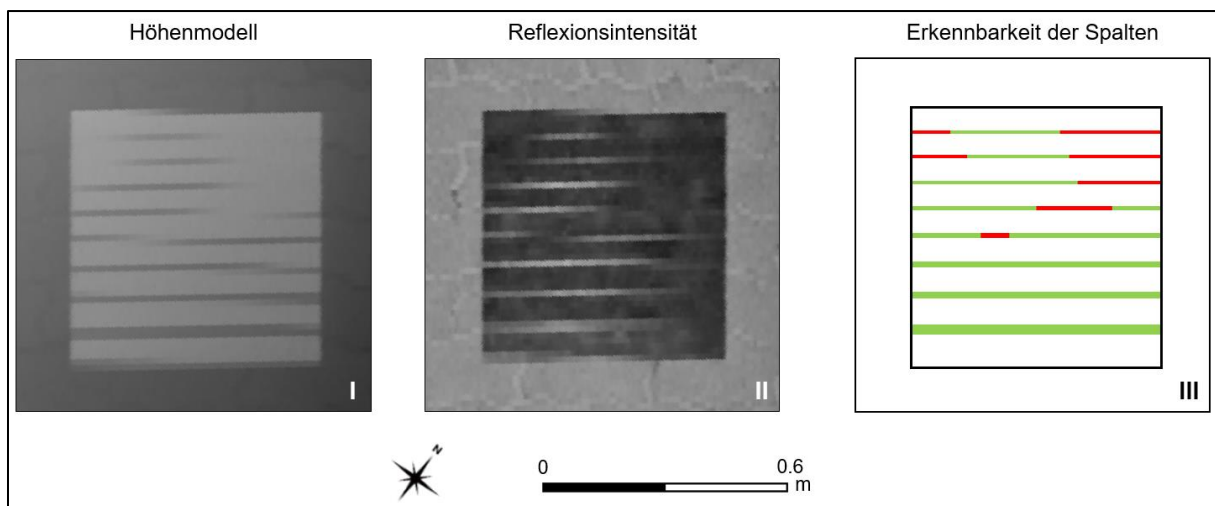


Abbildung 165: Kartierung der Erkennbarkeit von Spalten auf Basis des Höhenmodells (I). Nicht sicher erkennbare Teile der Spalten sind in der Teilabbildung III rot markiert. Erkennbare Anteile sind grün markiert. Für jede Prüfkörperaufnahme wurde nach dieser Vorgehensweise ermittelt, zu wieviel Prozent die unterschiedlich breiten Spalten erkennbar sind.

Die erzeugten Oberflächenmodelle wurden visuell analysiert, wobei jeder Spalt als Linienobjekt an seiner Soll-Position auf dem Prüfkörper mithilfe der Software QGIS abdigitalisiert wurde (siehe Kapitel 4.1.3). Je nachdem, ob der entsprechende Spalt im Oberflächenmodell

erkennbar war oder nicht, wurde das betreffende Linienobjekt in Segmente aufgeteilt und diese mit der Information „erkennbar“ und „nicht erkennbar“ attribuiert. Somit konnte für jede Spaltbreite in jedem aufgenommenen Szenario quantifiziert werden, welcher prozentuale Anteil erkennbar ist. Abbildung 165 zeigt diese Vorgehensweise exemplarisch an einer Prüfkörpermessung, die bei 50 km/h durchgeführt wurde und bei der die Spalten parallel zur Scanrichtung (in Fahrtrichtung) ausgerichtet waren. In Anhang VIII sind weitere Beispiele zu finden. Zusätzlich zur rein visuellen Analyse der Höhenmodelle wurden für alle Messszenarien Profilmessungen durchgeführt. Abbildung 166 zeigt beispielhaft eine solche Überlagerung von gemessenen Profilschnitten mit dem Sollprofil des Prüfkörpers.

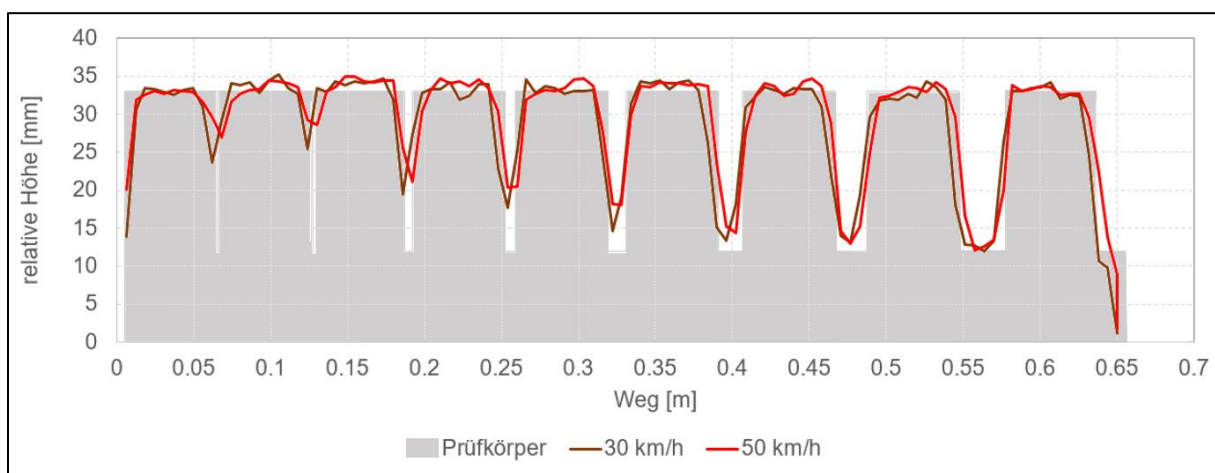


Abbildung 166: Exemplarische Profile durch den gescannten Prüfkörper. Die Ausrichtung der Spalten ist hierbei orthogonal zur Scanrichtung.

5.4.2 Erkenntnisse zur Erkennbarkeit von Oberflächendiskontinuitäten

Die Ergebnisse der Untersuchungen zur Erkennbarkeit von Oberflächendiskontinuitäten sind in Tabelle 15 zusammengefasst. Die dort aufgeführten Werte repräsentieren den Mittelwert der erkannten Anteile der Spalten über die Wiederholungsmessungen. Die Standardabweichung ist in Klammern hinter dem Mittelwert angegeben. Die Ergebnisse lassen den Schluss zu, dass bei den im kommunalen Straßennetz üblichen Befahrungsgeschwindigkeiten Oberflächendiskontinuitäten ab einer Öffnungsbreite von 1 mm aus dem PPS-Laserscan des I.R.I.S-Systems detektiert werden können. Während Spalten mit einer Breite ≥ 5 mm über alle untersuchten Szenarien hinweg sicher und überwiegend vollständig detektiert werden konnten, war dies für Spalten mit einer Breite ≤ 2 mm nur bedingt möglich. Zwar konnten selbst Anteile von 1 mm breiten Spalten erkannt werden, jedoch beeinflusste hier die Spotgröße des Lasers in Verbindung mit der Scanfrequenz und Befahrungsgeschwindigkeit sowie die relative

Ausrichtung der Spalten in Bezug zur Scanrichtung deren Erkennbarkeit. Hierbei spielt eine entscheidende Rolle, ob genügend Photonen des Lasers in den Spalt gelangen und ein entsprechend hoher Anteil auch wieder vom Detektor empfangen werden kann. Dies wiederum ist abhängig von verschiedenen Faktoren, wie beispielsweise der vom Auftreffwinkel abhängigen Ausprägung des Laserfootprints auf dem Messobjekt (Kreis oder Ellipse), des Auftreffpunkts des Lasers relativ zur Spaltöffnung (Messung zentral in den Spalt hinein oder anscheinende Messung der Spaltkante) und der Eindringtiefe des Lasers in den Spalt.

Tabelle 15: Prozentualer Anteil der erkannten Spalltlänge bei verschiedenen Messszenarien für verschiedene Spaltbreiten bei einer festen Spalttiefe von 20 mm. Angegeben sind der Mittelwert sowie die Standardabweichung in Klammern bezogen auf die Wiederholungsmessungen/-auswertungen. Die Auswertung erfolgte auf Oberflächenmodellen mit einer 5 mm x 5 mm Auflösung.

Spaltbreite [mm]	Szenario	Erfassungsgeschwindigkeit: 30 km/h		Erfassungsgeschwindigkeit: 50 km/h	
		Ausrichtung: parallel	Ausrichtung: orthogonal	Ausrichtung: parallel	Ausrichtung: orthogonal
1		46 (6)	61 (40)	26 (18)	80 (20)
1,8		60 (13)	100 (0)	43 (3)	100 (0)
5		64 (4)	100 (0)	64 (2)	100 (0)
7		99 (1)	99 (1)	98 (2)	98 (1)
11,5		100 (0)	99 (1)	98 (2)	100 (0)
16		100 (0)	100 (0)	100 (0)	100 (0)
20,5		100 (0)	100 (0)	100 (0)	100 (0)
30		100 (0)	100 (0)	100 (0)	100 (0)

5.5 Deep Learning basierte Oberflächenanalyse

Die nachfolgenden Kapitel beschreiben alle durchgeführten Untersuchungen, die zur Entwicklung eines Analysesystems führten, das mithilfe von Deep Learning eine automatische Kartierung von Oberflächenschäden auf der Grundlage von PPS-Oberflächenmodellen realisiert. Das Grundprinzip des Analyseansatzes sieht vor, dass spezialisierte DeepLabV3+ Modelle für verschiedene Aufgaben auf entsprechenden Datensätzen trainiert werden. So ist pro Datenkanal des PPS (Höhe und Reflexionsintensität) jeweils ein Modell für die Segmentierung von asphalt- und betonspezifischen Schäden zuständig. Zusätzlich segmentiert ein weiteres Modell pflasterspezifische Oberflächenschäden auf Grundlage der Höheninformation. In der Inferenzphase werden alle drei DeepLabV3+ Modelle auf PPS-Oberflächenmodellen angewendet und ihre Ausgaben zu einem gemeinsamen, georeferenzierten Segmentierungsergebnis zusammengeführt.

5.5.1 Erstellung eines Trainings- und Evaluationsdatensatzes

Modelle, die zur automatischen Datenanalyse trainiert werden, sollten möglichst gute Generalisierungseigenschaften aufweisen. Wie in den Kapiteln 2.4.1.2 und 2.4.3 beschrieben, benötigen sie dazu nicht nur Beispieldaten in großem Umfang, sondern auch eine hohe Diversität in der Ausprägung der Daten. Deshalb wurde mit verschiedenen I.R.I.S Systemen über den Zeitraum mehrerer Jahre ein umfangreicher Bestand an PPS-Laserscans aufgenommen und hunderte relevante 3D-Punktwolken selektiert (siehe Kapitel 3.2.5).

Für alle 2544 verfügbaren 3D-Punktwolken wurden mithilfe der linearen NN-Interpolation Oberflächenmodelle auf Basis der Höheninformation (Z) und der Reflexionsintensität (I) mit einer Auflösung von 5 mm x 5 mm berechnet. Zusätzlich wurden auf Basis des Z-Höhenmodells die vier in Kapitel 4.3 beschriebenen Oberflächenanalyseraster berechnet: Hillshade (HS), Positive Openness (PO), Black Top Hat Transform (BTHT) und Topographic Position Index (TPI).

In einem manuellen Prozess wurde mithilfe der Software ScaLabel (siehe Kapitel 4.1.4) eine Auswahl von mehr als 3800 Z- und I-Oberflächenmodellen pixelgenau gelabelt. Diesem Labelling liegt ein Klassenkatalog zugrunde, der neben 14 relevanten Schadensarten auch sechs Deckenarten (Materialklassen) und diverse Hilfsklassen, wie Fahrbahnmarkierungen, Vegetation, Bordsteine oder Einbauten, umfasst.

In Abbildung 167 ist je an einem Beispiel dargestellt, in welchem Detailgrad das Labelling der relevanten Schäden vorgenommen wurde. Die asphaltspezifischen Schadensarten, die anhand der lokalen Höhenvarianz erkannt werden können, sind offener Riss (I), offene Naht (II), Ausbruch (III) und Ausmagerung bzw. Kornverlust (IV). Für Pflasterdecken können folgende Schäden auf Basis der Höheninformation erkannt werden: schadhaftes Pflasterelement (V), fehlender Pflasterstein (VI), Flickstelle (VII), Lagerungsschaden (VIII), offene Pflasterfuge (IX) und Gefügebrauch (X). Anhand der Reflexionsintensität können eingelegte Flickstelle (XI), aufgelegte Flickstelle (XII), sanierter Riss (XIII) und Bindemittelanreicherung (XIV) erfasst werden. Die Schäden, die auf Asphalt vorkommen, unterscheiden sich morphologisch nicht wesentlich von denen auf Betondecken. Aus diesem Grund und in Ermangelung einer offiziellen Erfassungsrichtlinie für das kommunale Straßennetz (siehe Kapitel 2.1.3.2) wurden alle asphaltspezifischen Schadensarten bis auf Bindemittelanreicherung und offene Naht auch für Betonbauweisen gelabelt. Insgesamt wurden mehr als 27000 Schadstellen konturscharf auf den Oberflächenmodellen markiert.

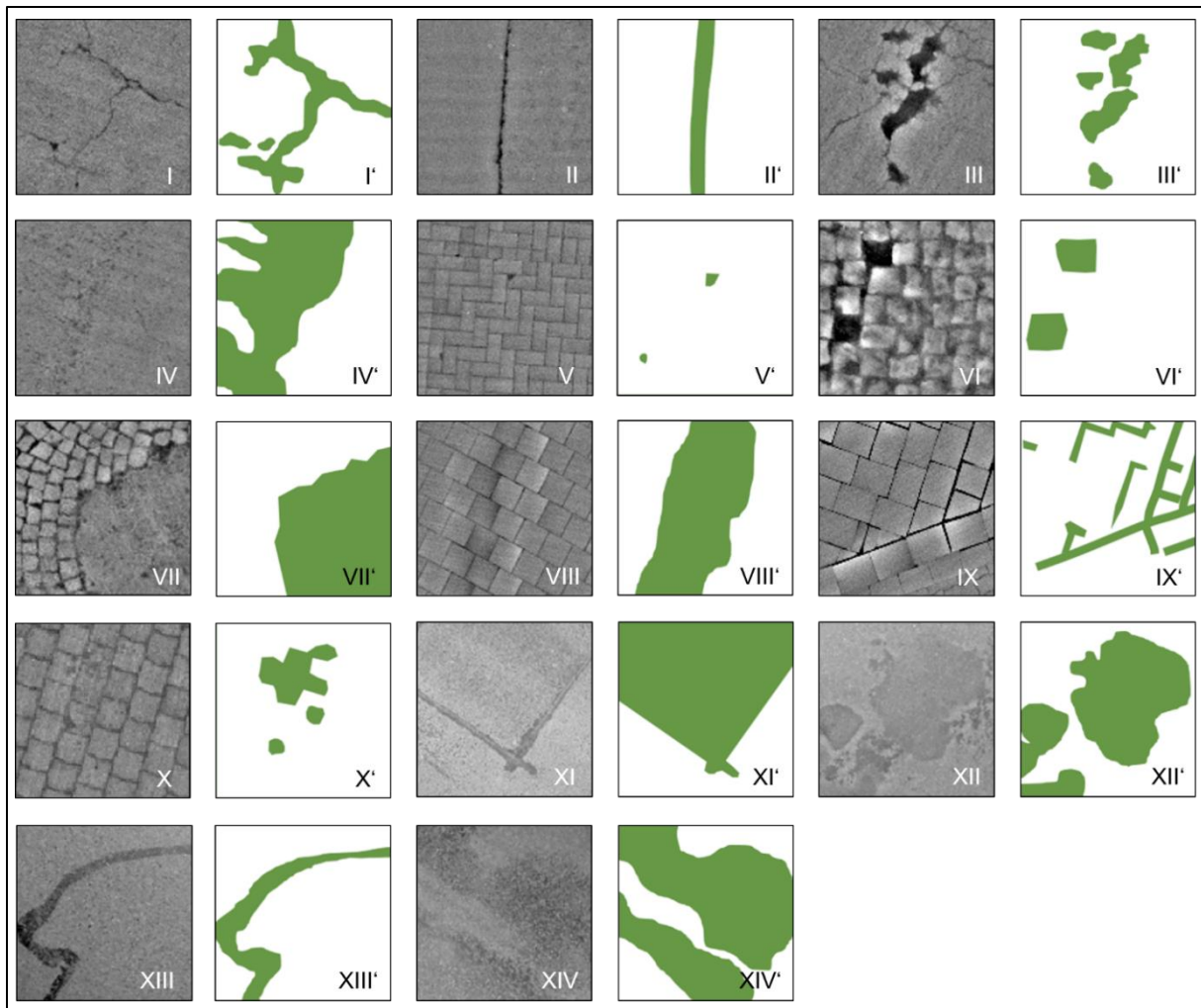


Abbildung 167: Dargestellt sind Beispiele für alle relevanten Oberflächenschäden im TPI-Raster (I bis X) und Intensitätsbild (XI bis XIV) sowie die entsprechende Labelmaske ('). Die Ausschnitte bilden 1 m² Oberfläche ab.

Die Gesamtheit der manuell gelabelten Oberflächenmodelle wurden anschließend in Form von drei Datensätzen organisiert: Z-A/B, Z-P und I-A/B. Die Buchstaben vor dem Bindestrich stehen für den jeweiligen Informationskanal (Höhe *Z* oder Intensität *I*). Die Buchstaben nach dem Bindestrich zeigen an, welche Deckschichtart (Asphalt *A*, Beton *B* oder Pflaster *P*) im Mittelpunkt des jeweiligen Datensatzes steht.

Der Z-A/B-Datensatz basiert auf Z-Oberflächenmodellen und legt den Fokus auf asphalt- und betonspezifische Schäden, die durch ihre Oberflächenmorphologie charakterisiert sind, sowie die Unterscheidung der Deckenarten Asphalt, Beton, Pflaster und Schotter. Er umfasst 1646 Oberflächenmodelle, wovon 228 zur Evaluation separiert wurden (siehe Abbildung 168).

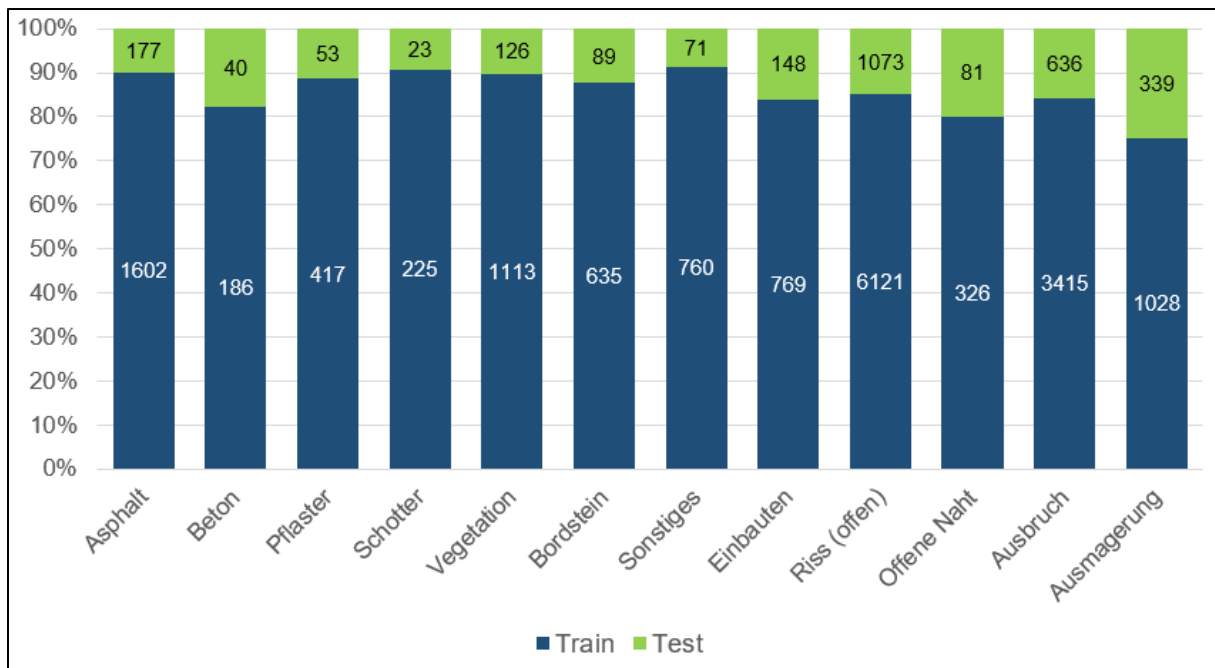


Abbildung 168: Zusammensetzung des Z-A/B-Datensatzes, der Asphalt- und Betonschäden adressiert, die durch ihre Oberflächenmorphologie charakterisiert sind. Die Zahlen entsprechen den gelabelten Objekten.

Der Z-P-Datensatz basiert ebenfalls auf Z-Oberflächenmodellen (siehe Abbildung 169). Er widmet sich explizit den pflasterspezifischen Schäden sowie der weiteren Unterscheidung des Pflasertyps in die Unterkategorien Natursteinpflaster, Betonsteinpflaster und Rasengittersteine. Dieser Datensatz umfasst 369 Oberflächenmodelle, von denen 28 nicht ins Training gingen, sondern zur Evaluation separiert wurden.

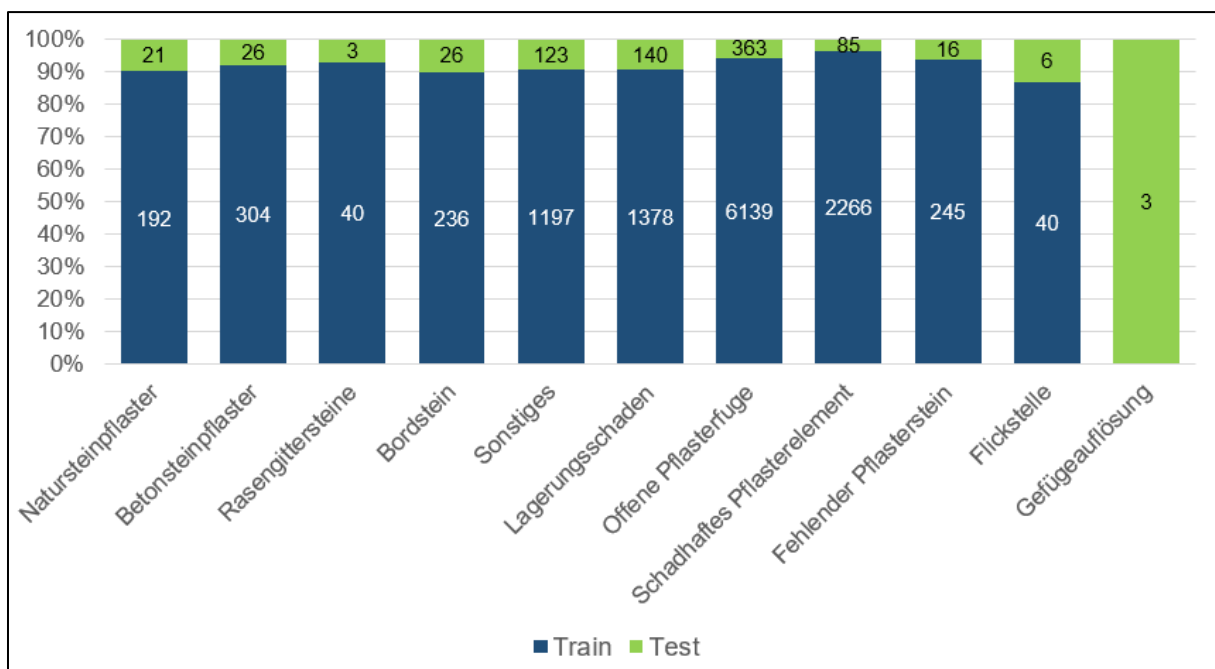


Abbildung 169: Zusammensetzung des Z-P-Datensatzes, der Pflasterschäden adressiert, die durch ihre Oberflächenmorphologie charakterisiert sind. Die Zahlen entsprechen den gelabelten Objekten.

Die Schadklasse Gefügeauflösung im Datensatz Z-P war in der Datenbasis so selten vertreten, dass sie im Training der Modelle nicht verwendet wurde. Der Vollständigkeit halber ist diese Klasse dennoch in Abbildung 169 aufgeführt.

Der dritte und letzte Datensatz (I-A/B) stützt sich auf I-Oberflächenmodelle und beinhaltet jene asphalt- und betonspezifischen Schäden, die sich in erster Linie aufgrund ihrer Reflexionseigenschaften von intakten Fahrbahnbereichen abgrenzen lassen. Er umfasst 1795 Intensitätsbilder, wovon 166 zur Evaluation separiert wurden (siehe Abbildung 170).

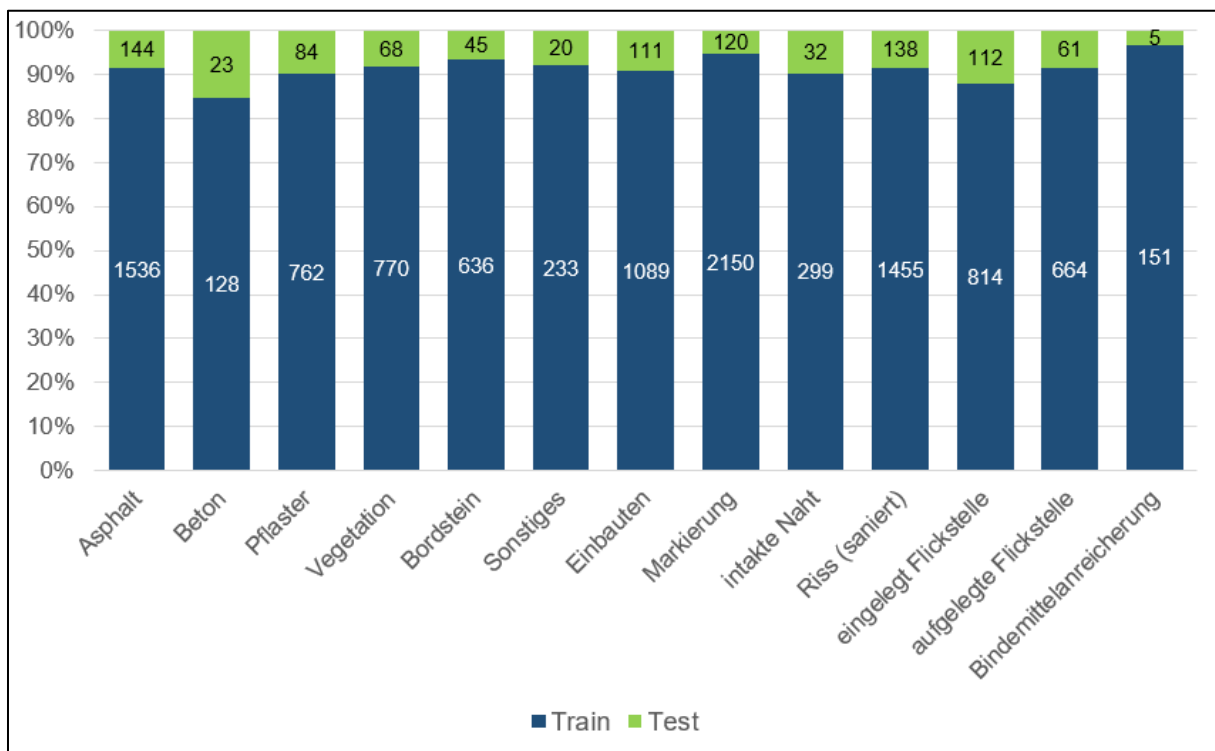


Abbildung 170: Zusammensetzung des I-A/B-Datensatzes, der Asphalt- und Betonschäden adressiert, die durch Reflexionsintensitätsunterschiede charakterisiert sind. Die Zahlen entsprechen den gelabelten Objekten.

Der als Testdaten separierte Anteil der drei Datensätze Z-A/B, Z-P und I-A/B dient dazu, unbekannte Daten zu simulieren und später im Rahmen einer Evaluation trainierte Modelle hinsichtlich ihrer Leistungsfähigkeit bewerten zu können (Hold-Out Splitting). Um Trainings- und Testdaten weitgehend zu entkoppeln, wurden zum Testen nur Daten gewählt, die mehrere Kilometer weit entfernt vom nächstgelegenen Trainingsbeispiel aufgenommen wurden.

Um einschätzen zu können, wie akkurat ein manuelles, pixelgenaues Labelling von Schäden überhaupt sein kann und mit welcher Unsicherheit die manuellen Labels behaftet sind, wurden mehrfach erzeugte Labelmasken miteinander verglichen. Bei STRICKER et al. (2021b) wurde dies mithilfe von drei Experten für eine vergleichbare Problemstellung untersucht. Dort wird das pixelgenaue Labelling von Straßenschäden in hochauflösenden Oberflächenfotos

behandelt. Mangels der Verfügbarkeit unterschiedlicher Experten wurde im Rahmen der vorliegenden Arbeit ein und dasselbe Oberflächenmodell im Laufe des sich über mehrere Monate erstreckenden Labelprozesses mehrfach manuell gelabelt. Zwischen den Zeitpunkten t_1 und t_2 lagen 21 Tage und zwischen t_2 und t_3 35 Tage. Zwar kann ein unbewusstes, subjektives Reproduzieren der vormals gezeichneten Konturen aus der Erinnerung heraus nicht gänzlich ausgeschlossen werden. Jedoch ist dies angesichts von mehr als 7000 gelabelter Risse auf mehr als 1600 Oberflächenmodellen des Z-A/B-Datensatzes nicht sehr wahrscheinlich.

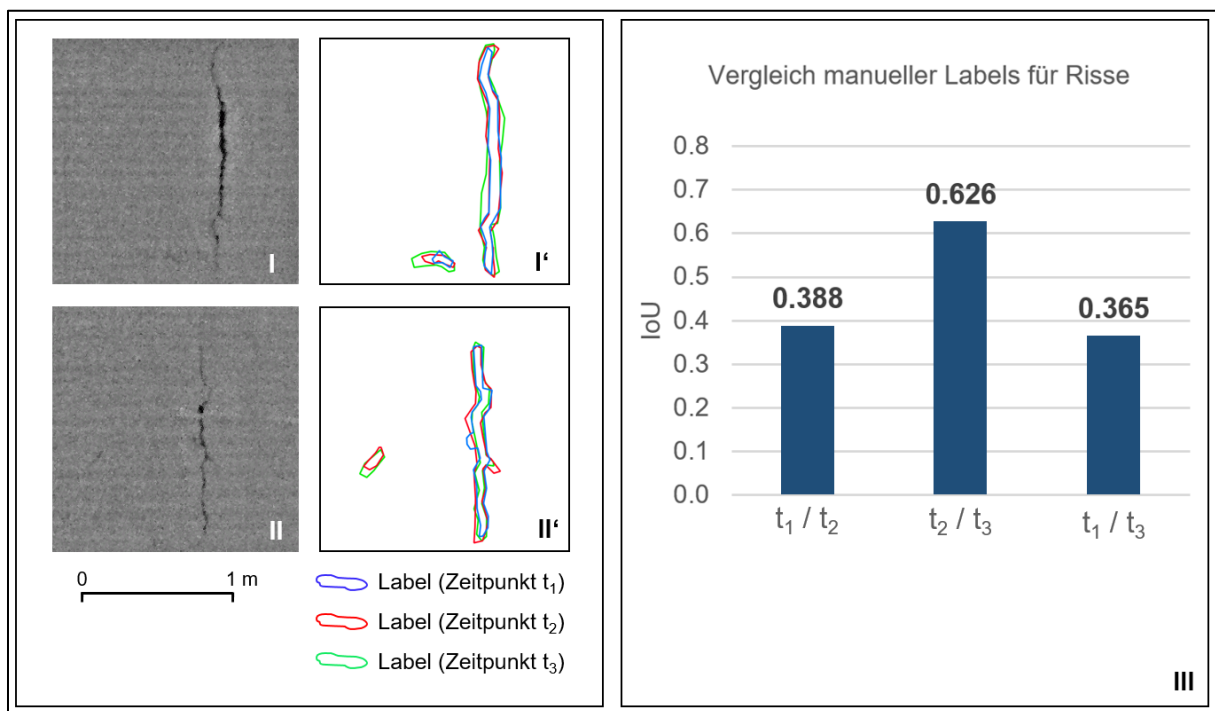


Abbildung 171: Überprüfung der Genauigkeit des Labellings mittels mehrfach manuell annotierter Schadstellen (I' und II'). In der Teilabbildung III ist die Übereinstimmung der Labels anhand der Metrik Intersection over Union (IoU) bewertet, die auch für die Modell-Evaluierung verwendet wird.

Auf dem für den Test selektierten Oberflächenmodell befinden sich zwei deutlich ausgeprägte Risse von ca. 1 m Länge und 5 mm bzw. 15 mm Öffnungsbreite sowie zwei sehr feine Risse von 0,05 m Länge und ca. 2 mm Öffnungsbreite (siehe Abbildung 171). Ansonsten sind keine Schäden vorhanden. Die Situation ist demnach als wenig komplex einzuschätzen. Der Vergleich der verschiedenen Labelzeitpunkte zeigt, dass die Stützpunkte der Konturen nur wenige Zentimeter voneinander entfernt sind und die Konturen in einem ähnlichen Abstand zum Zielobjekt gesetzt wurden (ca. 3 bis 4 cm). In den Labels von Zeitpunkt t_1 wurde einer der beiden kurzen und feinen Risse übersehen und demzufolge nicht gelabelt (siehe Abbildung 171 links unten). Nach visuellen Maßstäben ist die Qualität des Labellings hinsichtlich der Reproduzierbarkeit dennoch als gut einzuschätzen. Um dies zu quantifizieren wurde als Maß

für die Übereinstimmung der Labels jeweils der IoU (siehe Kapitel 4.5.3) berechnet. Die IoU-Werte fallen mit 0,36 bis 0,63 angesichts der qualitativen Einschätzung vergleichsweise niedrig aus. STRICKER et al. (2021b) kamen in ihrem Vergleich verschiedener Auswerter auf mIoU-Werte von 0,21 bis 0,57, wobei hier nicht eine einzelne Klasse, sondern das Mittel über mehrere pixelgenau gelabelte Schadklassen analysiert wurde. Die festgestellten IoU-Werte sind vor allem bei der abschließenden Evaluation der trainierten Modelle bzw. bei der Interpretation der erzielten IoU-Werte hilfreich (siehe Kapitel 5.5.2.2).

Das pixelgenaue Labelling eines einzigen Oberflächenmodells nahm 1 bis 15 Minuten in Anspruch, je nach Komplexität der vorhandenen Oberflächen, Objekte und Schäden.

5.5.2 Training und Evaluation von DeepLabv3+ Modellen

Im Folgenden wird beschrieben, wie mithilfe von Transfer Learning DeepLabv3+ Modelle trainiert werden, die in der Lage sind, PPS-Oberflächenmodelle semantisch zu segmentieren. Auch wenn die Wirkzusammenhänge innerhalb von CNN wissenschaftlich dokumentiert sind basiert ein wesentlicher Bestandteil des Architekturdesigns und des Hyperparametersettings auf Versuch und Irrtum (POULINAKIS et al. 2023). So wurden auch im Rahmen der vorliegenden Arbeit Voruntersuchungen mit dem Ziel durchgeführt, eine Pipeline zum Training von DeepLabv3+ Modellen zu erstellen. Der Begriff Pipeline beschreibt im Bereich des Maschinellen Lernens eine Abfolge von Schritten, die von der Datenextraktion und Datenvorverarbeitung, über das Modelltraining bis hin zur Bereitstellung des Modells reichen können.

Ein limitierender Faktor der Untersuchungen waren die Hardware-Ressourcen, insbesondere die Grafikkarte (GPU) des PCs, die für das Training verwendet wurde (siehe Kapitel 4.1). Um mit der verwendeten Hardware und den vorliegenden Datensätzen zielführend trainieren zu können, mussten Behelfs- bzw. Kompromisslösungen gefunden werden. Deshalb wird zunächst beschrieben, welche Voruntersuchungen durchgeführt wurden, um die finalen Modelle zu trainieren, die den Kern des automatischen Analysesystems zur Kartierung von Oberflächenschäden aus PPS-Oberflächenmodellen darstellen. Der daran anschließende Abschnitt beschreibt das Training und die Evaluation der finalen Modelle.

5.5.2.1 Voruntersuchungen

Obwohl die in Kapitel 5.5.1 beschriebenen Datensätze mehrere tausend semantisch annotierte Beispieldaten enthalten, genügt diese Datenmenge nicht, um derart robuste low- und mid-level-Merkmale zu extrahieren, wie sie beispielsweise auf Basis des ImageNet-Datensatzes gelernt werden können (DENG et al. 2009). Ein Modelltraining von Grund auf, bei dem alle Gewichte im Netz zufällig initialisiert und im Trainingsprozess angepasst werden, wurde daher nicht in Betracht gezogen. Da beispielsweise GOPALAKRISHNAN et al. (2017) Transfer Learning für vergleichbare Erkennungsaufgaben erfolgreich einsetzen, wurde dieser Ansatz auch für die Problemstellung der vorliegenden Arbeit gewählt.

Ein Trainingsprozess unter Nutzung eines mehrere tausend Bilder umfassenden Datensatzes nimmt mit der verfügbaren Hardware ein bis drei Tage in Anspruch. Um möglichst viele Trainingsprozesse durchführen zu können, wurden alle Voruntersuchungen auf reduzierten Z-A/B- und I-A/B-Trainingsdatensätzen durchgeführt (800 Z-Oberflächenmodelle und 839 I-Oberflächenmodelle).

Transfer Learning

Über die Online-Codeverwaltung Github ist der Quellcode der TensorFlow-Implementierung von DeepLabV3+ zugänglich und zur Nutzung unter Apache License 2.0 lizenziert (CHEN et al. 2021). Vortrainierte DeepLab-Modelle können über den TensorFlow DeepLab Model Zoo⁶ bezogen werden. Für die Voruntersuchungen wurden Xception65- und Xception71-Modelle heruntergeladen und als Backbone CNN in DeepLabv3+ eingesetzt. Die Merkmalsextraktion beider CNN ist auf Daten des ImageNet-Datensatzes vortrainiert (DENG et al. 2009) und sie unterscheiden sich im Wesentlichen nur im Hinblick auf ihre Tiefe bzw. Komplexität (65 vs. 71 Schichten).

Hyperparameter

Zunächst mussten die wesentlichen Hyperparameter eingestellt werden, die sich in universelle und architekturspezifische Hyperparameter unterscheiden lassen. Hierzu wurden in ersten Versuchen Trainingsprozesse mit variierenden Hyperparameter-Konfigurationen gestartet und die Segmentierungsergebnisse visuell überprüft. Zu den universellen Hyperparametern zählt die Lernrate. Die Erhöhung der Base Learning Rate vom Standardwert 0,0001 auf 0,0004 führte dazu, dass flächenhafte Klassen weniger lückenhaft segmentiert wurden und die Kontur bei

⁶ Website: https://github.com/tensorflow/models/blob/master/research/deeplab/g3doc/model_zoo.md#checkpoints-pretrained-on-imagenet

eher linienhaften Klassen engumschließender verlief. Um mit der verfügbaren Hardware trainieren zu können, mussten die Dimensionen der Eingangskanäle auf 1025 x 1025 Pixel skaliert und eine Batch Size von 2 gewählt werden. Die Batch Normalization konnte hardwarebedingt nicht mit trainiert werden. Zu den Hyperparametern, die spezifisch für die DeepLabV3+ Architektur sind, zählen die Parameter Atrous Rate und Output Stride. Beide Werte lassen sich sowohl beim Training als auch bei der Anwendung eines trainierten Modells modifizieren. Für das Training der Modelle wurden die empfohlenen Werte aus dem Quellcode übernommen. Für die Anwendung des trainierten Modells jedoch wurden die Pixelwerte für die Atrous Rates von 12, 24 und 36 auf 6, 12 und 18 sowie der Output Stride auf 8 gesetzt. Damit ging eine Verbesserung der Segmentierungsgenauigkeit einher.

Hinsichtlich des Backbone CNN konnte festgestellt werden, dass die Verwendung des tieferen Xception71-Netzes keinen signifikanten positiven Einfluss auf die Segmentierungsergebnisse hatte, dafür aber deutlich längere Trainingszeiten benötigte. Deshalb wurde für alle weiteren Trainingsprozesse ein auf ImageNet-Daten vortrainiertes Xception65 CNN zur Merkmalsextraktion eingesetzt. Tabelle 16 zeigt die Hyperparameter, mit denen alle Modelle trainiert wurden.

Tabelle 16: Übersicht über die wesentlichen Hyperparameter, die zum Training der DeepLabV3+ Modelle verwendet wurden.

Hyperparameter	Training des Modells	Anwendung des Modells
Base Learning Rate	0,0004	-
Crop Size	1025	1025
Steps (Iterations)	150000 - 300000	-
Batch Size	2	1
Fine Tune Batch Normalization	False	-
Backbone Model Variant	xception_65_imagenet	xception_65_imagenet
Atrous Rate 1	6	6
Atrous Rate 2	12	12
Atrous Rate 3	18	18
Output Stride	16	8
Decoder Output Stride	4	4

Streuung von Trainingsprozessen

Nachdem die Grundlagen geschaffen waren, um DeepLabV3+ Modelle auf dem verfügbaren Datenmaterial und mit der verfügbaren Hardware zu trainieren, sowie die ersten erzeugten Segmentierungsergebnisse die Annahme bestärkten, dass der Transfer-Learning-Ansatz zielführend ist, musste der Zeitpunkt für das Ende des Trainings bestimmt werden. Bei üppig dimensionierten Hardware-Ressourcen stellt sich diese Frage nicht zwingend, denn das

Training wird dann beendet, wenn keine Verbesserung des Modells auf dem Test- oder Validierungsdatensatz mehr festzustellen ist. TensorFlow stellt diese Early-Stopping-Methode auch standardmäßig zur Verfügung, jedoch ist es dazu nötig, in kurzen Intervallen den aktuellen Stand des trainierten Netzes (engl.: Checkpoint) zusätzlich in den Grafikspeicher zu laden und diesen Modell-Checkpoint auf die Trainings- und Testdaten anzuwenden. Dies war mit der verfügbaren Hardware und der gewählten Hyperparameter-Konfiguration nicht möglich.

Deshalb wurde zunächst in einem Experiment untersucht, wie stark die Streuung der Bewertungsmaße F1, IoU und MCC ist, wenn mehrere Modelle auf denselben Datensatz für die gleiche Anzahl an Iterationen trainiert werden, um mit dieser Erkenntnis in einem Folgeexperiment zu untersuchen, ab welcher Iterationszahl keine wesentliche Verbesserung der Leistungsfähigkeit auf dem Testdatensatz mehr eintritt.

Tabelle 17: Streuung der Modellperformance aufgrund zufälliger Initialisierungen während des Trainings von drei identisch parametrisierten Prozessen (V1, V2 und V3).

Iteration	Training	F1	mIoU	MCC
150K	V1	0,949	0,913	0,832
	V2	0,950	0,915	0,838
	V3	0,952	0,918	0,842
	Spannweite	0,003	0,005	0,011
300k	V1	0,951	0,916	0,838
	V2	0,951	0,916	0,841
	V3	0,953	0,919	0,846
	Spannweite	0,003	0,004	0,008

Da mit dem Transfer-Learning-Ansatz nur vergleichsweise wenige Parameter zufällig initialisiert werden, wurde eine vergleichsweise niedrige Streubreite der Trainingsergebnisse angenommen. Um sie zu quantifizieren, wurden nacheinander drei Trainingsprozesse auf den im Umfang reduzierten I-A/B-Datensatz gestartet. Für jedes Training (V1, V2 und V3) wurde nach 150000 und 300000 Iterationen ein Modell exportiert, auf den Testdatensatz angewendet und die in Kapitel 4.5.3 beschriebenen Metriken berechnet. Die entsprechenden Scores sowie deren Spannweite sind in Tabelle 17 aufgeführt. Am Beispiel dieser Trainingsprozesse lässt sich auch demonstrieren, wie die Erkenntnis aus diesem Experiment genutzt werden kann: Das

Modell V1 erreicht nach 150000 Iterationen einen mIoU von 0,913 und nach weiteren 150000 Iterationen einen mIoU von 0,916. Auf den ersten Blick ist eine Verbesserung des Modells eingetreten, jedoch liegt sie mit 0,003 Score-Punkten innerhalb der Spannweite von 0,004, die auf zufällige Initialisierungsprozesse während des Trainings zurückzuführen ist. Die festgestellten Spannweiten wurden u. a. in der Untersuchung zur Trainingsdauer berücksichtigt, die nachfolgend beschrieben ist.

Ende des Trainings

Da hardwarebedingt das Trainingsende nicht mithilfe von Early Stopping bestimmt werden konnte, wurde auf Basis des reduzierten I-A/B-Datensatzes untersucht, ab welcher Iterationszahl keine wesentliche Verbesserung der Leistungsfähigkeit auf dem Testdatensatz mehr festzustellen ist. Dazu wurden während des Trainings an sechs willkürlich festgelegten Zeitpunkten Modell-Checkpoints exportiert: nach 40k, 60k, 120k, 150k, 300k und 500k Iterationen (k steht nachfolgend für den Faktor 1000). Nach dem letzten Checkpoint-Export wurden die jeweiligen Modelle nacheinander auf dem Testdatensatz evaluiert.

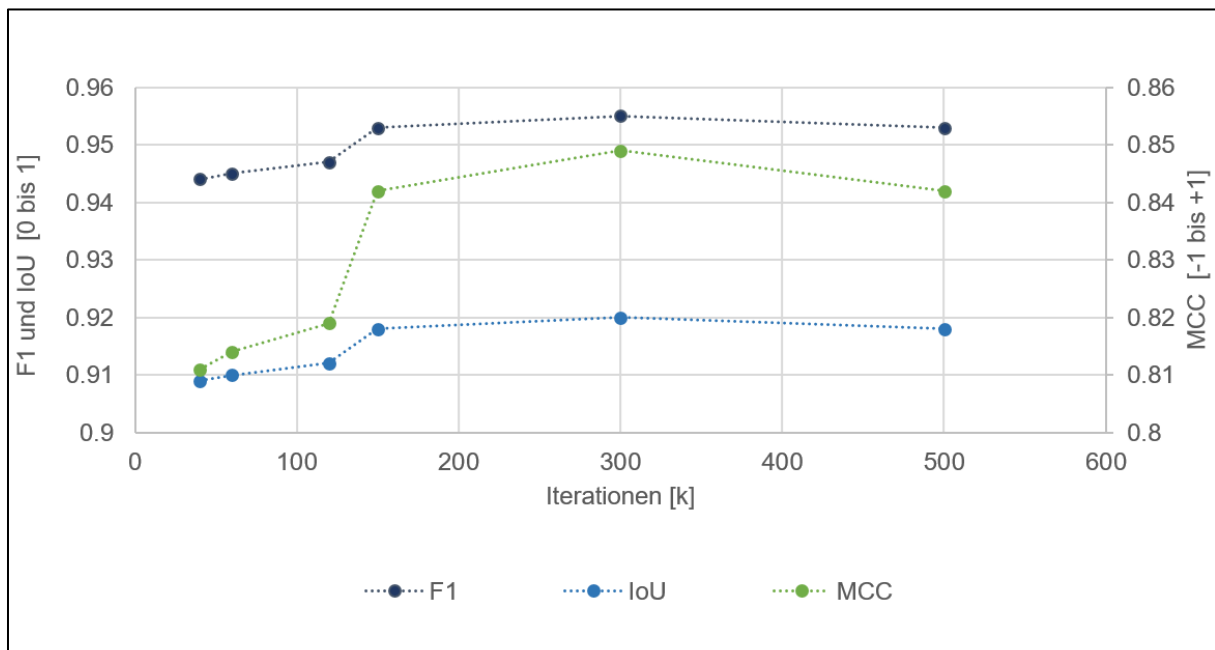


Abbildung 172: Einfluss der Trainingsdauer auf die Leistungsfähigkeit des Modells. Alle untersuchten Metriken zeigen einen Sprung bei 150k Iterationen. Spätestens ab 300k Iterationen beginnen die Scores zu sinken.

Abbildung 172 zeigt das Ergebnis dieser Voruntersuchung. Alle drei verwendeten Metriken zeigen ein Maximum bei 300k Iterationen. Bei 500k Iterationen kann bereits ein Sinken der Scores festgestellt werden. Da bei 150k Iterationen ein vergleichsweise großer Leistungszugewinn erreicht werden konnte, wurden die weiteren Trainingsprozesse der

Voruntersuchungen bei 150k Iterationen beendet. Aufgrund der willkürlich festgesetzten Zeitpunkte ist nicht sichergestellt, dass bei allen nachfolgenden Trainingsprozessen tatsächlich immer der optimale Zeitpunkt für das Ende des Trainings genutzt wurde. Es ist aufgrund der hohen Iterationszahl nicht auszuschließen, dass die trainierten Modelle bereits dazu tendieren, sich auf die Trainingsdaten überanzupassen.

Geomorphometrische Vorverarbeitung der Netzeingaben

Ein weiteres Experiment zielte darauf ab, herauszufinden, ob eine der vier geomorphometrischen Ansätze, die zum manuellen Labelling der Beispieldaten im Z-A/B- und Z-P-Datensatz verwendet wurden, für die finale Trainingspipeline zu favorisieren ist oder ob alle Ansätze vergleichbare Modelle hervorbringen.

Zu diesem Zweck wurden auf Basis des reduzierten Z-A/B-Datensatzes je ein Modell für 150k Iterationen trainiert, für dessen Eingabe Hillshade Raster, Topographic Position Index Raster, Positive Openness Raster und Black Top Hat Transform Raster verwendet wurden. Die Labelmasken wurden unter Berücksichtigung aller vier Ansätze im Labelling erzeugt, weshalb jeweils nur eine Labelmaske für die vier Netzeingabeoptionen existiert.

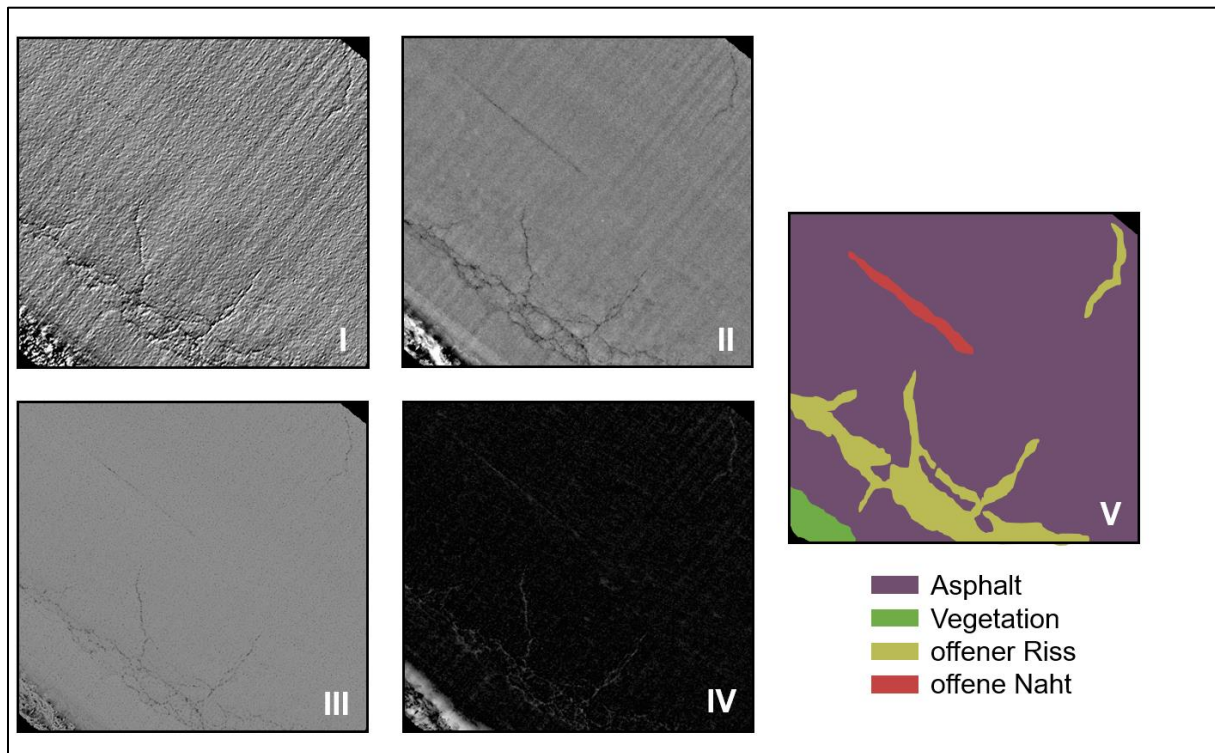


Abbildung 173: Vier verschiedene geomorphometrische Raster als Eingabeoptionen für das Training von Modellen. Untersucht wurden Hillshade Raster (I), Topographic Position Index Raster (II), Positive Openness Raster (III) und Black Top Hat Transform Raster (IV). Die Labelmasken (V) wurden unter Berücksichtigung aller vier Ansätze im Labelling erzeugt.

In Abbildung 173 sind die vier Eingabeoptionen exemplarisch dargestellt. Bereits im Zuge des manuellen Labellings wurde deutlich, dass nicht alle Situationen auf allen vier Analyserastern gleich gut eingeschätzt werden konnten. So hängt beispielsweise bei Hillshades die Sichtbarkeit von Schäden stark von ihrer relativen Orientierung zur virtuellen Beleuchtungsquelle ab. In Anhang IX, Anhang X, Anhang XI und Anhang XII ist jeweils die Konfusionsmatrix der trainierten Modelle aufgeführt. Wie in Tabelle 18 dargestellt, erzielt das Modell auf Basis des Topographic Position Index die besten Scores bei den relevanten Metriken. Die Unterschiede zu den anderen Ansätzen sind aber, unter Berücksichtigung der in Tabelle 17 aufgeführten zufallsbedingten Streubreite, eher gering.

Tabelle 18: Vergleich von Modellen, die auf Basis unterschiedlicher geomorphometrischer Vorverarbeitungen des Oberflächenmodells für jeweils 150k Iterationen trainiert wurden. In der Kopfzeile steht bei den Metriken in Klammern jeweils die zuvor ermittelte Spannweite der zufallsbedingten Streuung von Trainingsprozessen.

geomorphometrische Verarbeitung des Z-Oberflächenmodells	F1 (0,003)	mIoU (0,005)	MCC (0,011)
Topographic Position Index	0,974	0,939	0,867
Topographic Positive Openness	0,970	0,935	0,846
Black Top Hat Transform	0,967	0,932	0,832
Hillshade	0,964	0,929	0,821

Für die finale Trainingspipeline wurde der Z-A/B- und der Z-P-Datensatz aufgrund der Ergebnisse des vorstehend beschriebenen Experimentes auf TPI-Raster trainiert. Diese Raster enthalten, wie in Kapitel 4.3.4 beschrieben, Gleitkommazahlen, die negativ oder positiv sein können. Um DeepLabv3+ mit TPI-Rastern trainieren zu können, mussten die Gleitkommawerte in 8bit-Werte überführt werden. Unter der Prämisse, dass die Grauwert-Schritte des 8bit-Zielbildes wenigstens 0,5 mm darstellen können sollten, wurden Versuche mit verschiedenen Spreizungen (engl.: Stretch) durchgeführt (siehe Abbildung 174). Dabei erwies sich ein Minimalwert von -0,01 m und ein Maximalwert von +0,01 m als am besten geeignet. Bei den 256 verfügbaren Werten eines RGB-Bildkanals löst eine einzelne Grauwertstufe 0,08 mm auf. Schäden, wie Risse, offene Nähte, Ausmagerungen und Ausbrüche sind visuell eindeutig zu erkennen. Der verwendete Stretch-Bereich ermöglicht es, dass auch andere Objektklassen, wie z. B. Bordsteine, noch hinreichend genau abgegrenzt bzw. lokalisiert werden können.

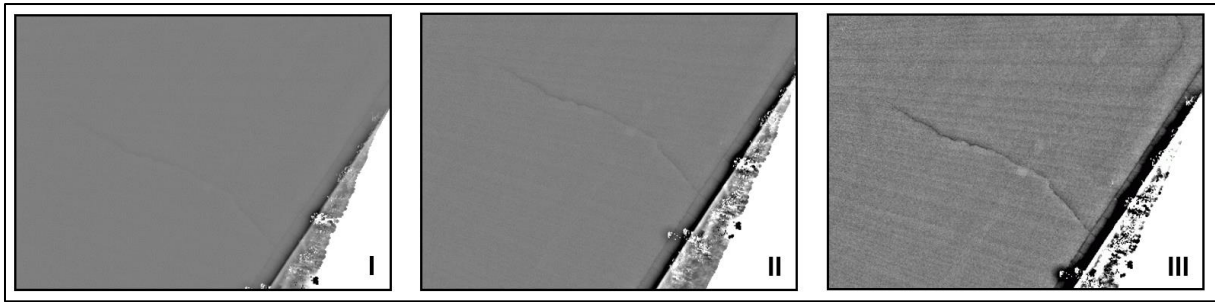


Abbildung 174: Verschiedene Stretch-Werte, um die Gleitkommazahlen der TPI-Raster für die Eingabe in DeepLabv3+ auf 256 Grauwerte abzubilden. Teilabbildung I zeigt eine Spreizung von +0,1 m bis -0,1 m. Teilabbildung II zeigt eine Spreizung von +0,04 m bis -0,04 m. Teilabbildung III zeigt eine Spreizung von +0,01 m bis -0,01 m.

5.5.2.2 Finale Trainingspipeline

Mit den im Rahmen der Voruntersuchungen gewonnenen Erkenntnissen wurden die finalen Trainingsprozesse durchgeführt. Dabei wurden unter Verwendung der beschriebenen Hyperparameter (siehe Tabelle 16) insgesamt drei DeepLabv3+ Modelle auf Basis der drei vollständigen Trainingsdatensätze trainiert (siehe Kapitel 5.5.1). Das teilweise gravierende Ungleichgewicht der Klassenbesetzung kann beim Training von DeepLabV3+ Modellen durch das Verwenden von Klassengewichten berücksichtigt werden. Diese Gewichte werden jeweils aus dem Kehrwert der Pixelanzahl der jeweiligen Klasse dividiert durch die Gesamtpixelanzahl im Datensatz berechnet und zur Berechnung des Loss' in der Kostenfunktion automatisch angebracht. So wird die Klasse Riss, die trotz der hohen Objektanzahl aufgrund ihrer Form- und Labelcharakteristik nur einen kleinen Anteil aller im Datensatz enthaltenen Pixelwerte ausmacht, mit einem Faktor von 280 gewichtet. Die Klasse Asphalt, die zwar hinsichtlich der Objektzahl seltener, aber dafür hinsichtlich der Pixel Anzahl wesentlich stärker vertreten ist, wurde mit dem Faktor 3 gewichtet. Damit werden Fehlklassifikationen von vergleichsweise selten vorkommenden Klassen im Trainingsprozess über die Loss-Funktion härter bestraft und das Lernen dieser Klassen ermöglicht.

Das Z-A/B-Modell und das Z-P-Modell wurden für 200k Iterationen, das I-A/B-Modell für 300k Iterationen trainiert. Ein Trainingsprozess für ein einziges Modell dauerte mit der verwendeten Hardware und den eingestellten Hyperparametern ca. 50 Stunden.

Um die drei finalen Modelle Z-A/B, I-A/B und Z-P zu testen, wurden sie auf die Oberflächenmodelle des jeweiligen Testdatensatzes angewendet und die Ergebnismasken mit den manuell gelabelten Referenzmasken verglichen. Die Anwendung eines Modells auf ein Eingabeoberflächenmodell dauert unabhängig von der Komplexität der vorhandenen Schäden, Materialien und Objekte ohne weitere softwareseitige Optimierungen weniger als 1 Sekunde.

Die Abbildung 175, Abbildung 176 und Abbildung 177 zeigen jeweils die normalisierte Konfusionsmatrix, die die Leistungsfähigkeit der Modelle Z-A/B, Z-P und I-A/B visualisiert. Die Werte in den Zellen der Hauptdiagonalen entsprechen dem Recall-Wert, der die Vollständigkeit der Segmentierung pro Klasse quantifiziert. Auffällig ist bei allen drei Modellen die starke Besetzung der Hauptdiagonale. Dies deutet auf eine brauchbare Klassifikationsleistung der Modelle auf den unabhängigen Testdatensätzen hin. Die globalen Metriken für die Modelle (siehe Tabelle 19) unterstützen diese Aussage.

Tabelle 19: Globale Bewertung der trainierten Modelle anhand dreier Metriken.

	Z-A/B-Modell	Z-P-Modell	I-A/B-Modell
F1	0,960	0,954	0,962
mIoU	0,928	0,918	0,932
MCC	0,924	0,906	0,919

Die Klassifikation der Deckschichtarten Asphalt, Beton, Pflaster und Schotter, sowie die Unterteilung von Pflasterarten im Z-P-Modell, funktioniert sehr zuverlässig. Ähnliches gilt für Objektklassen, wie Fahrbahnmarkierungen, Bordsteine oder Einbauten. Bei den Schadklassen zeigt sich folgendes Ergebnis: Tendenziell flächenhaft und scharf abgrenzbare Schäden, wie beispielsweise eingelegte Flickstellen, sind durch hohe Recall-Werte gekennzeichnet. Eher kleinräumig oder linienhaft auftretende Schäden, wie z. B. Risse, Nahtrisse oder Ausbrüche, weisen dagegen deutlich geringere Recall-Werte auf. Verwechslungen bzw. Fehlklassifikationen betreffen häufig genau die Grenzfälle, die auch im manuellen Auswerteprozess schwer zu entscheiden sind.

Z-A/B-Modell

In Bezug auf die Schadklassen sind im Z-A/B-Modell folgende Fehlklassifikationen festzustellen: Kleinräumig oder linienhaft auftretende Schäden werden am häufigsten mit den Deckenarten, auf denen sie vorkommen, verwechselt (Asphalt, Beton). Dies können Fälle sein, in denen Schadstellen tatsächlich komplett übersehen werden, oder in denen die Modelle enganschließendere Objektgrenzen erzeugen, als die während des Labellings um die betreffenden Schadstellen herum gezeichneten Masken aufweisen.

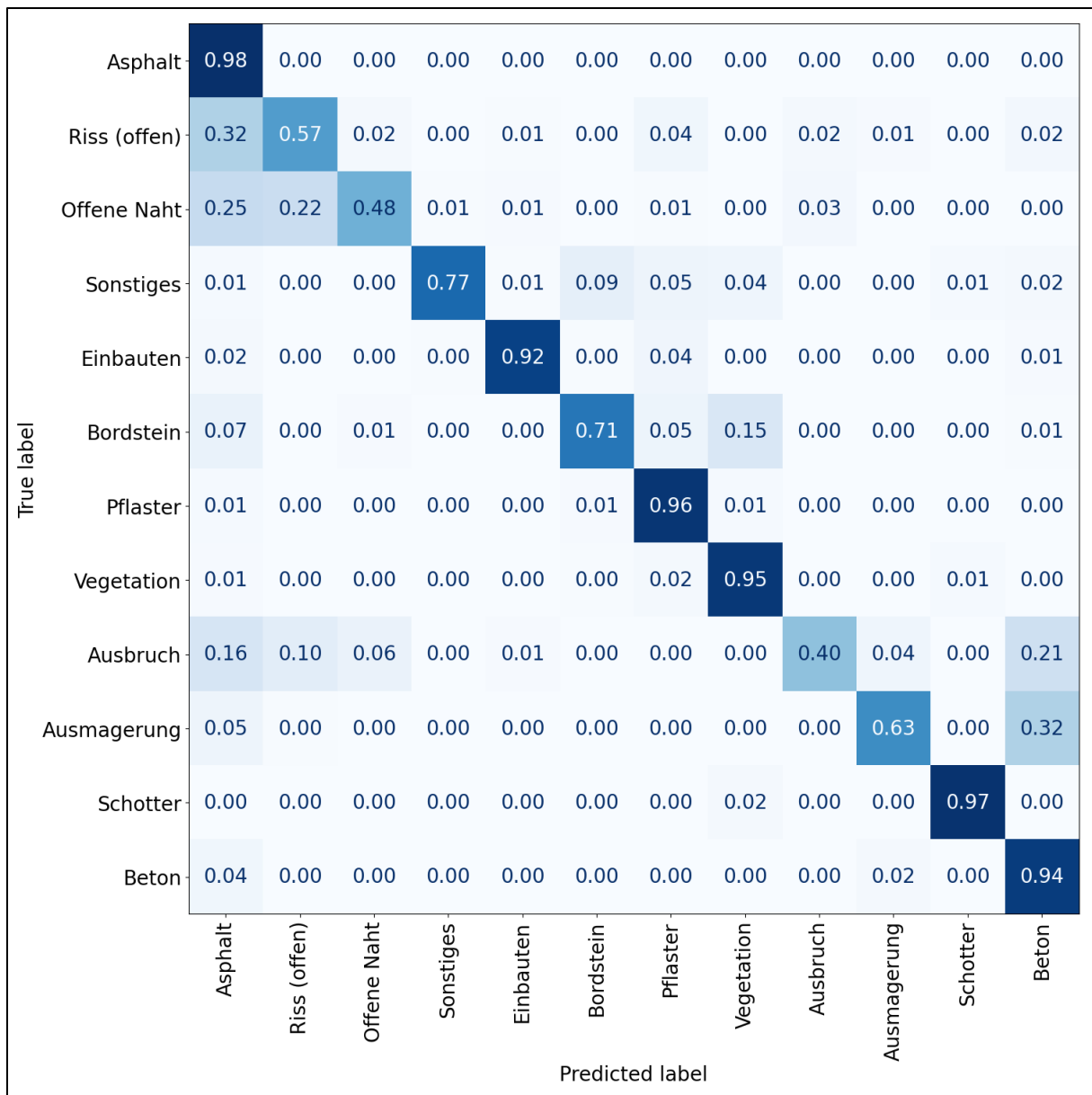


Abbildung 175: Konfusionsmatrix des Z-A/B-Modells zur Analyse von asphalt- und betonspezifischen Schäden auf Basis der Höheninformation. Es erreicht einen globalen IoU Score von 0,928, einen F1 Score von 0,960 und einen MCC Score von 0,924 auf dem Testdatensatz.

Pixel, die als offene Naht gelabelt sind, werden häufig mit offenen Rissen verwechselt. Dies ist insofern nachvollziehbar, als dass sich die beiden Klassen morphologisch nur darin unterscheiden, dass offene Nähte typischerweise einen geradlinigeren Verlauf haben als Risse. Für die Pixel, die als Ausbruch gelabelt sind, kommen Fehlklassifikationen mit den Schadklassen vor, mit denen Ausbrüche oft gemeinsam auftreten bzw. aus denen sich Ausbrüche im Laufe der Zeit entwickeln (offene Risse, offene Nähte und Ausmagerungen).

True label	Asphalt	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
	eingelegte Flickstelle	0.14	0.79	0.00	0.00	0.01	0.00	0.01	0.01	0.00	0.00	0.02	0.00	0.02
	Markierung	0.05	0.00	0.93	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00
	Riss (saniert)	0.21	0.03	0.00	0.63	0.01	0.00	0.04	0.03	0.00	0.00	0.01	0.00	0.04
	Pflaster	0.01	0.00	0.00	0.00	0.94	0.02	0.00	0.00	0.03	0.00	0.00	0.00	0.00
	Vegetation	0.01	0.00	0.00	0.00	0.08	0.90	0.00	0.00	0.01	0.00	0.00	0.00	0.00
	Einbauten	0.01	0.02	0.00	0.00	0.01	0.00	0.95	0.00	0.00	0.01	0.00	0.00	0.00
	intakte Naht	0.28	0.06	0.00	0.12	0.00	0.00	0.00	0.52	0.00	0.00	0.02	0.00	0.00
	Bordstein	0.08	0.00	0.04	0.00	0.07	0.08	0.00	0.00	0.71	0.00	0.00	0.00	0.02
	Sonstiges	0.03	0.00	0.00	0.00	0.01	0.05	0.01	0.00	0.03	0.80	0.00	0.00	0.06
	aufgelegte Flickstelle	0.05	0.04	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.81	0.00	0.06
	Bindemittel-anreicherung	0.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.73	0.00
	Beton	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.96
		Asphalt	eingelegte Flickstelle	Markierung	Riss (saniert)	Pflaster	Vegetation	Einbauten	intakte Naht	Bordstein	Sonstiges	aufgelegte Flickstelle	Bindemittel-anreicherung	Beton
		Predicted label												

Abbildung 176: Konfusionsmatrix des I-A/B-Modells zur Analyse von asphalt- und betonspezifischen Schäden auf Basis der Intensitätsinformation. Es erreicht einen globalen IoU Score von 0,932, einen F1 Score von 0,962 und einen MCC Score von 0,919 auf dem Testdatensatz.

I-A/B-Modell

In Bezug auf die Schadklassen, sind im I-A/B-Modell folgende Fehlklassifikationen festzustellen: sanierte Risse werden teilweise als intakte Nähte oder eingelegte Flickstellen fehlklassifiziert. Dies ist insofern nachvollziehbar, als dass bei allen drei Klassen Bitumen verwendet wird, um entweder einen offenen Riss zu versiegeln, eine baubedingte Naht zu vergießen oder die Schnittfuge einer flächenhaften Instandsetzung abzudichten. Auch um Schachtdeckel herum wird häufig ein bitumenhaltiges Fugenband aufgebracht, weshalb auch Fehlklassifikationen zur Klasse Einbauten erklärt werden können. Die Unterscheidung

zwischen eingelegten und aufgelegten Flickstellen führt zu brauchbaren Ergebnissen. Bindemittelanreicherungen, die nur auf Asphaltdecken auftreten, werden typischerweise als intakter Asphalt fehlklassifiziert. Die Gründe sind hierbei vor allem in der unscharfen Abgrenzbarkeit des Schadens zu suchen. Ferner ähneln Asphaltbereiche, die von abtrocknender Nässe betroffen sind, optisch Oberflächen, bei denen Bindemittel austritt.

Z-P-Modell

Das Z-P-Modell basiert auf einem deutlich kleineren Datensatz, als das Z-A/B-Modell und das I-A/B-Modell. Die globalen Metriken für dieses Modell sind im Vergleich zu den beiden anderen Modellen auch etwas niedriger (siehe Tabelle 19). Die Recall-Werte der Klasse schadhaftes Pflasterelement sind mit 0,19 sehr niedrig und es treten häufig Fehlklassifikationen mit der Deckenart auf, auf der diese Schäden vorkommen (Betonsteinpflaster). Es handelt sich aber auch um die mit Abstand am feinsten gelabelte Schadklasse (siehe Abbildung 167). Außerdem treten Verwechslungen zu den anderen pflasterspezifischen Schäden auf, vornehmlich zu Lagerungsschäden. Lagerungsschäden wiederum werden auch in den meisten Fällen als die Deckenart fehlklassifiziert, auf der sie typischerweise entstehen (Betonsteinpflaster). Da Lagerungsschäden und offene Pflasterfugen ursächlich zusammenhängen, sind die Verwechslungen dieser Klassen nachvollziehbar. Die Fehlklassifikation von offenen Pflasterfugen als Natursteinpflaster kann dadurch erklärt werden, dass die Fugen gerade bei sehr altem Natursteinpflaster (Kopfsteinpflaster) etwas breiter sind und oft nicht mehr bis zur Steinoberkante verfüllt sind. Flickstellen auf Pflasterdecken werden sehr zuverlässig erkannt. Sie sind von allen pflasterspezifischen Schäden auch die, die am präzisesten abgrenzbar sind und dazu eher großflächig auftreten. Fehlende Pflastersteine sind eher kleinräumig, aber auch klar abgrenzbar. Neben den Verwechslungen mit den beiden Pflasterarten, bei denen fehlende Steine typischerweise vorkommen, wird diese Klasse auch als schadhaftes Pflasterelement oder offene Pflasterfuge fehlklassifiziert. Die Verwechslungen beim Z-A/B-Modell sind mutmaßlich darauf zurückzuführen, dass der Datensatz angesichts der Vielzahl möglicher Ausprägungen von Pflasterdecken und damit auch -schäden noch zu klein ist.

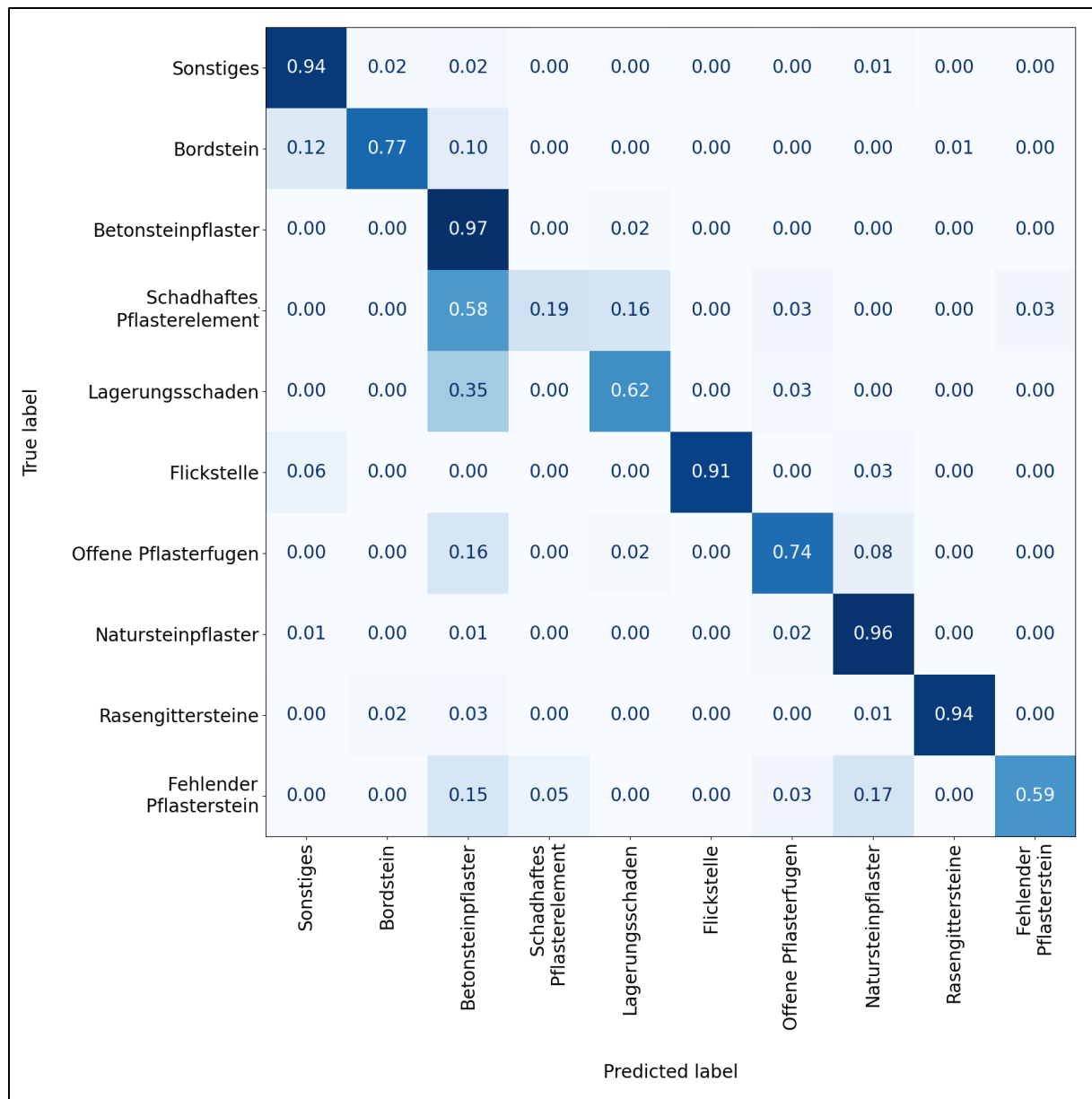


Abbildung 177: Konfusionsmatrix des Z-P-Modells zur Analyse von pflasterspezifischen Schäden auf Basis der Höheninformation. Es erreicht einen globalen IoU Score von 0,918, einen F1 Score von 0,954 und einen MCC Score von 0,906 auf dem Testdatensatz.

In Tabelle 20 sind die Ergebnisse in Form der klassenweisen IoU Scores aufgelistet. Die Segmentierung der ungeschädigten Deckschichtarten Asphalt, Beton, Pflaster und Schotter, ist mit IoU Werten zwischen 0,85 und 0,97 zuverlässig. Dies gilt sowohl für das Z-A/B- als auch für das I-A/B-Modell. Die Trennung von Pflasterarten (Z-P-Modell) ist ebenfalls auf einem hohen Niveau möglich. Ähnliches gilt für Objektklassen, wie Fahrbahnmarkierungen, Vegetation, Bordsteine, Einbauten und sonstige Objekte (z. B. parkende Fahrzeuge), wobei hier die IoU-Werte mit 0,40 bis 0,86 bereits etwas niedriger liegen.

Bei den Schadklassen zeigt sich ein ähnliches Ergebnis, wie zuvor anhand der Konfusionsmatrizen beschrieben: Flächenhaft auftretende und scharf abgrenzbare Schäden, wie

z. B. Flickstellen, sind durch IoU Werte zwischen 0,57 und 0,87 gekennzeichnet. Schadklassen, wie z. B. Ausmagerungen, Lagerungsschäden und Bindemittelanreicherungen, die sich zwar flächenhaft ausbilden, aber nur unscharf abgrenzbar sind, weisen IoU-Werte zwischen 0,39 und 0,59 auf. Die niedrigsten IoU-Werte entfallen auf punkthaft, kleinräumig oder linienhaft auftretende Schäden, wie z. B. Risse, offene Nähte oder Pflasterfugen, Ausbrüche und schadhafte Pflasterelemente. Dies bedeutet aber nicht, dass diese Schäden unzureichend detektiert werden. Insbesondere bei eher kleinräumigen, dünnen und linearen Strukturen hängt der IoU stark davon ab, in welchem Abstand um den Schaden herum das Referenzlabel manuell erzeugt wurde und in welchem Abstand um den Schaden herum die Prädiktion durch das Modell erfolgt. Dieser Effekt wurde bereits in der Untersuchung zur Genauigkeit der manuellen Labels deutlich (siehe Abbildung 171): IoU-Werte können relativ gering ausfallen, wenn die Objektgrenzen mit einem unterschiedlich breiten Saum um die Mittellinie eines Schadens herum gesetzt wurden, auch wenn die Lokalisierung des Schadens an sich korrekt ist.

Tabelle 20: Evaluation der drei trainierten Modelle auf dem unabhängigen Testdatensatz mittels globaler und klassenweiser Intersection over Union (IoU).

Z-A/B-Modell		Z-P-Modell		I-A/B-Modell	
mIoU: 0,928		mIoU: 0,918		mIoU: 0,932	
Klassen	IoU	Klasse	IoU	Klasse	IoU
Asphalt	0,964	Natursteinpflaster	0,941	Asphalt	0,967
Beton	0,910	Betonsteinpflaster	0,952	Beton	0,945
Pflaster	0,920	Rasengittersteine	0,917	Pflaster	0,856
Schotter	0,888	Bordstein	0,618	Vegetation	0,826
Vegetation	0,829	Sonstiges	0,856	Bordstein	0,515
Bordstein	0,549	Lagerungsschaden	0,402	Sonstiges	0,634
Sonstiges	0,654	Offene Pflasterfuge	0,317	Einbauten	0,798
Einbauten	0,779	Schadhaftes Pflasterelement	0,133	Markierung	0,824
Riss (offen)	0,442	Fehlender Pflasterstein	0,496	intakte Naht	0,403
Offene Naht	0,299	Flickstelle	0,871	Riss (saniert)	0,418
Ausbruch	0,261	Gefügebrauflösung	-	ingelegt Flickstelle	0,750
Ausmagerung	0,399			aufgelegte Flickstelle	0,575
				Bindemittelanreicherung	0,592

5.5.3 Automatische Kartierung von Oberflächenschäden

Um die DeepLabv3+ Segmentierungsergebnisse, die im Bildkoordinatensystem des jeweiligen Oberflächenmodells vorliegen, in eine georeferenzierte Kartierung (Vektordaten) zu überführen und letztlich eine Ergebnisdatei im Sinne der Straßenzustandsbewertung (siehe

Kapitel 2.1.3.3) zu erzeugen, wurde eine angepasste Verarbeitungs-Pipeline entwickelt. Im Folgenden wird dieser Ablauf vorgestellt (siehe Abbildung 178).

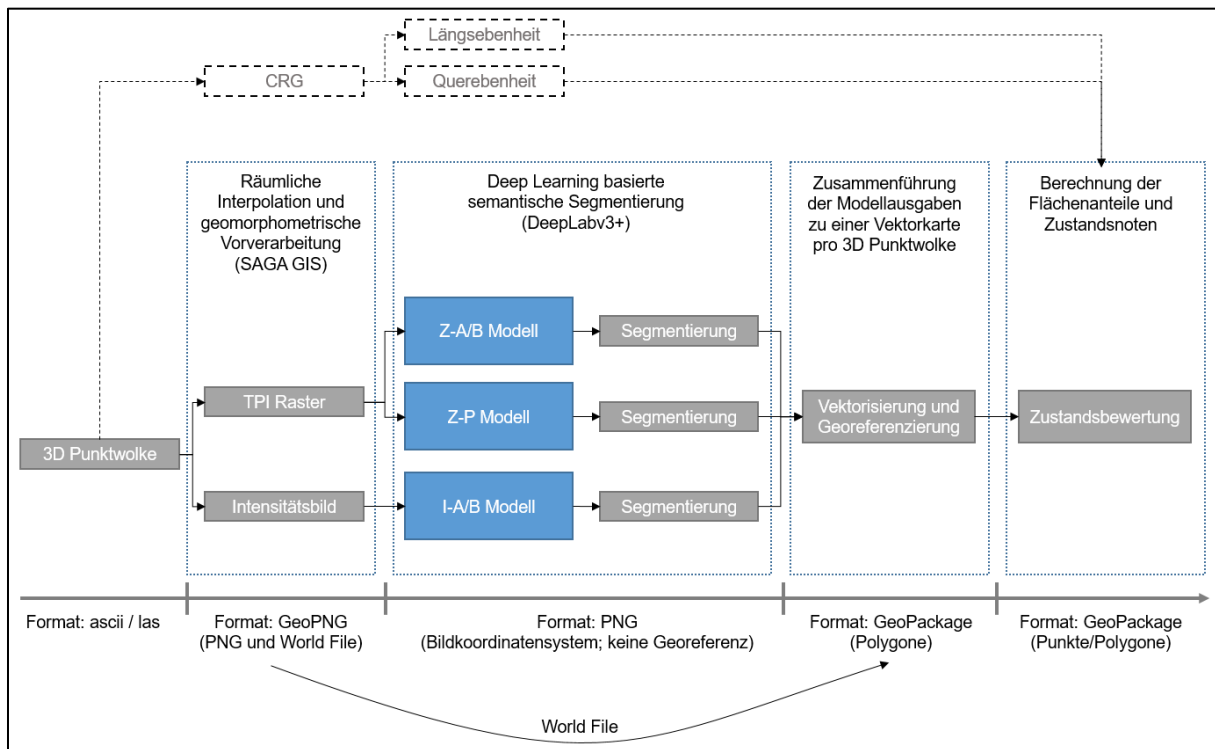


Abbildung 178: Ablaufschema der automatischen Zustandsbewertung auf Basis von Mobile LiDAR-Daten. Die gestrichelten Linien sollen zeigen, dass die Ebenheitsbewertung auch auf Basis der Laserscans realisiert werden kann (siehe Kapitel 2.5.1), aber nicht im Fokus dieser Arbeit steht.

Zunächst werden nach erfolgter Messung die Trajektorien des MLS-Systems optimiert und damit 3D-Punktwolken in einem 10 m Intervall erzeugt. Jeder 10 m lange und ca. 4,2 m breite Laserscan-Abschnitt wird mittels linearer NN-Interpolation zu zwei Oberflächenmodellen interpoliert: ein Oberflächenmodell mit absoluten Höhen und ein Oberflächenmodell mit den Werten der Reflexionsintensität. Das Intensitätsraster wird ohne weitere Verarbeitungsschritte direkt in ein 8 bit GeoPNG-Format konvertiert, wobei die Bildinformation in einem PNG mit Alpha-Kanal und die Georeferenz in einem sogenannten World File gespeichert wird. Bei letzterem handelt es sich um eine Textdatei, die u. a. die Pixelgröße und die Geokoordinaten der linken oberen Bildecke speichert (sechs Parameter einer Affintransformation von Pixel- zu Geokoordinaten). Auf Grundlage des Höhenrasters wird der TPI berechnet und die Ergebnisse auf 8 bit gestreckt, sodass auch hier ein GeoPNG als Ausgabeformat erzeugt wird. Die beiden zu einem Laserscan-Abschnitt gehörenden PNGs werden anschließend mit den jeweiligen DeepLabv3+ Modellen segmentiert. Im Ergebnis liegen drei Segmentierungsmasken vor. Mithilfe der Affintransformation aus dem World-File werden aus den Pixelmasken der

einzelnen Klassen Geoobjekte in Form von Polygonen berechnet und als GeoPackage gespeichert. Damit liegt eine georeferenzierte Segmentierung der Oberflächenmodelle vor (thematische Kartierung). Die Information, zu welchem Laserscan-Abschnitt die Geoobjekte gehören, wird im gesamten Prozess als Attribut mitgeführt, sodass auf Basis der realen Flächen von Schäden und Fahrbahnoberfläche die Zustandsgrößen gemäß der Erfordernisse des kommunalen Straßennetzes berechnet werden können (FGSV 2012): pro 10 m Abschnitt wird der Flächenanteil berechnet, der von Rissen (RISS), aufgelegten Flickstellen (AFLI) und sonstigen Oberflächenschäden (OFS) betroffen ist. Die berechneten Zustandsgrößen werden, wie in Kapitel 2.1.3.3 beschrieben, normiert und zum Teilzielwert Substanz (TWSUB) verknüpft. Damit liegt eine regelwerkskonforme Zustandsbewertung als georeferenzierte Ergebnisdatei vor. Für eine Kartendarstellung kann diese Ergebnisdatei entweder als Punktfolge im 10 m Intervall (2D-Schwerpunkt des Laserscan-Abschnitts) oder in Form von 10 m Polygonen (2D-konvexe Hülle um den Laserscan-Abschnitt) ausgegeben werden (siehe Abbildung 179). In Anhang XI sind verschiedene Beispiele aufgeführt.

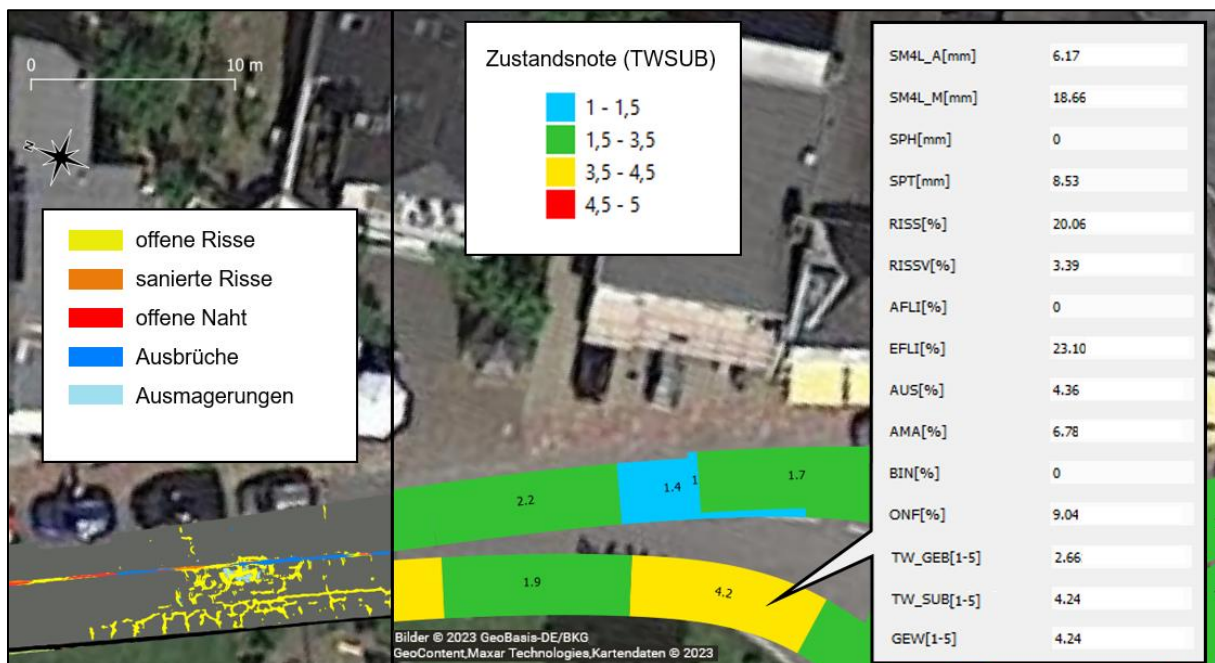


Abbildung 179: Beispiel für eine aus den PPS-Oberflächenmodellen automatisch erzeugte, objektscharfe Schadenskartierung (links) und die daraus abgeleitete georeferenzierte Ergebnisdatei in Form von 10 m Polygonen (rechts). Jedes Polygon der Ergebnisdatei ist mit den Flächenanteilen der Schäden im jeweiligen 10 m Abschnitt attribuiert. Die Zustandsgrößen der Ebenheit wurden in diesem Beispiel auch aus dem Laserscan berechnet (siehe Kapitel 2.5.1). Im rechten Teil der Abbildung sind die Ergebnisdateisegmente nach dem Substanzwert (TWSUB) farbkodiert dargestellt.

Mit dem beschriebenen Ansatz können für einen 1 km langen Messabschnitt innerhalb von wenigen Minuten Schäden, Objekte und Deckenarten automatisch kartiert werden. Ein geübter

Auswerter benötigt dagegen für die bislang übliche, weitaus abstrahierendere 1 m²-weise Erfassung der relevanten Schäden auf einer 1 km langen Strecke, je nachdem, wie komplex die jeweils vorhandene Schadenssituation ist, 15 bis 60 Minuten.

Während die verwendete GPU ca. 0,7 Sekunden pro DeepLabV3+ Modell zur semantischen Segmentierung eines PNGs benötigt, steigt die Rechenzeit auf einer CPU auf ca. 58 Sekunden pro DeepLabV3+ Modell und Bild. Eine Nutzung des Analysesystems auf einer CPU ist also prinzipiell möglich, für eine effiziente Analyse ist eine GPU aber unerlässlich.

5.5.4 Vergleich zu Referenzsystemen der Zustandserfassung

Mithilfe des in Kapitel 3.2.6 beschriebenen Datensatzes und des in Kapitel 4.5.4 erläuterten Analysekonzepts war es möglich, zwei unterschiedliche Aspekte zu untersuchen: Zum einen konnte überprüft werden, ob die Fehler der automatischen Erfassung von Oberflächenschäden innerhalb des Toleranzbereichs liegen, den die BAST für eine manuelle, von geschultem Personal durchgeführte Inspektion definiert (siehe Kapitel 2.1.4). Zum anderen konnte das im Rahmen der vorliegenden Arbeit entwickelte, automatische Analysesystem gegen nationale und internationale Referenzsysteme zur Straßenzustandserfassung getestet werden.

In einem ersten Schritt wurde die Trajektorie des I.R.I.S 16 Systems so verändert, dass der Heading-Winkel stetig nach Norden verläuft. Die 3D-Punktwolken, die mit diesen modifizierten Trajektorien berechnet wurden, verloren zwar ihre korrekte Georeferenz, aber die Oberflächenmodelle (GeoPNGs), die basierend auf den 10 m langen Scanabschnitten interpoliert wurden, lagen in der Folge räumlich synchronisiert zu den TP3-Oberflächenbilddaten des I.R.I.S 3 Systems und zu den LCMS2-Daten des I.R.I.S 17 Systems vor. Nachfolgend werden die Begriffe PPS-System (I.R.I.S 16), TP3-System (I.R.I.S 3) und LCMS2-System (I.R.I.S 17) verwendet. Nachdem die Daten der drei Systeme in einem einheitlichen Format vorlagen, wurden die Oberflächenmodelle des PPS-Systems gemäß ZTV ZEB-Stb (FGSV 2006) mit der in Kapitel 4.1.5 beschriebenen Software regelwerkskonform durch geschultes Auswertepersonal der Firma Lehmann+Partner GmbH inspiziert. Da es sich bei der Evaluierungsstrecke um eine Asphaltdecke handelt, musste im Zuge der manuellen Auswertung zunächst das 3 x 10 Zellen umfassende Auswerteraster für jedes Bild seitlich in die Fahrstreifengrenzen eingepasst werden (siehe Kapitel 2.1.2.3). Bei der Evaluierungsstrecke betraf dies vor allem den rechten Fahrbahnrand, der hier i. d. R. durch eine gepflasterte Entwässerungsrinne definiert ist. Typischerweise wird der linke Fahrstreifenrand durch eine Fahrbahnmarkierung definiert. Da diese auf der Evaluierungsstrecke nicht vorhanden ist, wurde das Auswerteraster am linken

Bildrand der PPS-Oberflächenmodelle festgesetzt. Wie in Kapitel 4.5.4 beschrieben und in Abbildung 154 dargestellt, wurde dieses auf Basis des PPS-Systems definierte Auswerteraster genutzt, um die Ergebnisse verschiedener Auswerteansätze bzw. verschiedener Sensoren in einem gemeinsamen Bezugssystem vergleichen zu können.

Vergleich von manueller und automatischer Erfassung von Oberflächenschäden

Im Zuge der manuellen Einpassung des Auswerterasters wurden vom Inspektionspersonal jene Rasterzellen mit den Schadensarten aktiviert, die innerhalb der betreffenden Zellen vorkommen (siehe Abbildung 180). Die manuelle Auswertung der ca. 2,4 km langen Strecke, die sowohl das Einpassen des Rasters in die Fahrstreifengrenzen als auch das Aufnehmen der Oberflächenschäden umfasst, dauerte ca. 120 Minuten, da pro 10 m Abschnitt beide Informationskanäle (Höhe und Intensität) inspiziert werden mussten.

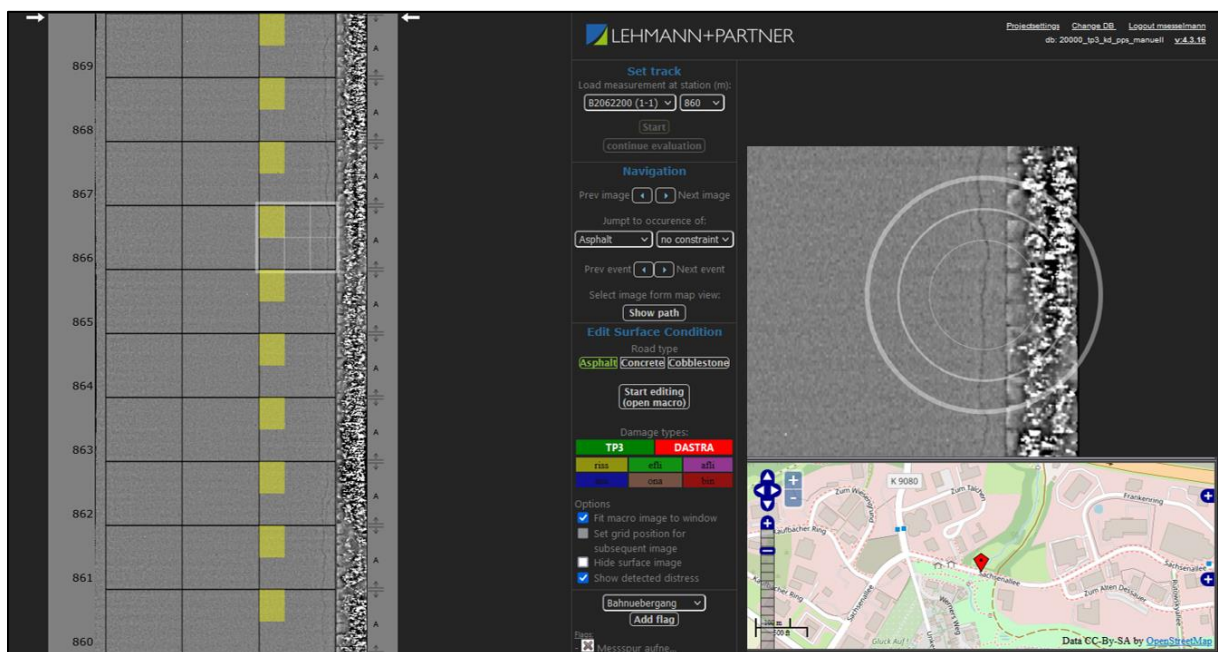


Abbildung 180: Manuelle Auswertung der Oberflächenschäden auf Basis von begradierten PPS-Oberflächenmodellen. Im Beispiel sind jene Rasterzellen gelb markiert, in denen Risse von der auswertenden Person erkannt wurden.

In einem zweiten Schritt wurden die begradierten PPS-Oberflächenmodelle mit den in Kapitel 5.5.2.2 beschriebenen DeepLabv3+ Modellen analysiert. Die Segmentierungsergebnisse der Modelle wurden zusammengeführt und vektorisiert (siehe Abbildung 178). Vor dem Georeferenzierungsschritt wurden die Schadstellenpolygone mit dem Auswerteraster, das durch die zuvor beschriebene manuelle Auswertung in seiner relativen Lage auf jedem einzelnen Oberflächenmodell bereits definiert war, im Bildkoordinatensystem räumlich verschnitten. Hiermit wurden die konturscharf segmentierten Schäden in das gerasterte TP3-

Auswerteformat überführt. In Abbildung 181 ist beispielhaft ein 10 m langes PPS-Oberflächenmodell in der TP3-Auswertesoftware dargestellt. Die gelben, unregelmäßigen Polygone repräsentieren die vom Z-A/B-Modell segmentierten offenen Risse. Die schwarzen Rasterzellen, die mit einem gelben Rechteck in der linken oberen Ecke markiert sind, entsprechen den Zellen, die aufgrund der Verschneidung mit den Schadstellenpolygonen automatisch mit der Schadensart RISS aktiviert wurden. Im Beispiel sind 10 von 30 Zellen aktiviert, was für den durch das Oberflächenmodell repräsentierten 10 m Abschnitt einem Flächenanteilwert von 33 % für die Zustandsgröße RISS entspricht. Nach Anwendung der Normierungsfunktion für Ortsdurchfahrten würde hier der Zustandswert ZW-RISS = 5,0 resultieren.

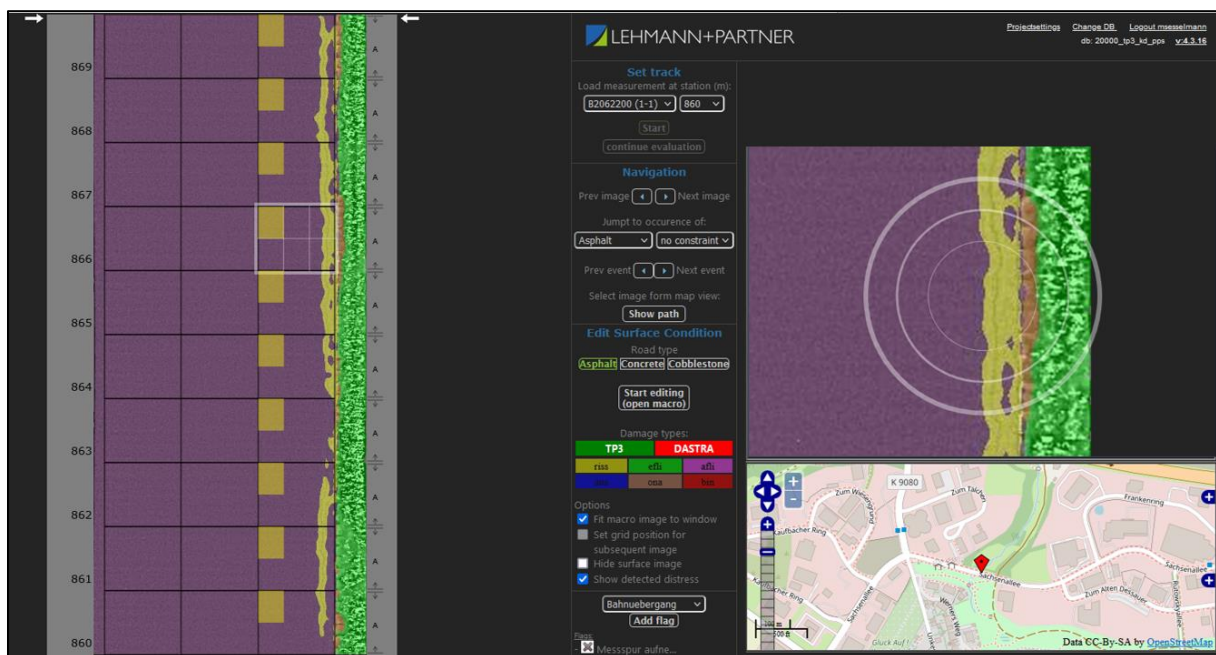


Abbildung 181: Automatische Auswertung der Oberflächenschäden auf Basis von begradtigten PPS-Oberflächenmodellen. Im Beispiel sind jene Rasterzellen gelb markiert, in denen Risse vom Z-A/B-Modell erkannt wurden.

Die automatische Auswertung dauerte mit der in Kapitel 4.1 beschriebenen GPU ca. 8 Minuten. Die Anwendung der DeepLabv3+ Modelle auf einer CPU (Intel i7) hingegen würde ca. 700 Minuten in Anspruch nehmen.

Nun lag eine manuelle Erstauswertung und eine automatische Zweitauswertung auf demselben Datenmaterial und unter Verwendung desselben Auswerterasters vor. So konnte das Instrument der Eigenüberwachung angewendet werden, um zu überprüfen, ob die vorgeschriebenen Toleranzen, die die BASt für die Erfassungsqualität durch geschultes Auswertepersonal festgelegt hat, auch durch den Deep Learning basierten Ansatz der automatischen Erfassung von Oberflächenschäden eingehalten werden können. Zunächst wurden gemäß der ZTV ZEB-Stb (FGSV 2006) aus den Aktivierungen des Analyserasters die beiden für die Eigen-

überwachung relevanten Zustandswerte ZW-RISS und ZW-FLI für 100 m lange Straßenabschnitte berechnet. Hierbei wurden die Normierungsfunktionen für Ortsdurchfahrten verwendet, die mit den Normierungsfunktionen des Arbeitspapiers AP 9 K 3.1 (FGSV 2016) für kommunale Straßen der Funktionsklasse A identisch sind (siehe Abbildung 15). Aus der mit dem PPS-System aufgenommenen Strecke wurden 20 aufeinanderfolgende 100 m Abschnitte gewählt, die hinsichtlich ihrer Zustandswerte im Bereich zwischen 2,0 und 4,0 lagen. Anschließend wurde pro Abschnitt die Differenz zwischen Erst- und Zweitauswertung gebildet und darüber der Mittelwert und die Standardabweichung berechnet. Die zulässigen Mittelwerte und Standardabweichungen sind mit 0,5 Notenwerten im Regelwerk festgelegt (siehe Kapitel 2.1.4).

Tabelle 21: Vergleich einer manuellen und automatischen Auswertung von PPS-Daten nach den Anforderungen der Eigenüberwachung gemäß ZTV ZEB-Stb (FGSV 2006). Für eine 2 km lange Strecke wurden die Zustandswerte (Noten) der beiden verglichenen Auswertungen für 100 m lange Abschnitte berechnet und die Differenzen gebildet. Über die 20 berechneten Differenzen wurden Mittelwert $\bar{x}$ und Standardabweichung σ bestimmt.

100 m Abschnitt	manuelle Auswertung		automatische Auswertung		Statistik für 2 km (Differenzen)			
					ZW-RISS		ZW-FLI	
	ZW-RISS	ZW-FLI	ZW-RISS	ZW-FLI	$\bar{x}$	σ	$\bar{x}$	σ
1	3,0	2,4	2,6	2,4	0,2	0,3	0,0	0,1
2	4,0	1,3	3,4	1,3				
3	3,7	1,4	3,6	1,1				
4	2,5	1,5	2,3	1,5				
5	3,9	2,0	3,3	2,0				
6	3,6	1,9	3,5	1,6				
7	3,8	1,4	3,7	1,4				
8	3,1	1,3	3,0	1,2				
9	4,9	2,1	4,4	2,2				
10	5,0	3,0	4,9	2,8				
11	2,2	5,0	2,1	4,9				
12	2,4	3,8	2,2	3,9				
13	1,7	5,0	1,5	5,0				
14	1,3	5,0	1,2	5,0				
15	1,9	5,0	1,8	5,0				
16	1,6	5,0	1,5	5,0				
17	1,8	5,0	1,9	5,0				
18	2,9	4,8	3,1	4,6				
19	3,6	2,2	2,9	2,3				
20	3,9	3,6	3,3	3,7				

Wie in Tabelle 21 dargestellt, werden die Toleranzen durch die automatische Auswertung eingehalten. Demzufolge liefert das Deep Learning basierte Analysesystem auf der Evaluierungsstrecke eine Erfassungsqualität, die den Anforderungen der Straßenzustands-

erfassung in Deutschland genügt. Das zugehörige Protokoll der Eigenüberwachung ist im Anhang XV zu finden.

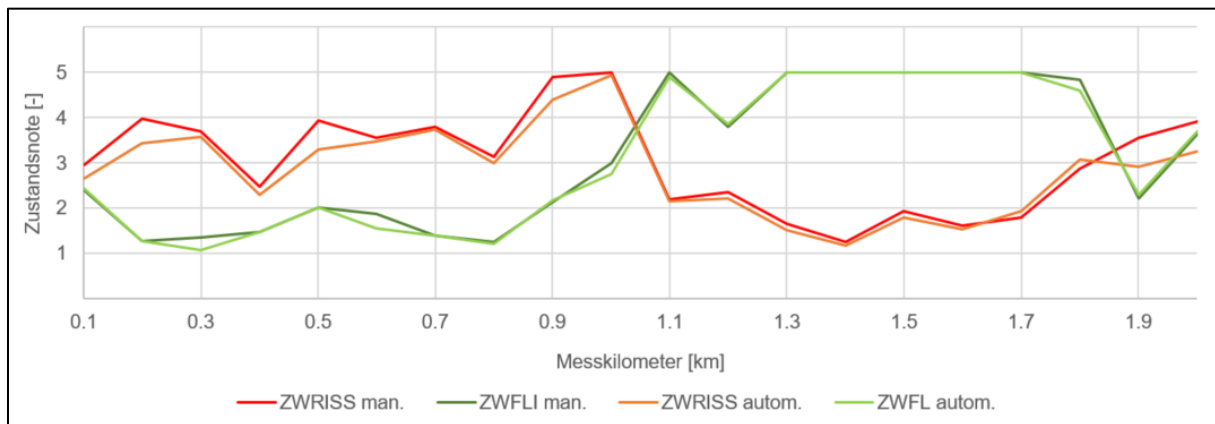


Abbildung 182: Ergebnisse des Vergleichs von automatischer und manueller Erfassung der Oberflächenschäden nach dem Verfahren der Eigenüberwachung, bei der die Zustandswerte für Risse und Flickstellen in 100 m langen Auswerteabschnitten einer Erst- und Zweitmessung miteinander verglichen werden. Die Erstmessung wurde hierbei durch geschultes Personal und die Zweitmessung durch das Deep-Learning-Verfahren erzeugt. Beide Auswertungen erfolgten auf denselben PPS-Oberflächenmodellen.

Die Kurvenverläufe von Abbildung 182 deuten an, dass das Deep Learning basierte Analysesystem tendenziell eher Fehler des Typs 2 (False Negatives) produziert. Die automatische Zustandsbewertung unterschätzt das Vorhandensein von Schäden im Vergleich zur manuellen Auswertung also leicht. Im ersten Teil der Evaluierungsstrecke treten Risse nahezu kontinuierlich entlang der Entwässerungsrinne auf. Solche Risse werden automatisch teilweise nur unvollständig erkannt, während sie bei der manuellen Auswertung im Oberflächenmodell detektiert werden. Das deutet darauf hin, dass derartige Beispiele im Trainingsdatensatz noch nicht ausreichend vertreten sind. Die geringen Abweichungen zwischen manueller und automatischer Auswertung für die Zustandswerte der Flickstellen lässt sich damit erklären, dass die Flickstellen der Evaluierungsstrecke für einen Auswerter zweifelsfrei erkennbar sind: sie zeichnen sich durch deutliche Materialunterschiede (Reflexionsintensität) aus und sind kaum erodiert. Außerdem kommen wenige Sonderfallsituationen vor (z. B. Deckensanierungen oder in sich verschachtelte Flickstellen unterschiedlicher Sanierungsmaßnahmen). Hinzu kommt, dass die Flickstellen auf ca. 400 m der Evaluierungsstrecke so großflächig vorhanden sind, dass sie durch die Normierung mit einem Zustandswert von ZW-FLI = 5,0 benotet wurden. Auch wenn die eigentlich erkannten Zustandsgrößen zwischen manueller und automatischer Analyse in diesem Bereich voneinander abweichen, führen sie durch die Bewertungssystematik ab einer Prozentzahl von > 25 % zur selben Bewertung.

Vergleich zu nationalen und internationalen Referenzsystemen

In Deutschland stellen hochauflösende Oberflächenbilder, die mit einem System aufgenommen wurden, das durch die BASt geprüft und zugelassen wurde, den höchsten Qualitätsstandard im Bereich der Erfassung von Straßenschäden dar. Im kommunalen Straßennetz werden solche Daten i. d. R. nicht erhoben, da vor allem die detaillierte Auswertung der Oberflächenbilder gemäß ZTV ZEB-Stb (FGSV 2006) für die jeweiligen Baulastträger nicht wirtschaftlich ist. In den nachfolgend beschriebenen Analysen wird die Erfassung von Oberflächenschäden aus Oberflächenbildern, die mit dem von der BASt zugelassenen Messsystem I.R.I.S 3 aufgenommen wurden, als Referenz genutzt, um die Leistungsfähigkeit des im Rahmen der vorliegenden Arbeit entwickelten Analysesystems mit dem nationalen Stand der Technik zu vergleichen.

In internationaler Perspektive ist das LCMS2 das am weitesten verbreitete System zur automatischen Straßenzustandserfassung. Anders als in Deutschland ist in europäischen Nachbarländern, wie Österreich, Frankreich, Tschechien oder Polen, die Zustandserfassung mit LCMS2-Sensoren und der integrierten automatischen Schadenserkenkung möglich und wird von zahlreichen Dienstleistern angeboten. Weltweit nutzen Dienstleister in 45 Ländern LCMS2-basierte Lösungen zur systematischen Überwachung des Straßenzustands. Bislang existieren keine unabhängigen Studien zur Leistungsfähigkeit des LCMS2-Systems im Hinblick auf die Anforderungen der ZTV ZEB-Stb (FGSV 2006) oder der E EMI (FGSV 2012).

Beide vorstehend aufgeführten Referenzsysteme nehmen die Straßenoberfläche in einer Auflösung von ca. 1 mm x 1 mm auf. Damit sind die Referenzdaten deutlich hochauflösender als die Oberflächenmodelle auf Basis von PPS-Laserscans (5 mm x 5 mm).



Abbildung 183: Skizze über die für den Vergleich verfügbaren Evaluierungsstreckenabschnitte der drei verwendeten Messsysteme.

Wie in Kapitel 3.2.6 beschrieben und in Abbildung 183 skizziert, zeicheten I.R.I.S 3 und I.R.I.S 16 nacheinander fahrend denselben Streckenabschnitt auf. Das mit dem LCMS2 ausgerüstete Messfahrzeug I.R.I.S 17 konnte aus logistischen Gründen erst ca. 1 Monat später

und nur auf einem ca. 1,2 km langen Teilstück der Evaluierungsstrecke Daten aufnehmen. In diesem Zeitraum sind Zustandsveränderungen nicht zu erwarten, sodass ein Vergleich der Erfassungsergebnisse möglich ist, obwohl keine synchrone Datenaufnahme aller verglichenen Systeme gegeben ist.

Die mit dem Messfahrzeug I.R.I.S 3 aufgenommenen TP3-Oberflächenbilder wurden gemäß ZTV ZEB-Stb (FGSV 2006) mit der in Kapitel 4.1.5 beschriebenen Software regelwerkskonform durch geschultes Auswertepersonal der Firma Lehmann+Partner GmbH ausgewertet. Diese Daten werden nachfolgend mit dem Begriff TP3-System bezeichnet. Die Lage des Auswerterasters wurde dabei von der manuellen Auswertung der PPS-Oberflächenmodelle (I.R.I.S 16) übernommen.

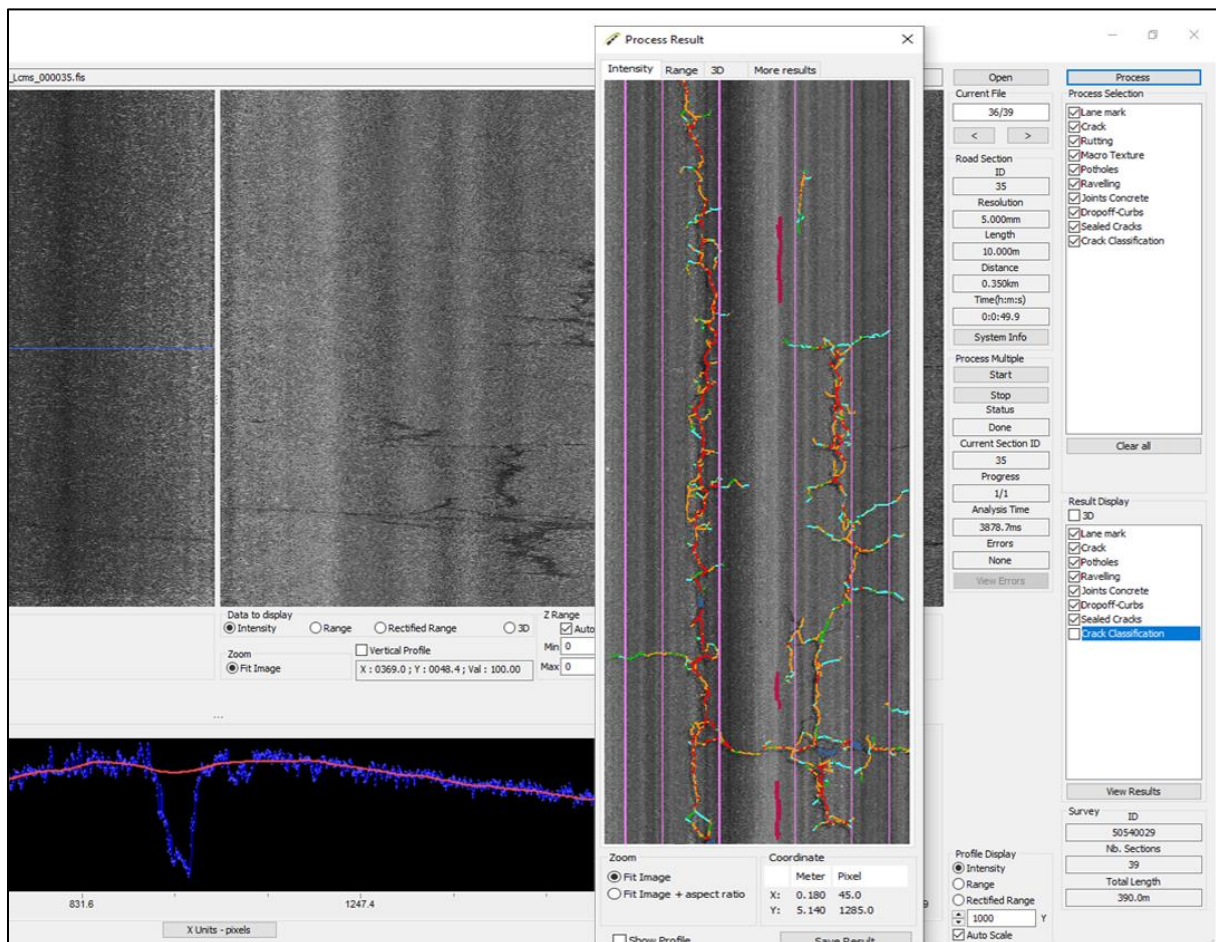


Abbildung 184: Ausschnitt der Benutzeroberfläche der Software LcmsRoadInspect, mit der Bilddaten aus dem LCMS-Rohdatenformat exportiert sowie die automatische Schadenserkennung parametrisiert und gestartet werden kann.

Die mit dem Messfahrzeug I.R.I.S 17 aufgenommenen LCMS2-Daten liegen im proprietären fis-Format vor und wurden unter Verwendung der an die beiden Sensoren gekoppelten Auswertesoftware LcmsRoadInspect prozessiert (siehe Abbildung 184). Dabei wurden sowohl

Oberflächenbilder im 10 m Intervall erzeugt, als auch die automatische Schadenserkenkung mit Default-Parametern gestartet. Die verwendeten Parameter sind im Anhang XV aufgeführt. Pro Oberflächenbild werden die detektierten Schäden in XML-Dateien gespeichert. Darin ist die Geometrie der erkannten Schäden (Linien, Polygone) im Bildkoordinatensystem abgelegt. Durch Verschneiden der Schadensgeometrien mit dem Auswerteraster der TP3-Erfassungssoftware konnten die LCMS2-Detektionen auf das in Deutschland vorgeschriebene Standardformat übertragen werden. Auch hier wurde wieder die Position des Auswerterasters, die aus der manuellen Auswertung des PPS-Systems resultierte, zugrundegelegt.

In der messtechnischen kommunalen Zustandsbewertung werden Flickstellen, wie in Kapitel 2.1.3.3 beschrieben, nur in ihrer aufgelegten Form berücksichtigt. Da LCMS2 nicht zwischen eingelegten und aufgelegten Flickstellen unterscheidet, wurde für den nachfolgenden Vergleich das allgemeinere Merkmal Flickstelle verglichen. Für das TP3-System und das PPS-System ergibt sich der jeweilige Flickstellenwert, wenn entweder eine eingelegte oder eine aufgelegte Flickstelle erkannt wurde. Außerdem wurden offene und sanierte Risse separat betrachtet, um einen möglichst differenzierten Vergleich der Systeme zu ermöglichen. Die Oberflächenschäden Ausmagerung und Bindemittelanreicherung kamen auf der Evaluierungsstrecke nicht vor und wurden auch von keinem der verglichenen Systeme detektiert.

Abbildung 185 zeigt den fast 1,2 km langen Abschnitt der Evaluierungsstrecke, der mit dem im Rahmen der vorliegenden Arbeit entwickelten Analysesystem sowie zwei Referenzsystemen im Hinblick auf Oberflächenschäden ausgewertet wurde. Es lässt sich ein hohes Maß an Übereinstimmung zwischen der Schadenserfassung durch das TP3-System und durch das PPS-System feststellen. Der auf Basis von 10 m langen Auswerteabschnitten berechnete Korrelationskoeffizient des Flächenanteils beträgt für offene Risse 0,86, für Flickstellen 0,88, für Ausbrüche 0,79 und für sanierte Risse 0,59. Zwischen den Zustandsgrößen, die durch das TP3-System und durch das PPS-System erfasst wurden, besteht ein mittelstarker bis stark positiver Zusammenhang. Die vergleichsweise niedrige Korrelation für die sanierten Risse lässt sich damit erklären, dass auf der Evaluierungsstrecke manche Risse mit sehr schmalen, länglich aufgelegten Flickstellen saniert wurden, die vom I-A/B-Modell eher als sanierter Riss erkannt werden.

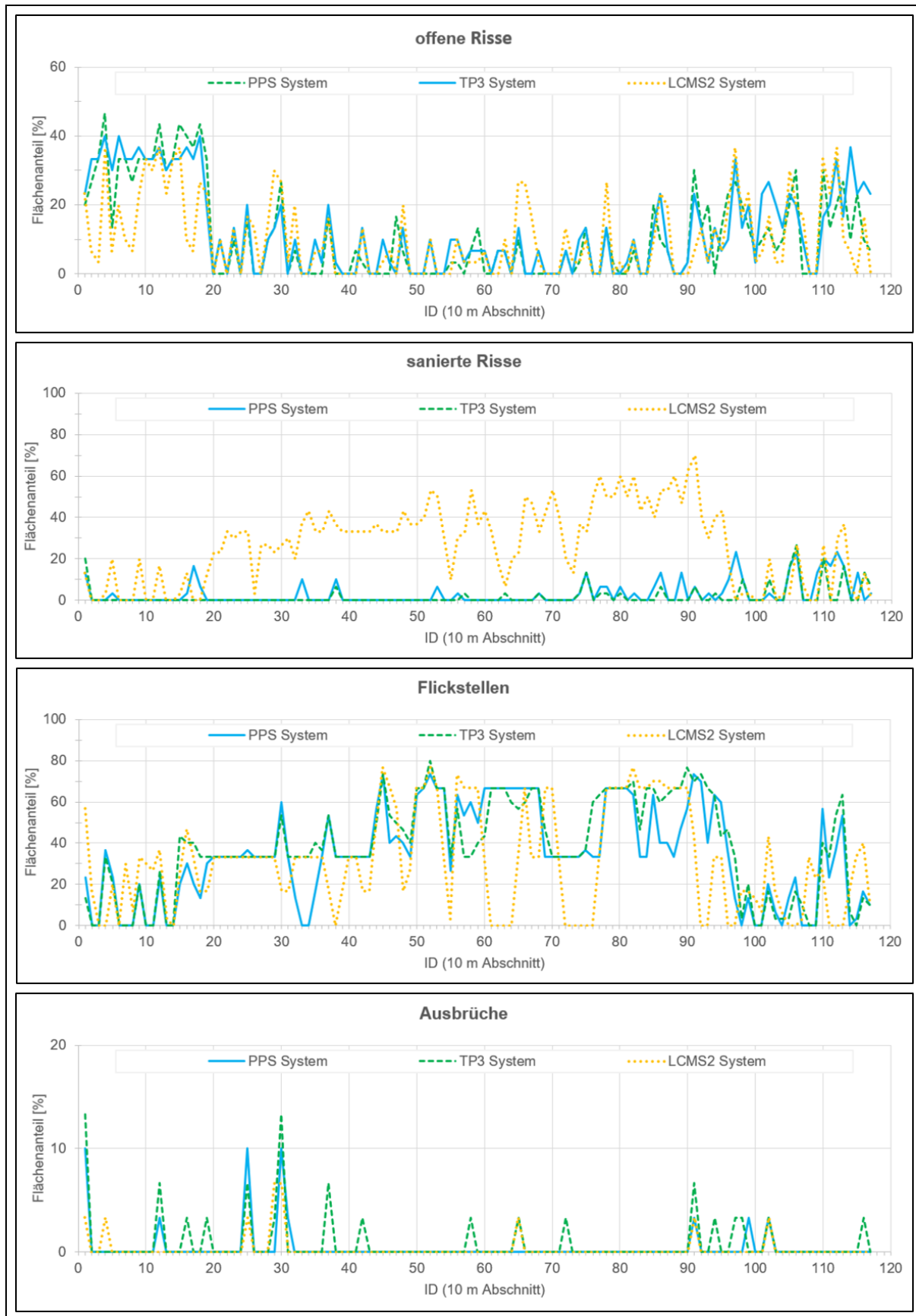


Abbildung 185: Vergleich der Deep Learning basierten Auswertung auf Grundlage von PPS-Oberflächenmodellen (PPS-System) mit einem nationalen (TP3-System) und internationalen Referenzsystem (LCMS2-System). Die Graphen zeigen die vier auf der Evaluierungsstrecke vorhandenen Schadensarten.

Ein Beispiel dafür ist an Abschnitt-ID 17 gegeben: Während das TP3-System ca. 40 % der Abschnittsfläche von Flickstellen betroffen ausweist und keine sanierten Risse erkennt, liefert die Auswertung des PPS-Systems ca. 17 % Flächanteil für sanierte Risse und 24 % Flächenanteil für Flickstellen, da hier fälschlicherweise die eigentlichen Flickstellen als sanierte Risse erkannt wurden (siehe Abbildung 186).

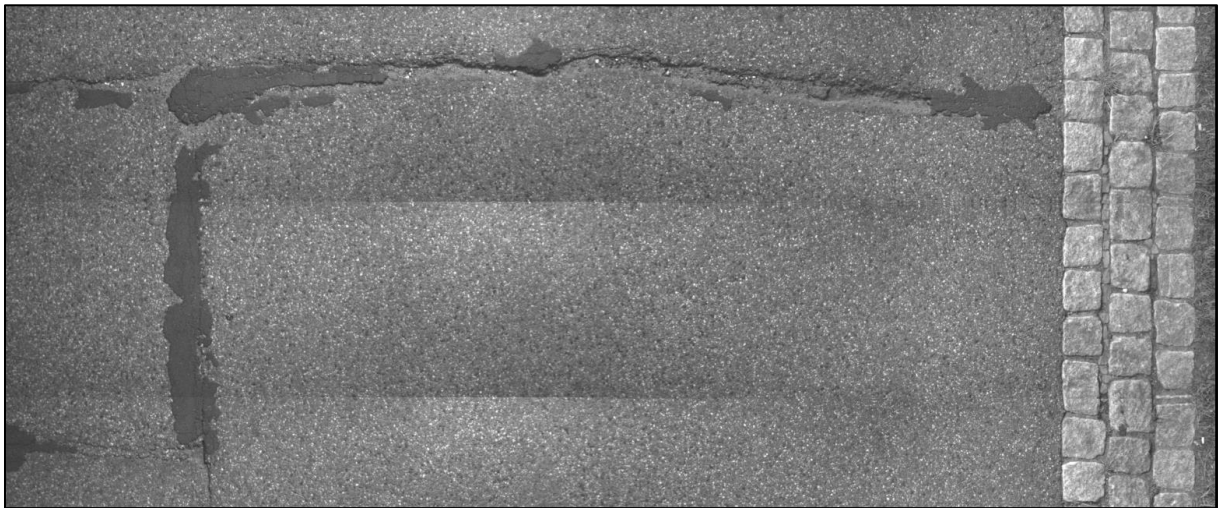


Abbildung 186: Aufgelegte Flickstellen, die aufgrund ihrer Form vom PPS-System als sanierte Risse fehlklassifiziert werden. Dargestellt ist ein Ausschnitt aus einem Oberflächenbild des TP3-Systems.

Die Ergebnisse des PPS-Systems und des LCMS2-Systems weisen ein geringeres Maß an Übereinstimmung auf. Der auf Basis der 10 m langen Auswerteabschnitten berechnete Korrelationskoeffizient des Flächenanteils beträgt für offene Risse 0,66, für Flickstellen 0,43, für Ausbrüche 0,64 und für sanierte Risse -0,08. Zwischen den Zustandsgrößen, die durch das LCMS2-System und durch das PPS-System erfasst wurden, besteht in den meisten Schadensarten ein mittelstarker positiver Zusammenhang. Für sanierte Risse liegt dagegen ein sehr schwach negativer Zusammenhang vor. LCMS2 neigt auf der untersuchten Strecke stark dazu, die Fugenbänder von Flickstellen als sanierten Riss zu erkennen. In den Graphen von Abbildung 185 ist dies vor allem im Bereich der Abschnitt-IDs 15 bis 95 zu sehen: Dort liegen die Flächenanteile für Flickstellen im TP3-System und PPS-System in einem Bereich zwischen 40 % bis 60 %. Für dieselben Abschnitte steigt in den Ergebnissen des LCMS2-Systems der Flächenanteil sanierten Risse auf ein Niveau von ca. 40 % an, welches weder im TP3-System noch im PPS-System für sanierte Risse so detektiert wird. Abbildung 187 zeigt ein Beispiel für eine Situation, in der eine idealtypische eingelegte Flickstelle vom LCMS2-System nicht detektiert, dafür jedoch das die Flickstelle umgebende Fugenband als sanierte Riss fehlklassifiziert wird. Damit ist auch die vergleichsweise niedrige Übereinstimmung von

LCMS2-System und TP3-System bezüglich der einzelnen Schadklassen zu erklären: Für offene Risse beträgt der Korrelationskoeffizient 0,66, für Flickstellen 0,45, für Ausbrüche 0,67 und für sanierte Risse -0,02.

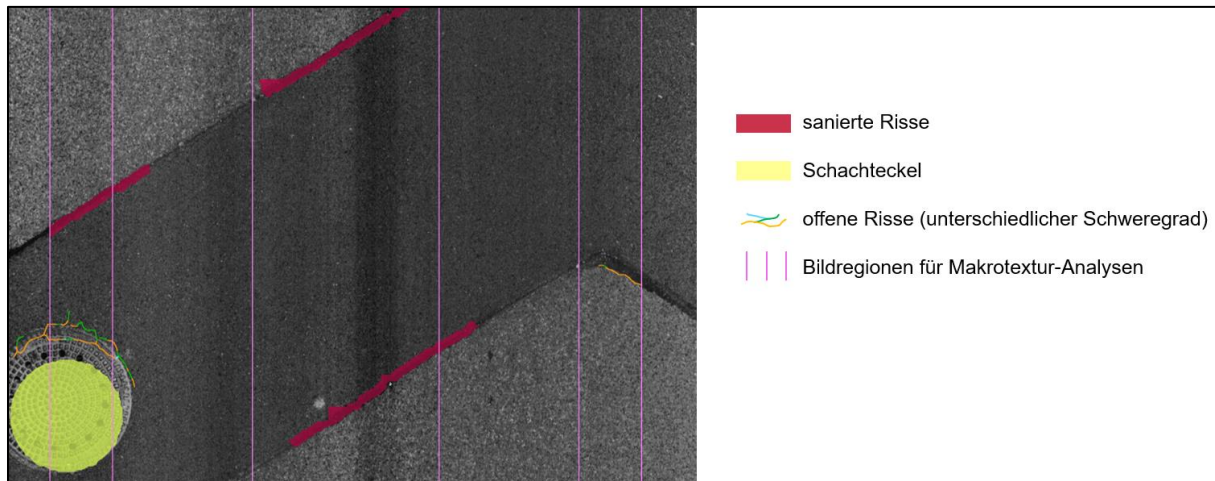


Abbildung 187: LCMS2-Intensitätsbild mit überlagerten Schadstellen und Objekten, die automatisch erkannt wurden. Die eingelegte Flickstelle wurde nicht als solche erkannt. Das Fugenband wurde teilweise als sanierter Riss fehlklassifiziert. Auch kommt es im Bereich des Schachtdeckels zu falsch-positiven Rissdetektionen.

Wie in Kapitel 3.2.6 beschrieben, erfassen alle drei verwendeten Systeme auf leicht unterschiedlichen Breiten von 4 m bis 4,5 m. Auch kann während einer Messung in der Praxis nicht exakt dieselbe Fahrlinie eingehalten werden. Die Trajektorien weichen im Mittel ca. 12 cm voneinander ab. Unterschiede in der Zustandserfassung, die in Abbildung 185 festzustellen sind, können also zum Teil auch daher rühren, dass die Daten in lateraler Richtung nicht vollständig deckungsgleich sind und das Auswerteraster des Systems mit der mittleren Erfassungsbreite zugrundegelegt wurde.

Die berechneten Korrelationskoeffizienten verweisen darauf, dass mit dem PPS-System eine höhere Übereinstimmung zu dem nationalen Referenzsystem (TP3-System) erzielt werden konnte, als mit dem LCMS2-System (siehe Tabelle 22). Mit den verwendeten Default-Parametern konnte auf der Evaluierungsstrecke mit dem LCMS2-System keine adäquate Auswertung gemäß den Erfordernissen der deutschen Straßenzustandserfassung erreicht werden. Von einem Finden optimaler Parameter wurde abgesehen, da dies über die Fragestellungen der Arbeit hinaus geht und das LCMS2-System aufgrund der verwendeten Laserklasse in Deutschland im öffentlichen Raum nicht ohne erheblichen Absicherungsaufwand eingesetzt werden kann.

Tabelle 22: Korrelationskoeffizienten der verglichenen Systeme für einzelne Schadensarten.

Systemvergleich	Offene Risse	Flickstellen	Ausbrüche	Sanierte Risse
TP3 vs. PPS	0,86	0,88	0,79	0,59
LCMS2 vs. PPS	0,66	0,43	0,64	-0,08
TP3 vs. LCMS2	0,66	0,45	0,67	-0,02

Am schnellsten lagen die Zustandserfassungsergebnisse beim LCMS2-System vor. Die Bilderzeugung und -auswertung dauerte für die ca. 1,2 km lange Strecke etwa 10 Minuten (CPU). Beim TP3-System ist die Erzeugung der Oberflächenbilder in ca. 2 Minuten möglich (CPU), jedoch benötigte die manuelle Auswertung ca. 60 Minuten, sodass insgesamt die Bereitstellung der erfassten Oberflächenschäden hier am längsten dauerte. Betrachtet man die Rechenzeit, die rein für die Schadstellensegmentierung benötigt wird, ist das im Rahmen der vorliegenden Arbeit entwickelte Analysesystem bestehend aus drei DeepLabv3+ Modellen in der Lage, eine semantische Segmentierung in ca. 5 Minuten zu erzeugen (GPU). Jedoch sind die 3D-Punktwolkenerzeugung und Oberflächenmodellierung, die im Vorfeld nötig sind, vergleichsweise rechenaufwändige Prozesse, sodass das PPS-System hinsichtlich der Gesamtprozesszeit eine Mittelstellung zwischen den beiden Referenzsystemen einnimmt. Die Ergebnisse lagen insgesamt nach etwa 30 Minuten vor, wobei anzumerken ist, dass die Erzeugung von Punktwolken und Oberflächenmodellen CPU-basiert und ohne Optimierungen (z. B. Multithreading) implementiert ist.

Der in diesem Kapitel durchgeführte Vergleich erhebt nicht den Anspruch, die Leistungsfähigkeit der Referenzsysteme umfassend zu bewerten. Dafür ist der Umfang der Strecke zu klein und zu wenig divers. Dennoch kann festgehalten werden, dass die im Rahmen der vorliegenden Arbeit entwickelte Deep Learning basierte Auswertung von Straßenschäden auf Grundlage von PPS-Oberflächenmodellen auf einer dem System unbekannten Evaluierungsstrecke ähnliche Ergebnisse liefert, wie die beiden verglichenen Referenzsysteme, obwohl die Auflösung der verwendeten Oberflächenmodelle fünf Mal geringer ist, als die Oberflächenbilder der Referenzsysteme.

In Anhang XV sind repräsentative Ausschnitte aus den Messergebnissen der drei verglichenen Systeme gegenübergestellt.

5.5.5 Erkenntnisse zur Oberflächenanalyse mithilfe von Deep Learning

Im Rahmen der Voruntersuchungen konnte ein optimales Hyperparameter-Setting gefunden werden, mit dem erfolgreich leistungsfähige DeepLabv3+ Modelle trainiert werden konnten. Aufgrund der Größe und Auflösung der zu verarbeitenden Bilder kam selbst eine leistungsstarke GPU insofern an ihre Grenzen, als dass nur eine kleine Batch-Größe eingestellt werden und keine zeitgleiche Evaluierung der Modelle während des Trainings (Early Stopping) stattfinden konnte. Deshalb wurde aus den Erkenntnissen der Voruntersuchungen eine Trainingspipeline definiert: Da der Early-Stopping-Ansatz Hardware-bedingt nicht genutzt werden konnte, musste ein adäquater Zeitpunkt (Iteration) identifiziert werden, an dem das Training beendet werden kann. Hierzu wurden nach willkürlich definierten Iterationen Modelle exportiert und getestet. Zusätzlich wurde die Streuung der Trainingsprozesse quantifiziert, die durch die zufällige Initialisierung der nicht vortrainierten Parameter bedingt ist. Damit konnte besser eingeschätzt werden, ab welcher Score-Differenz sich zwei verglichene Modelle tatsächlich unterscheiden. Nach 150k bis 300k Iterationen konnte im Rahmen der Voruntersuchungen keine weitere Optimierung der Detektoren festgestellt werden. Es kann nicht ausgeschlossen werden, dass hier bereits ein Overfitting auf die Trainingsdaten einsetzt. Bei der Segmentierung der Höheninformation konnten die besten Ergebnisse mit Modellen erzielt werden, die auf TPI-Rastern trainiert wurden. Die Unterschiede zwischen den betrachteten morphometrischen Vorverarbeitungen der Höhenmodelle waren gering.

Mit den gefundenen Hyperparametern wurden insgesamt drei Modelle trainiert, die alle auf dem unabhängigen Testdatensatz evaluiert wurden und Scores erreichen, die im Bereich der menschlichen Leistungsfähigkeit für das Problem der pixelgenauen Segmentierung von Oberflächenschäden aus Bilddaten liegen.

Eingebettet in eine entsprechende Prozesskette, können die auf Bildbasis erzeugten Segmentierungsergebnisse in ein georeferenziertes Vektordatenformat überführt werden. Damit liegen die Ergebnisse als konturscharf kartierte Geoobjekte vor und können auf unterschiedliche Weise in GIS-Anwendungen weiterverarbeitet werden. So ist es beispielsweise möglich, eine Ergebnisdatei im Sinne der kommunalen Straßenzustandserfassung abzuleiten, wie sie typischerweise in weiterführenden Pavement-Management-Systemen im Rahmen der Straßenerhaltung genutzt wird.

An einer Evaluierungsstrecke wurde der entwickelte Ansatz gegen ein nationales und ein internationales Referenzsystem getestet. Dazu wurden die Daten aller eingesetzten Systeme synchronisiert und die PPS-Laserscans so konvertiert, dass ein Vergleich mit den Referenzsystemen möglich war. Der Vergleich macht deutlich, dass auf der

Evaluierungsstrecke mit dem eingesetzten MLS-System vergleichbare Ergebnisse zu den Referenzsystemen erzielt werden konnten, trotz nominell deutlich geringerer Auflösung der zugrundeliegenden PPS-Oberflächenmodelle im Vergleich zu den Daten des TP3-Systems und des LCMS2-Systems. Das PPS-System liefert dabei für die Erfordernisse der kommunalen Straßenerhaltung in Deutschland bessere Ergebnisse als das LCMS2-System, da es dem TP3-System ähnlichere Ergebnisse produziert und die Richtlinien für das kommunale Straßennetz grundsätzlich an der Erfassungssystematik der ZTV ZEB-Stb (FGSV 2006) orientiert sind.

Zusätzlich wurde überprüft, ob die Deep Learning basierte Auswertung auf Grundlage von PPS-Oberflächenmodellen die Qualitätskriterien der Eigenüberwachung gemäß ZTV ZEB-Stb (FGSV 2006) erfüllt. Für die Evaluierungsstrecke wurden die kritischen Toleranzen eingehalten. Damit kann von einer gleichwertigen Auswertequalität der trainierten DeepLabv3+ Modelle im Vergleich zu einem geschulten Auswerter ausgegangen werden.

5.6 Zusammenfassung der Ergebnisse

In Tabelle 23, Tabelle 24 und Tabelle 25 sind alle Untersuchungen und wesentlichen Ergebnisse in Bezug auf die Forschungsfragen der vorliegenden Arbeit zusammengefasst.

Tabelle 23: Zusammenfassung der Ergebnisse (Teil 1 von 2).

Forschungsfrage	Untersuchungen	Ergebnis
Können alle relevanten Schadenstypen aus den LiDAR-Daten des PPS detektiert werden?	- Anforderungsanalyse: Richtlinien der Zustandserfassung (FGSV 2012, 2016) - visuell-manueller Abgleich von Mess- und Zustandsdaten - Erstellung eines Objektkatalogs für das spätere Labelling von Trainingsdaten	Alle bewertungsrelevanten Schadensarten für Asphalt-, Pflaster- und Beton-Bauweisen können aus den verfügbaren Datenkanälen des PPS detektiert werden.
Wie hoch ist die absolute Positionsgenauigkeit der mit dem Mobilen Laserscanning System erfassten Laserscans der Fahrbahnoberfläche und implizit der daraus abgeleiteten Schadstellen?	- Passpunktaufnahme mit handgeführtem GNSS-RTK auf acht innerorts und außerorts gelegenen Kreisstraßen - Aufnahme mit MLS-System I.R.I.S 13 - Vergleich von Passpunktmessungen aus Mobile Laserscan und handgeführtem GNSS-RTK	Die Laserscans weisen eine absolute Lagegenauigkeit von im Mittel 4 cm ($\sigma = 3$ cm) und eine vertikale Genauigkeit von im Mittel 5 cm ($\sigma = 5$ cm) auf.

Tabelle 24: Zusammenfassung der Ergebnisse (Teil 2 von 3).

Forschungsfrage	Untersuchungen	Ergebnis
Mit welcher Präzision kann die Geometrie eines Messobjekts abgebildet werden?	3D-Modellierung eines Prüfkörpers als Referenz 8-fach Messung eines Prüfkörpers mit MLS-System I.R.I.S 3 - TIN-Vermaschung der Scans mittels Delaunay-Triangulation - ICP-Matching der TINs auf das Referenzmodell - Differenzbildung: gematchter Scan zu Referenzmodell	Die durch das Messsystem verursachte Streuung (RMSE) liegt im Bereich $\leq 0,84$ mm (nur horizontale und geneigte Oberflächen des Prüfkörpers). Werden auch Bereiche des Prüfkörpers betrachtet, die messprinzipbedingt nur mit verminderter Genauigkeit abgebildet werden können, beträgt die Streuung $\leq 1,31$ mm.
Welche räumliche Interpolationsmethode ist für die Erzeugung von digitalen Oberflächenmodellen aus 3D-Straßenoberflächen-scans im Rahmen der gegebenen Problemstellung und der verfügbaren Daten am besten geeignet?	- Vergleich deterministischer Verfahren anhand von zwei Testarealen - Kriterien: Erhaltung der Oberflächeneigenschaften, geringe Parametersensitivität, Rechenkomplexität, visuelles Erscheinungsbild - Identifikation der optimalen Methode in Anlehnung an Nutzwertanalyse	Die lineare Natural-Neighbor-Interpolation ist hinsichtlich der Kriterien Parameteranzahl, Laufzeit und vertikale Genauigkeit am besten geeignet, um Oberflächenmodelle aus PPS-Laser-scans zu generieren.
Bis zu welchem Ausmaß können Oberflächendiskontinuitäten bzw. Schadstellenstrukturen in laserscanbasierten Oberflächenmodellen räumlich aufgelöst und als solche erkannt werden?	- Anfertigung eines Prüfkörper mit definierten Spaltbreiten (1 - 30 mm) - Wiederholte Messungen des Prüfkörpers bei verschiedenen Geschwindigkeiten (30 und 50 km/h) und verschiedenen Ausrichtungen (längs und quer) - Interpolation der 3D-Punktwolken zu Oberflächenmodellen (Raster DEM) - Visuelle Inspektion der Raster und statistische Auswertung des Anteils erkennbarer Spalten	Aus den auf PPS-Laserscans basierenden Oberflächenmodellen können Spalten ab einer Öffnungsbreite von 1 mm detektiert werden. Eine reduzierte Geschwindigkeit begünstigt die Vollständigkeit der Erkennbarkeit von Spalten ≤ 2 mm Breite. Spalten mit einer Breite ≥ 5 mm sind sicher und überwiegend vollständig in den PPS-Oberflächenmodellen erkennbar.
Welches Hyperparameter-Setting ist für ein erfolgreiches Training von DeepLabv3+ Modellen auf Basis der vorliegenden Eingabedaten geeignet?	- Wiederholte Trainingsprozesse mit variierenden Hyperparameter-Einstellungen ausgehend von den Default-Werten des Quellcodes - Numerische Evaluation und visuelle Überprüfung der Segmentierungsergebnisse	Für die verwendete Hardware und die Datenbasis konnte eine Hyperparameter-Konfiguration identifiziert werden, mit der Modelle erfolgreich trainiert werden können. Nach 300k Iterationen verbesserte sich die Leistung der Modelle nicht mehr.

Tabelle 25: Zusammenfassung der Ergebnisse (Teil 3 von 3).

Forschungsfrage	Untersuchungen	Ergebnis
Welcher Streuung aufgrund von zufälligen Initialisierungsprozessen innerhalb der DeepLabv3+ Netzwerkstruktur unterliegt die Leistungsfähigkeit trainierter Modelle?	- Start von drei identisch parametrisierten Trainingsprozessen auf derselben Datengrundlage - Evaluation der entstandenen Modelle nach 150k und 300k Iterationen auf dem Testdatensatz	Es ist von einer zufälligen Streuung mit einer Spannweite von 0,003 Punkten für F1, 0,005 für IoU und 0,011 für MCC auszugehen, wenn DeepLabv3+ für 150k bis 300k Iterationen trainiert wird.
Welche Algorithmen der digitalen Geländeanalyse eignen sich zur Erkennung von Straßenschäden aus Oberflächenmodellen und welchen Einfluss hat die Wahl einer geomorphometrischen Vorverarbeitung der Höheninformation auf die Leistungsfähigkeit eines entsprechenden DeepLabv3+ Modells?	- Recherche von Analogien von Schädern zu Formen des Georeliefs - Identifikation und Parametrisierung der entsprechenden Algorithmen der digitalen Geländeanalyse: Hillshade, Topographic Openness, Top Hat Transform und Topographic Position Index - Training und Evaluierung von Deep-Learning-Modellen mit unterschiedlich vorverarbeiteten Höhenmodellen	Modelle, die auf Basis des Topographic Position Index trainiert wurden, liefern die höchsten F1-, IoU- und MCC-Werte. Die Unterschiede sind unter Berücksichtigung der Zufallskomponente des Trainings eher gering.
Mit welcher Qualität ist eine vollautomatische Erfassung von Oberflächenschäden unter Nutzung des semantischen Segmentierungsnetzwerks DeepLabv3+ möglich?	- Training von DeepLabv3+ Modellen auf umfangreichen, diversen Datensätzen - Transfer-Learning-Ansatz (vortrainiertes Backbone-CNN) - Evaluierung der Modelle auf separatem Testdatensatz	Es wurden drei Modelle trainiert, mit deren Hilfe Oberflächenarten, Objekte und Schäden für Asphalt-, Beton- und Pflasterdecken erkannt werden können. Die globale IoU-Metrik verweist auf eine Segmentierungsleistung der Modelle auf menschlichem Niveau.
Wie leistungsfähig ist ein auf Mobile Laserscanning und Deep Learning basierender, automatischer Inspektionsprozess im Vergleich zu einer konventionellen Inspektion im kommunalen Straßennetz?	- Grundlage für die Evaluierung ist die Systematik der Schadenserfassung in einem 1 m Raster gemäß ZTV ZEB-Stb (FGSV 2006) - Vergleich von manueller und automatischer Auswertung einer 2 km langen Strecke im Stil einer Eigenüberwachung gemäß ZTV ZEB-Stb (FGSV 2006) - Vergleich des entwickelten Analysesystems mit zwei Referenzsystemen (1,2 km Strecke)	Die Qualität der automatischen Schadenserfassung befindet sich innerhalb des Toleranzbereichs, den die Eigenüberwachung für geschultes Auswertepersonal definiert und nimmt dabei deutlich weniger Prozesszeit in Anspruch. Im Hinblick auf die deutschen Erfassungsstandards sind die Ergebnisse einer automatisch erzeugten Zustandserfassung auf Basis von PPS-Laserscans besser als die des internationalen Referenzsystems LCMS2.

6 Schlussbetrachtung

Die Ergebnisse der Arbeit zeigen, dass Mobile-Mapping-Systeme, die ein Applanix POS LV 420 und einen Fraunhofer PPS integriert haben, in der Lage sind, schnell und flächendeckend hochgenaue Laserscans von Fahrbahnoberflächen aufzunehmen. Werden die 3D-Punktwolken in adäquate Oberflächenmodelle überführt, können sie nicht nur auf Längs- und Querebenheit, sondern auch auf Oberflächenschäden messtechnisch analysiert und bewertet werden. Alle Schäden, die für die systematische Erhaltung kommunaler Straßen bewertungsrelevant sind, können bei innerortstypischen Geschwindigkeiten aus PPS-Oberflächenmodellen erkannt werden.

Das im Rahmen der vorliegenden Arbeit entwickelte Deep Learning basierte Analysesystem realisiert eine automatische Detektion von Oberflächenschäden auf Grundlage geometrischer und radiometrischer Informationen von PPS-Laserscans. Es ist nicht nur in der Lage, alle relevanten Schadensarten zu klassifizieren, sondern auch Schäden konturscharf als Geoobjekte mit einer absoluten Genauigkeit im Zenti- bis Dezimeterbereich zu kartieren. Damit liefert das Verfahren im Hinblick auf die Lage sowie die Längen- und Flächenausdehnung von Schadstellen, ein deutlich detaillierteres Detektionsniveau als die bisher manuelle Erfassung aus Kamerabildern durch geschultes Personal.

Übertragen auf die Auswertesystematik, welche die Richtlinien für das kommunale Straßennetz vorsehen konnte gezeigt werden, dass die automatisch erzeugte Zustandsbewertung dasselbe Qualitätsniveau einer durch manuelle Experten erzeugten Inspektion aufweist und dabei deutlich weniger Prozesszeit in Anspruch nimmt. Der reduzierte Zeit- und Personalbedarf für eine Inspektion durch eine Automatisierung der Schadstellenerkennung ermöglicht engere Inspektionsintervalle. Dadurch können Schäden frühzeitig erkannt und potentiell kostengünstiger durch zielgenaue Instandhaltungsmaßnahmen saniert werden. Schwankungen der Erfassungsqualität, wie sie bei einer manuellen Inspektion beispielsweise aufgrund von Ermüdung auftreten, können mit der automatisierten Erfassung von Oberflächenschäden eliminiert werden, wodurch eine objektivere Zustandsbewertung erreicht werden kann.

Trotz dieser Vorteile hängt die Effektivität einer Deep Learning basierten Inspektion von der Qualität und Zuverlässigkeit der trainierten Modelle und der ihnen zugrundeliegenden Daten ab. Ist eine bestimmte Oberflächensituation unzureichend oder gar nicht in den Trainingsdaten enthalten, wird das Modell in der Inferenzphase systematisch falsche Entscheidungen treffen. Bislang können solche Fehler nur im Zuge einer geeigneten Qualitätssicherung durch Fachpersonal erkannt und behandelt werden. Darüber hinaus wurden folgende Grenzen des Verfahrens im Rahmen der Untersuchungen festgestellt: Um sehr feine Risse zuverlässig

erkennen zu können, ist eine vergleichsweise niedrige Befahrungsgeschwindigkeit von ≤ 30 km/h vorteilhaft. Während dies für Innerortsbereiche kein Problem darstellt, ist dies für die Erfassung von Außerortsstraßen in der Praxis durchaus eine Einschränkung. Um die Anzahl gemessener Profile pro Sekunde angemessen zu steigern und somit in Fahrtrichtung eine dichtere Abtastung der Oberfläche auch bei höheren Fahrgeschwindigkeiten zu erreichen, müsste die Drehzahl des PPS Spiegelpolygonrades von derzeit 100 Hz auf mindestens 200 Hz erhöht werden. Hier böte sich Potential für eine technische Weiterentwicklung der Sensorik. Die Punktdichte pro Profil lässt sich auch bei aktuellen PPS Varianten bereits auf 1800 Punkte verdoppeln (FRAUNHOFER INSTITUTE FOR PHYSICAL MEASUREMENT TECHNIQUES IPM 2022).

Auch wenn Mobile Laserscanning nicht durch die vorherrschende Beleuchtungssituation determiniert wird, so existieren doch exogene Faktoren, die die Qualität der Oberflächenerfassung und der automatischen Erfassung des Oberflächenzustandes beeinflussen. So führt stehendes Wasser auf der Fahrbahnoberfläche dazu, dass an den betreffenden Stellen Lücken im Laserscan entstehen. Zwar legen die geltenden Richtlinien nahe, eine messtechnische Zustandserfassung generell nur bei trockenen Situationen durchzuführen, jedoch ist dies in der Praxis, insbesondere bei größeren kommunalen Straßennetzen, nicht immer möglich. Vor allem bereichsweise auftretende abtrocknende Nässe kann zu einer Zunahme von falsch positiven Detektionen von aufgelegten Flickstellen und Bindemittelanreicherungen führen, da hier eine gewisse Ähnlichkeit im Intensitätsbild gegeben ist. Ferner kann es vorkommen, dass PPS-Messungen auf stark retroreflektierenden Materialien die Oberflächenerfassung und -modellierung negativ beeinflussen. In manchen Aufnahmesituationen sorgen der Markierungsfarbe beigemischte Glasperlen zur Erhöhung der Nachsichtbarkeit von Fahrbahnmarkierungen bei Nässe dafür, dass die Entfernungsmessung des PPS verfälscht wird. In der Folge entstehen kleinräumige Depressionen im PPS-Oberflächenmodell, die zu falsch positiven Detektionen von Ausbrüchen führen können.

Folgende Aspekte wurden im Rahmen dieser Arbeit nicht betrachtet und bieten Raum für weiterführende Untersuchungen: Andere länderspezifische Auswertestandards, wie die von Österreich oder der Schweiz, erfordern Informationen zu den detektierten Schadstellen, die über die Anforderungen der deutschen Regelwerke hinausgehen. So muss beispielsweise eine Analyse des Schweregrads der Schädigung sowie eine feinere Differenzierung von beispielsweise Rissen in Einzel-, Netz-, Längs- oder Querrisse erfolgen, die mit dem entwickelten Verfahren noch nicht umgesetzt ist. Darüber hinaus ist die Modellierung von Unsicherheiten ein wichtiges Thema bei der Automatisierung von Prozessen. Die im Rahmen der Arbeit verwendete DeepLabv3+ Architektur gibt standardmäßig kein Unsicherheitsmaß für

die Prädiktion an. STRICKER et al. (2021a) schlagen u. a. den Monte Carlo Dropout Ansatz von GAL & GHAHRAMANI (2016) als potentiell zielführend zur Modellierung von Unsicherheiten im Kontext der Segmentierung von Straßenschäden aus Bilddaten vor. Eine weitere Möglichkeit wäre, ein Ensemble von DeepLabv3+ Modellen anzuwenden, die auf unterschiedlichen Teilen des Trainingsdatensatzes trainiert wurden (LAKSHMINARAYANAN, PRITZEL & BLUNDELL 2017). So wie menschliche Experten mit leicht unterschiedlichem Wissensstand eindeutige Situationen übereinstimmend bewerten, würde auch das Modell-Ensemble bei typischen Oberflächenschäden übereinstimmend entscheiden. In Grenzfällen dagegen sollten die einzelnen Modelle des Ensembles voneinander abweichende Vorhersagen treffen. Dies könnte als Indikator für die Sicherheit bzw. Unsicherheit einer Entscheidung im Rahmen einer expertengestützten Qualitätssicherung genutzt werden. Auch wenn eine robuste Unsicherheitsschätzung die Laufzeit einer Deep Learning basierten Schadstellenerkennung bei Verwendung derzeitiger Rechentechnik deutlich erhöht (STRICKER et al. 2021a), könnte dies durch zukünftige Weiterentwicklungen auf Seiten der Hardware durchaus kompensiert werden. Abschließend ist folgendes festzuhalten: Die Ergebnisse der Arbeit zeigen, dass es möglich ist, mit einem einzigen aktiven Sensor eine ganzheitliche Datengrundlage (Ebenheit und Oberflächenschäden) für die kommunale Straßenzustandserfassung zu erfassen, die nicht den Nachteilen passiver Kamerasensorik unterliegt. Eine Deep Learning basierte Erfassung von Oberflächenschäden auf der Grundlage von PPS-Oberflächenmodellen kann im Sinne einer systematischen Straßenerhaltung zu einer schnelleren und objektiveren Zustandsbewertung beitragen.

Literaturverzeichnis

- ABADI, M., AGARWAL, A., BARHAM, P., BREVDO, E., CHEN, Z., CITRO, C., CORRADO, G. S., DAVIS, A., DEAN, J., DEVIN, M., GHEMAWAT, S. & GOODFELLOW, I. (2016): TensorFlow: a system for large-scale machine learning. Proceedings of the 12th USENIX conference on Operating Systems Design and Implementation (OSDI'16), pp. 265-283.
- AGHDAM, H. H. & HERAVI, E. J. (2017): Guide to Convolutional Neural Networks. A Practical Application to Traffic-Sign Detection and Classification. Berlin, Schweiz: Springer.
- AIT (2019): Straßenzustandserfassung mit dem RoadSTAR. Abgerufen am 08. August 2022 von https://www.ait.ac.at/fileadmin//mc/mobility/downloads/Rail_Road_Infrastructure/RoadSTAR_-_Leistungskatalog.pdf
- ALASTAL, A. I. & SHAQFA, A. H. (2022): GeoAI Technologies and Their Application Areas in Urban Planning and Development: Concepts, Opportunities and Challenges in Smart City (Kuwait, Study Case). Journal of Data Analysis and Information Processing, 10, pp. 110-126.
- ALLOGHANI, M., AL-JUMEILY, D., MUSTAFINA, J., HUSSAIN, A. & ALJAAF, A. J. (2020): A Systematic Review on Supervised and Unsupervised Machine Learning Algorithms for Data Science. In: M. BERRY, A. MOHAMED & B. YAP [Hrsg.], Supervised and Unsupervised Learning (pp. 3-21). Cham: Springer.
- ALZUBAIDI, L., ZHANG, J., HUMAIDI, A. J., AL-DUJAILI, A., DUAN, Y., AL-SHAMMA, O., SANTAMARÍA, J., FADHEL, M. A., AL-AMIDIE, M. & FARHAN, L. (2021): Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. Journal of Big Data, 8(53), pp. 1-74.
- AMHAZ, R., CHAMBON, S., IDIER, J. & BALTAZART, V. (2016): Automatic crack detection on two-dimensional pavement images: An algorithm based on minimal path selection. IEEE Transactions on Intelligent Transportation Systems, 17(10), pp. 2718-2729.
- APPLANIX (2019): POS LV. Abgerufen am 22. Februar 2023 von <https://www.applanix.com/downloads/products/specs/POS-LV-Datasheet.pdf>

- ASTM E1703/E1703M-10 (2015): Standard Test Method for Measuring Rut-Depth of Pavement Surfaces Using a Straightedge. West Conshohocken, Pennsylvania, USA: ASTM International.
- BADRINARAYANAN, V., KENDALL, A. & CIPOLLA, R. (2017): SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(12), pp. 2481-2495.
- BALCK, H., SCHÜLLER, H., BALMBERGER, M. & ROSSOL, C. (2017): Verfahren zur Zusammenführung von Informationen unterschiedlicher Netzanalysesysteme. Bremen: Carl Schünemann Verlag GmbH.
- BAST (2018a): ZEB - Zustandserfassung und -bewertung von Straßen. Abgerufen am 22. Februar 2023 von https://www.bast.de/DE/Strassenbau/Fachthemen/ZEB.pdf;jsessionid=D333ADAD0965FAB2A2F4EF85D856B1A6.live11314?__blob=publicationFile&v=2
- BAST (2018b): Annual Report 2018. Bundesanstalt für Straßenwesen. Bremen: Fachverlag NW.
- BAST (2022): Referenzstrecke für Messsysteme zur Erfassung der Griffigkeit. Abgerufen am 22. Februar 2023 von https://www.durabast.de/durabast/DE/Referenzstrecken/Griffigkeit/Griffigkeit_node.html;jsessionid=F7061E284CCD71844F099ED8C227379A.live21302
- BATER, C. W. & COOPS, N. C. (2009): Evaluating error associated with LiDAR-derived DEM interpolation. Computers & Geosciences, 35(2), pp. 289-300.
- BAUER, M. (2017): Vermessung und Ortung mit Satelliten. Globale Navigationssatellitensysteme (GNSS) und andere satellitengestützte Navigationssysteme. Berlin: Wichmann.
- BAYERISCHE VERMESSUNGSVERWALTUNG (2022a): Digitales Geländemodell. Abgerufen am 25. August 2022 von https://geodatenonline.bayern.de/geodatenonline/krhuvUCGRJa3eNONjQkco9G-H0kKy6-DgLfsvLNvr7Q_MO6N4SQAE-GMB6pjZsAl5mDahPG5N80NGg6PmB0SsJiLW2H0YejF/krhd0/ZsA27/eNO55/PG5bd

- BAYERISCHE VERMESSUNGSVERWALTUNG (2022b): Digitale Orthophotos. Abgerufen am 25. August 2022 von https://geodatenonline.bayern.de/geodatenonline/krhuvUCGRJa3eNONjQkco9G-H0kKy6-DgLfsvLNvr7Q_MO6N4SQAE-GMB6pjZsAlg1yZ7NiPBpaJJguNZCObt6nWxLXC6C9Q/krhd0/ZsA27/eNO55/NiPd4
- BELIKOV, V. V., IVANOV, V. D., KONTOROVICH, V. K., KORYTNIK, S. A. & SEMENOV, A. Y. (1997): The non-Sibsonian interpolation: a new method of interpolation of the value of a function on an arbitrary set of points. *Computational Mathematics and Mathematical Physics*, 37(1), pp. 9-15.
- BENINI, E. & PONZA, R. (2010): Nonparametric fitting of aerodynamic data using smoothing thin-plate splines. *American Institute of Aeronautics and Astronautics Journal*, 48(7), pp. 1403-1419.
- BENZ, U. C., HOFMANN, P., WILLHAUCK, G., LINGENFELDER, I. & HEYNEN, M. (2004): Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(3-4), pp. 239-258.
- BERALDIN, J., BLAIS, F. & LOHR, U. (2010): Laser scanning technology. In: G. VOSSelman & H. G. MASS [Hrsg.], *Airborne and terrestrial laser scanning* (pp. 1-42). Caithness: Whittles Publishing.
- BERALDIN, J.-A. (2004): Integration of Laser Scanning and Close-Range Photogrammetry - The Last Decade and Beyond. *Proceedings 20th ISPRS Congress (Istanbul, 12-23 July 2004)*, pp. 12-23.
- BERTHOLD, D., GROßMANN, A., HEIMRICH, N., NABER, T., TRÜMPER, J., FROHBÖSE, B. & VILLARET, S. (2023): Ermittlung der Vergleichs- und Wiederholstreubereiche bei der Erfassung von Substanzmerkmalen (Oberfläche). FE 04.0261/2012/MRB. Bergisch-Gladbach: Bundesanstalt für Straßenwesen (BASt) – in Vorbereitung.
- BESL, P. J. & MCKAY, N. D. (1992): A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14, pp. 239-256.

- BEUTEL, A., MØLHAVE, T. & AGARWAL, P. K. (2010): Natural neighbor interpolation based grid DEM construction using a GPU. Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems (GIS '10), pp. 172-181.
- BHUNIA, G. S., SHIT, P. K. & MAITI, R. (2018): Comparison of GIS-based interpolation methods. Journal of the Saudi Society of Agricultural Sciences, 17, pp. 114-126.
- BIENERT, A., GEORGI, L., KUNZ, M., MAAS, H.-G. & VON OHEIMB, G. (2018): Comparison and Combination of Mobile and Terrestrial Laser Scanning for Natural Forest Inventories. Forests, 9(7), pp. 1-25.
- BILL, R. (2016): Grundlagen der Geo-Informationssysteme. Berlin-Offenbach: Wichmann.
- BJORCK, J., GOMES, C., SELMAN, B. & WEINBERGER, K. Q. (2018): Understanding Batch Normalization. Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS), pp. 7705-7716.
- BLANKENBACH, J. (2017): As-built Dokumentation für BIM. In: T. LUHMANN & C. SCHUMACHER [Hrsg.], Photogrammetrie – Laserscanning – Optische 3D-Messtechnik – Beiträge der Oldenburger 3D-Tage (pp. 290-298). Berlin: Wichmann/VDE Verlag.
- BLASCHKE, T. & STROBL, J. (2001): What's wrong with pixels? Some recent developments interfacing remote sensing and GIS. GIS - Zeitschrift für Geoinformationssysteme, 6, pp. 12-17.
- BLEßMANN, W., BÖHM, S., ROSAUER, V. & SCHÄFER, V. (2010): ZTV BEA-StB – Handbuch und Kommentar. Bonn: Kirschbaum Verlag.
- BLUMENFELD, T. J. (2020): Analyse und Prognose von Straßenzustandsdaten mit Hilfe von probabilistischen Methoden und der Strukturgleichungsmodellierung. Dissertation, TU Darmstadt, Department of Civil and Environmental Engineering Science.
- BÖHNER, J. & CONRAD, O. (2011): Satelliten- und flugzeuggestützte Oberflächenparameter zur Klassifikation von stadtklimatisch relevanten Stadtstrukturen in SAGA. In: W. KRESSE, C. JÜRGENS, L. PLÜMER, F. ROTTENSTEINER & E. SEYFERT [Hrsg.], Photogrammetrie - Fernerkundung - Geoinformation (Vol. 5, pp. 397-397). Stuttgart: E. Schweizerbart'sche Verlagsbuchhandlung.

- BOLT, H., MORRIS, J., PEDRALI, M., ANTÃO, P. & SOARES, C. G. (2010): Techniques for Human Reliability Evaluation. In: C. G. SOARES [Hrsg.], Safety and Reliability of Industrial Products, Systems and Structures (pp. 141-156). Boca Raton: CRC Press.
- BRUNKEN, H. (2021): Stereo vision-based road condition monitoring. Dissertation, Technische Universität Berlin, Fakultät IV - Elektrotechnik und Informatik, Berlin.
- CAKIR, L. (2011): Modified quadratic shepard and radial basis functions interpolations for determination of orthometric heights from global positioning system (GPS) heights. International Journal of the Physical Sciences, 6(36), pp. 8183-8187.
- CANZLER, U. & WINKLER, B. (2012): Weiterentwicklung der automatisierten Merkmalerkennung im Rahmen des TP3. Bremerhaven: Wirtschaftsverlag NW.
- CĂTEANU, M. & CIUBOTARU, A. (2020): Accuracy of Ground Surface Interpolation from Airborne Laser Scanning (ALS) Data in Dense Forest Cover. ISPRS International Journal of Geo-Information, 9(224), pp. 1-20.
- CHAMBON, S. & MOLIARD, J. M. (2011): Automatic road pavement assessment with image processing: Review and comparison. International Journal of Geophysics, pp. 1-20.
- CHAZETTE, P., TOTEMS, J., HESPEL, L. & BAILLY, J.-S. (2016): Principle and Physics of the LiDAR Measurement. In: N. BAGHDADI & M. ZRIBI [Hrsg.], Optical Remote Sensing of Land Surface (pp. 201-247). London: Elsevier.
- CHEN, L.-C., PAPANDREOU, G., KOKKINOS, I., MURPHY, K. & YUILLE, A. L. (2018a): DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution and Fully Connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(4), pp. 834-848.
- CHEN, L.-C., PAPANDREOU, G., SCHROFF, F. & ADAM, H. (2017): Rethinking Atrous Convolution for Semantic Image Segmentation. Abgerufen am 23. Februar 2023 von <https://arxiv.org/pdf/1706.05587.pdf>
- CHEN, L.-C., ZHU, Y., PAPANDREOU, G., HUI, H., COLLINS, M. D. & LIU, T. (2021): DeepLab: Deep Labelling for Semantic Image Segmentation. Abgerufen am 23. Februar 2023 von <https://github.com/tensorflow/models/tree/master/research/deeplab>

- CHEN, L.-C., ZHU, Y., PAPANDREOU, G., SCHROFF, F. & ADAM, H. (2018b): Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. European Conference on Computer Vision – ECCV 2018 (pp. 833-851). Springer International Publishing.
- CHICCO, D. & JURMAN, G. (2020): The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics, 21(6), pp. 1-13.
- CHOLLET, F. (2017): Xception: Deep Learning with Depthwise Separable Convolutions. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1800-1807.
- CHOLLET, F. (2021): Deep Learning with Python. New York: Manning.
- CLARKE, K. C. (1995): Analytical and Computer Cartography. Upper Saddle River (New Jersey): Prentice Hall (Pearson Education, Inc.).
- CLOUDCOMPARE (2022): github. Abgerufen am 21. November 2022 von <https://github.com/CloudCompare/CloudCompare>
- CONRAD, O., BECHTEL, B., BOCK, M., DIETRICH, H., FISCHER, E., GERLITZ, L., WEHBERG, J., WICHMANN, V. & BÖHNER, J. (2015): System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. Geoscientific Model Development, 8, pp. 1991-2007.
- CORDTS, M., OMRAN, M., RAMOS, S., REHFELD, T., ENZWEILER, M., BENENSON, R., FRANKE, U., ROTH, S. & SCHIELE, B. (2016): The cityscapes dataset for semantic urban scene understanding. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1-11.
- CSURKA, G., LARLUS, D. & PERRONNIN, F. (2013): What is a good evaluation measure for semantic segmentation? The British Machine Vision Conference (BMVC), pp. 1-11.
- DE BLASIIS, M. R., DI BENEDETTO, A. & FIANI, M. (2020): Mobile Laser Scanning Data for the Evaluation of Pavement Surface Distress. Remote Sensing, 12(6), pp. 1-25.
- DEGLE, S. (2006): Arbeit und Sehen. Eine interdisziplinäre Erklärung von Veränderungen des Sehens durch Bildschirmarbeit. Dissertation, Universität Augsburg, Wirtschaftswissenschaftliche Fakultät.

- DELL'ACCIO, F. & DI TOMMASO, F. (2016): Scattered data interpolation by Shepard's like methods: classical results and recent advances. Dolomites Research Notes on Approximation - Proceedings of Kernel-based Methods and Function Approximation., 9, pp. 32-44.
- DENG, J., DONG, W., SOCHER, R., LI, L.-J., LI, K. & FEI-FEI, L. (2009): ImageNet: A large-scale hierarchical image database. IEEE conference on computer vision and pattern recognition, pp. 248-255.
- DEPPE, O., DORNER, G., KÖNIG, S., MARTIN, T., VOIGT, S. & ZIMMERMANN, S. (2017): MEMS and FOG Technologies for Tactical and Navigation Grade Inertial Sensor - Recent Improvements and Comparison. Sensors, 17(3), pp. 1-22.
- DESTATIS (2023): Etwa jedes achte Unternehmen nutzt künstliche Intelligenz (Pressemitteilung Nr. 453 vom 27. November 2023). Abgerufen am 13. Dezember 2023 von https://www.destatis.de/DE/Presse/Pressemitteilungen/2023/11/PD23_453_52911.html
- DIN EN 13036 (2020): Oberflächeneigenschaften von Straßen und Flugplätzen - Prüfverfahren. Berlin: Beuth Verlag.
- DIN EN 60825 (2013): Sicherheit von Lasereinrichtungen - Beiblatt 13: Messungen zur Klassifizierung von Lasereinrichtungen (IEC/TR 60825-13:2011). Berlin: Beuth Verlag.
- DIN ISO 3534-1 (2009): Statistik - Begriffe und Formelzeichen - Teil 1: Wahrscheinlichkeit und allgemeine statistische Begriffe (ISO 3534-1:2006). Berlin: Beuth-Verlag.
- DINAS, S. & BAÑÓN, J. M. (2014): A review on delaunay triangulation with application on computer vision. International Journal of Computer Science and Engineering, 3(2), pp. 9-18.
- DÖBEL, I., LEIS, M., VOGELSANG, M. M., NEUSTROEV, D., PETZKA, H., RIEMER, A., RÜPING, S., VOSS, A., WEGELE, M. & WELZ, J. (2018): Maschinelles Lernen - Kompetenzen, Anwendungen und Forschungsbedarf. München: Fraunhofer-Gesellschaft.
- DOGO, E. M., AFOLABI, O. J. & TWALA, B. (2022): On the Relative Impact of Optimizers on Convolutional Neural Networks with Varying Depth and Width for Image Classification. Applied Sciences, 12(23), pp. 1-36.

- DÖLLNER, J. (2020): Geospatial Artificial Intelligence: Potentials of Machine Learning for 3D Point Clouds and Geospatial Digital Twins. PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science, 88, pp. 15-24.
- DONATO, G. & BELONGIE, S. (2002): Approximation Methods for Thin Plate Spline Mappings and Principal Warps. In: A. HEYDEN, G. SPARR, M. NIELSEN & P. JOHANSEN [Hrsg.], Computer Vision - ECCV 2002: 7th European Conference on Computer Vision, Copenhagen, Denmark, May 28-31, 2002., Proceedings, Part III, Lecture Notes in Computer Science, pp. 21 - 31.
- DONEUS, M. (2013): Openness as Visualization Technique for Interpretative Mapping of Airborne Lidar Derived Digital Terrain Models. Remote Sensing, 5(12), pp. 6427-6442.
- DÖRN, S. (2018): Programmieren für Ingenieure und Naturwissenschaftler - Intelligente Algorithmen und digitale Technologien. Berlin: Springer Vieweg.
- DRENTH, K., TAN, J. Y., DRENTH, M. & KIT, O. J. (2020): LCMS-2 Measurements of the Quality of Road Markings. In: C. RAAB [Hrsg.], Proceedings of the 9th International Conference on Maintenance and Rehabilitation of Pavements—Mairepav9. Lecture Notes in Civil Engineering, 76, pp. 943-952.
- DUAN, W., CHIANG, Y.-Y., LEYK, S., UHL, J. H. & KNOBLOCK, C. A. (2020): Automatic alignment of contemporary vector data and georeferenced historical maps using reinforcement learning. International Journal of Geographical Information Science, 34(4), pp. 824-849.
- DUMITRU, P. D., PLOPEANU, M. & BADEA, D. (2013): Comparative study regarding the methods of interpolation. Recent Advances in Geodesy and Geomatics Engineering, pp. 45-52.
- DYNATEST US, INC. (2022): Multi Functional Vehicle. Abgerufen am 10. August 2022 von <https://dynatest.com/equipment/multi-functional-vehicle/>
- EICHLER, J. & EICHLER, H.-J. (1991): Laser - Grundlagen Systeme Anwendungen. G. Herzinger & H. Weber [Hrsg.], Heidelberg: Springer.
- EICHLER, J., DÜNKEL, D. & EPPICH, B. (2004): Die Strahlqualität von Lasern – Wie bestimmt man Beugungsmaßzahl und Strahldurchmesser in der Praxis? Lasertechnik Journal, 1(2), 63-66.

- EISENBACH, M., STRICKER, R., SEICHTER, D., AMENDE, K., DEBES, K., SESSELMANN, M., EBERSBACH, D., STÖCKERT, U. & GROSS, H.-M. (2017): How to get pavement distress detection ready for deep learning? A systematic approach. International Joint Conference on Neural Networks (IJCNN), pp. 2039-2047.
- EISENFÜHR, F., WEBER, M. & LANGER, T. (2010): Rationales Entscheiden. Berlin, Heidelberg: Springer-Verlag.
- EVERINGHAM, M., ESLAMI, S. M., VAN GOOL, L., WILLIAMS, C. K., WINN, J. & ZISSERMAN, A. (2015): The PASCAL Visual Object Classes Challenge: A Retrospective. International Journal of Computer Vision, 111(1), pp. 98-136.
- FAHRMEIR, L., HEUMANN, C., KÜNSTLER, R., PIGEOT, I. & TUTZ, G. (2016): Statistik - Der Weg zur Datenanalyse. Berlin Heidelberg: Springer Spektrum.
- FAN, Z., LI, C., CHEN, Y., MASCI, P. D., CHEN, X., ZHU, G. & LOPRENCIPE, G. (2020): Ensemble of Deep Convolutional Neural Networks for Automatic Pavement Crack Detection and Measurement. Coatings, 10(152), pp. 1-14.
- FARMAKIS-SEREBRYAKOVA, M. & HURNI, L. (2020): Comparison of Relief Shading Techniques Applied to Landforms. ISPRS International Journal of Geo-Information, pp. 1-22.
- FERNANDEZ-MORAL, E., MARTINS, R., WOLF, D. & RIVES, P. (2018): A New Metric for Evaluating Semantic Segmentation: Leveraging Global and Contour Accuracy. 2018 IEEE Intelligent Vehicles Symposium (IV), pp. 1051-1056.
- FGSV (2006): Zusätzliche Technische Vertragsbedingungen und Richtlinien zur Zustandserfassung und -bewertung von Straßen. Köln: FGSV Verlag GmbH.
- FGSV (2007): Technische Prüfvorschriften für Griffigkeitsmessungen im StraßenbauTeil: Seitenkraftmessverfahren (SKM) - TP Griff-StB (SKM) . Köln: FGSV Verlag GmbH.
- FGSV (2009a): Technische Prüfvorschriften für Ebenheitsmessungen auf Fahrbahnoberflächen in Längs- und Querrichtung, Teil: Berührungslose Messungen (TP Eben - Berührungslose Messungen). Köln: FGSV Verlag GmbH.
- FGSV (2009b): Merkblatt für die Herstellung von Oberflächentexturen auf Verkehrsflächen aus Beton - M OB. Köln: FGSV Verlag GmbH.

- FGSV (2012): Empfehlungen für das Erhaltungsmanagement von Innerortsstraßen. Köln: FGSV Verlag GmbH.
- FGSV (2015): Arbeitspapiere zur Systematik der Straßenerhaltung. Reihe K: Kommunale Straßen (AP9 K/3.1). Köln: FGSV Verlag GmbH.
- FGSV (2016): Arbeitspapiere zur Systematik der Straßenerhaltung. Reihe K: Kommunale Straßen. Köln: FGSV Verlag GmbH.
- FGSV (2020): Technische Prüfvorschriften für die Erfassung von Substanzmerkmalen (Oberfläche) mit schnellfahrenden Messsystemen (Teil: Bildaufnahme- und Auswertetechnik). Köln: FGSV Verlag.
- FRANKE, R. (1982): Scattered data interpolation: tests of some methods. *Mathematics of Computation*, 38, pp. 181-200.
- FRANKE, R. & NIELSON, G. (1980): Smooth Interpolation of Large Sets of Scattered Data. *International Journal for Numerical Methods in Engineering*, 15, pp. 1691-1704.
- FRAUNHOFER INSTITUTE FOR PHYSICAL MEASUREMENT TECHNIQUES IPM (2018): A close-up inspection: Detecting cracks in road surfaces. Abgerufen am 22. Februar 2023 von <https://www.ipm.fraunhofer.de/content/dam/ipm/en/PDFs/reports/AR-articles-2018-2019/AR-2018-crack-detection-road-surface-laserscanning.pdf>
- FRAUNHOFER INSTITUTE FOR PHYSICAL MEASUREMENT TECHNIQUES IPM (2022): Pavement Profile Scanner PPS / PPS-Plus. Abgerufen am 22. Februar 2023 von <https://www.ipm.fraunhofer.de/content/dam/ipm/en/PDFs/product-information/OF/MTS/Pavement-Profile-Scanner-PPS.pdf>
- FREEMAN, J., BENTLEY, K., MOORE, R., CHAMBERS, S. & QUESENBERRY, G. (2019): Open Geospatial Consortium - Discussion Papers. Abgerufen am 27. November 2022 von https://portal.ogc.org/files/?artifact_id=82553
- FRITSCH, D. (2015): Punktwolken, und was dann? *Vermessung & Geoinformation*, 1, pp. 54-66.
- FUGRO N.V. (2022): Fugro Roadware. Abgerufen am 10. August 2022 von <https://www.fugro.com/our-services/asset-integrity/roadware>

- GAL, Y. & GHAHRAMANI, Z. (2016): Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. Proceedings of the 33rd International Conference on Machine Learning (PMLR), 48, pp. 1050-1059.
- GARNERO, G. & GODONE, D. (2013): Comparisons between different interpolation techniques. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-5(W3), pp. 139-144.
- GENSER, A., SPIELHOFER, R., NITSCHKE, P. & KOUVELAS, A. (2022): Ride comfort assessment for automated vehicles utilizing a road surface model and Monte Carlo simulations. Computer-Aided Civil and Infrastructure Engineering, 37(20), pp. 1316-1334.
- GÉZERO, L. & ANTUNES, C. (2019): Road Rutting Measurement Using Mobile LiDAR Systems Point Cloud. ISPRS International Journal of Geo-Information, 8(9), pp. 1-16.
- GLENNIE, C. (2007): Rigorous 3D error analysis of kinematic scanning LIDAR systems. Journal of Applied Geodesy, 1(3), pp. 147-157.
- GLENNIE, C. (2009): Kinematic Terrestrial Light-Detection and Ranging System for Scanning. Transportation Research Record, 2105(1), pp. 135-141.
- GONZALES-INCA, C., CALLE, M., CROGHAN, D., TORABI HAGHIGHI, A., MARTTILA, H., SILANDER, J. & ALHO, P. (2022): Geospatial Artificial Intelligence (GeoAI) in the Integrated Hydrological and Fluvial Systems Modeling: Review of Current Applications and Trends. Water, 14, pp. 1-38.
- GOPALAKRISHNAN, K., KHAITAN, S. K., CHOUDHARY, A. & AGRAWAL, A. (2017): Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection. Construction and Building Materials(157), pp. 322-330.
- GORDON, B. (2008): Zur Bestimmung von Messunsicherheiten terrestrischer Laserscanner. Dissertation, Technischen Universität Darmstadt, Fachbereich Bauingenieurwesen und Geodäsie.
- GRÄFE, G. (2009): Kinematische Anwendungen von Laserscannern im Straßenraum. Dissertation, Universität der Bundeswehr München, Fakultät für Bauingenieur- und Vermessungswesen.

- GREJNER-BRZEZINSKA, D. (1999): Direct Exterior Orientation of Airborne Imagery with GPS/INS System: Performance Analysis. *Navigation*, 4, pp. 261-270.
- GROßMANN, A., SCHELLER, S., WILHELM, S., UECKERMAN, A., JAKOBI, V., REITERER, A., LEIDINGER, M. & FREY, C. (2023): Fortschreibung von Qualitätsstandards zur Abnahme von Ebenheitsmesssystemen für ZEB- und Abnahmemessungen vor dem Hintergrund neuer Erfassungstechnologien. FE 04.0326/2018. Berglisch-Gladbach: Bundesanstalt für Straßenwesen (BASt) – in Vorbereitung.
- GROVES, P. (2013): Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems. Boston: Artech.
- GRUNAU, W. (2022): Künstliche Intelligenz in Geodäsie und Geoinformatik. Potenziale und Best-Practice-Beispiele. Berlin: Wichmann Verlag.
- GUAN, H., LI, J., CAO, S. & YU, Y. (2016): Use of mobile LiDAR in road information inventory: a review. *International Journal of Image and Data Fusion*, 7(3), 219-242.
- GUYOT, A., LENNON, M. & HUBERT-MOY, L. (2021): Objective comparison of relief visualization techniques with deep CNN for archaeology. *Journal of Archaeological Science: Reports*(38), pp. 1-8.
- HAIGERMOSER, A., LUBER, B., RAUH, J. & GRÄFE, G. (2015): Road and track irregularities: measurement, assessment and simulation. *Vehicle System Dynamics*, pp. 878-957.
- HE, K., ZHANG, X., REN, S. & SUN, J. (2016): Deep Residual Learning for Image Recognition. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770-778.
- HE, Y., LIANG, B., YANG, J., LI, S. & HE, J. (2017): An Iterative Closest Points Algorithm for Registration of 3D Laser Scanner Point Clouds with Geometric Features. *Sensors*, 17, pp. 1-16.
- HELLER INGENIEURGESELLSCHAFT MBH (2018): Rohdaten der Straßenzustandserfassung - Formatdefinition auf XML-Basis. Abgerufen am 21. Januar 2023 von https://itzeb.heller-ig.de/files/Rohdatenformat_180517_V_3_0_1_8_v_Doc_1_0_1.pdf
- HELLER INGENIEURGESELLSCHAFT MBH (2022): Portal - IT-ZEB Server. Teilprojekt 4 (TP4). Abgerufen am 06. Januar 2023 von https://itzeb.heller-ig.de/zeb_tp4.html

- HENGL, T. & EVANS, I. S. (2009): Mathematical and Digital Models of the Land Surface. In: T. HENGL & H. I. REUTER [Hrsg.], Geomorphometry - Concepts, Software, Applications. Developments in Soil Science (Vol. 33, pp. 31-63). Amsterdam: Elsevier.
- HENGL, T. & REUTER, H. I. (2009): Geomorphometry: Concepts, Software, Applications. Developments in Soil Science (Vol. 33). Amsterdam: Elsevier.
- HENNES, M. (2007): Konkurrierende Genauigkeitsmaße – Potential und Schwächen aus der Sicht des Anwenders. Allgemeine Vermessungsnachrichten, 4, pp. 136-146.
- HEYWOOD, I., CORNELIUS, S. & CARVER, S. (2006): An introduction to geographical information systems. Harlow: Pearson Prentice Hall.
- HILLER, J. K. & SMITH, M. (2008): Residual relief separation: digital elevation model enhancement for geomorphological mapping. Earth Surface Processes and Landforms, 33(14), pp. 2266-2276.
- HINÜBER, E. (2002): Inertiale Messsysteme mit faseroptischen Kreisel für Fahrdynamik und Topologiedaten-Erfassung. Automobiltechnische Zeitschrift ATZ, 104, pp. 584-591.
- HINZEN, A. (1993): Der Einfluß des menschlichen Fehlers auf die Sicherheit der Eisenbahn. Dissertation, Rheinisch-Westfälische Technische Hochschule Aachen, Verkehrswissenschaftliches Institut.
- HOFFMEISTER, W. (2000): Investitionsrechnung und Nutzwertanalyse. Eine entscheidungsorientierte Darstellung mit vielen Beispielen und Übungen. Stuttgart, Berlin, Köln: Kohlhammer.
- HOFMANN, S. (2017): Potential von LiDAR Mobile Mapping für hochgenaue Karten. Dissertation, Leibniz Universität Hannover, Fachrichtung Geodäsie und Geoinformatik.
- HONG, T., LEE, M., KOO, C., JEONG, K. & KIM, J. (2017): Development of a method for estimating the rooftop solar photovoltaic (PV) potential by analyzing the available rooftop area using Hillshade analysis. Applied Energy, 194, pp. 320-332.
- HORN, B. K. (1981): Hill Shading and the Reflectance Map. Proceedings of the IEEE, 69(1), pp. 14-47.

- HOU, L., SAMARAS, D., KURC, T. M., GAO, Y., DAVIS, J. E. & SALTZ, J. H. (2016): Patch-based Convolutional Neural Network for Whole Slide Tissue Image Classification. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2424-2433.
- INFRALAB (2019): Abnahmemessungen ARAN. Abgerufen am 22. Februar 2023 von <https://infralab.ch/de/project/abnahmemessungen/>
- ISHII, M. & SATO, A. (2018): Layer-Wise Weight Decay for Deep Neural Networks. In: M. PAUL, C. HITOSHI & Q. HUANG [Hrsg.], Image and Video Technology (Lecture Notes in Computer Science) (pp. 276-289). Cham: Springer.
- JACOT, A., LINDENMANN, H. P. & SEILER, L. (2007): Grundlagen zur Revision der Griffigkeitsnormen (Vol. 1202). Bern: Eidgenössisches Departement für Umwelt, Verkehr, Energie und Kommunikation UVEK.
- JÄHNE, B. (2012): Digitale Bildverarbeitung und Bildgewinnung. Berlin: Springer.
- JANOWICZ, K., GAO, S., MCKENZIE, G., HU, Y. & BHADURI, B. (2020): GeoAI: spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond. International Journal of Geographical Information Science, 34(4), pp. 625-636.
- JEKELI, C. (2001): Inertial Navigation systems with geodetic applications. Berlin: de Gruyter.
- JOHNSON, N. S., VULIMIRI, P. S., TO, A. C., ZHANG, X., BRICE, C. A., KAPPES, B. B. & STEBNER, A. P. (2020): Invited review: Machine learning for materials developments in metals additive manufacturing. Additive Manufacturing, 36, pp. 1-30.
- JURGIEL, B., VERCHERE, P., TOURIGNY, E. & BECERRA, J. (2022): Profile Tool. Abgerufen am 23. Februar 2023 von <https://github.com/PANOimagen/profiletool>
- JUTZI, B. (2007): Analyse der zeitlichen Signalform von rückgestreuten Laserpulsen. Dissertation, Technischen Universität München, Fachgebiet Photogrammetrie und Fernerkundung.
- KAHMEN, H. (2011): Angewandte Geodäsie: Vermessungskunde. Berlin: De Gruyter.
- KAPELA, R., ŚNIATAŁA, P., TURKOT, A., RYBARCZYK, A., POŻARYCKI, A., RYDZEWSKI, P., WYCZAŁEK, M. & BŁOCH, A. (2015): Asphalt surfaced pavement cracks detection based on histograms of oriented gradients. 22nd International Conference Mixed Design of Integrated Circuits & Systems (MIXDES), pp. 579-584.

- KAPLAN, E. & HEGARTY, C. (2017): Understanding GPS/GNSS: Principles and Applications. Boston: Artech House.
- KARACA, Y., CATTANI, C. & MOONIS, M. (2017): Comparison of Deep Learning and Support Vector Machine Learning for Subgroups of Multiple Sclerosis. Computational Science and Its Applications – ICCSA 2017, pp. 142-153.
- KELLER, W. & BORKOWSKI, A. (2019): Thin plate spline interpolation. Journal of Geodesy, 93, pp. 1251-1269.
- KEMPF, M. (2020): Fables of the past: landscape (re-)constructions and the bias in the data. Documenta Praehistorica, 47, pp. 476-492.
- KERN, F. (2003): Automatisierte Modellierung von Bauwerksgeometrien aus 3D-Laserscanner-Daten. Dissertation, Technische Universität Braunschweig, Fachbereich Bauingenieurwesen.
- KIM, T. & RYU, S. K. (2014): Review and analysis of pothole detection methods. Journal of Emerging Trends in Computing and Information Sciences, pp. 603-608.
- KIMURA, N., YOSHINAGA, I., SEKIJIMA, K., AZECHI, I. & BABA, D. (2020): Convolutional Neural Network Coupled with a Transfer-Learning Approach for Time-Series Flood Predictions. Water , 12(1), pp. 1-19.
- KLINGBEIL, L. & ROMANOVAS, M. (2012): Rekursive Bayesische Schätzverfahren zur Lokalisierung von Personen innerhalb von Gebäuden. zfv – Zeitschrift für Geodäsie, Geoinformation und Landmanagement, 6, pp. 365-376.
- KNITTER, D., BROZIO, J., HAMER, W., DUTTMANN, R., MÜLLER, J. & NAKOINZ, O. (2019): Transformations and Site Locations from a Landscape Archaeological Perspective: The Case of Neolithic Wagrien, Schleswig-Holstein, Germany. Land, 8(4), pp. 1-17.
- KOHLSTOCK, P. (2011): Topographie: Methoden und Modelle der Landesaufnahme. Berlin/New York: de Gruyter.
- KORNBLITH, S., SHLENS, J. & LE, Q. V. (2019): Do better imagenet models transfer better? Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2661-2671.

- KOTULAK, K., FROŃ, A., KRANKOWSKI, A., PULIDO, G. O. & HENRANDEZ-PAJARES, M. (2017): Sibsonian and non-Sibsonian natural neighbour interpolation of the total electron content value. *Acta Geophysica*, 65, pp. 13-28.
- KRÄHENBÜHL, P. & KOLTUN, V. (2011): Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials. *Advances in Neural Information Processing Systems (NIPS)*, 24, pp. 1-9.
- KRAUSE, G. (2000): Abgrenzung von Mängelklassen zur Kennzeichnung von wahrscheinlichen Schadensursachen mit Hilfe meßtechnisch erfaßter Bestands- und Zustandsmerkmale (Vol. 790). Bonn: Bundesministerium für Verkehr und digitale Infrastruktur.
- KRITTANAWONG, C., JOHNSON, K. W., ROSENSON, R. S., WANG, Z., AYDAR, M., BABER, U., MIN, J. K., TANG, W. H., HALPERIN, J. L. & NARAYAN, S. M. (2019): Deep learning for cardiovascular medicine: a practical primer. *European Heart Journal*, 40(25), pp. 2058–2073.
- KRIZHEVSKY, A., SUTSKEVER, I. & HINTON, G. E. (2012): ImageNet classification with deep convolutional neural networks. In: P. L. BARTLETT, F. PEREIRA, C. J. BURGESS, L. BOTTOU & K. Q. WEINBERGER [Hrsg.], *Advances in Neural Information Processing Systems: 26th Annual Conference on Neural Information Processing Systems (NIPS)* (Vol. 25, pp. 1097-1105). Lake Tahoe: Curran Associates, Inc.
- KRUSE, R., BORGELT, C., BRAUNE, C., KLAWONN, F., MOEWES, C. & STEINBRECHER, M. (2015): Allgemeine neuronale Netze. In: R. KRUSE, C. BORGELT, C. BRAUNE, F. KLAWONN, C. MOEWES & M. STEINBRECHER [Hrsg.], *Computational Intelligence* (pp. 33-42). Wiesbaden: Springer Vieweg.
- KUHLMANN, H. & HOLST, C. (2017): Flächenhafte Abtastung mit Laserscanning. In: W. SCHWARZ [Hrsg.], *Ingenieurgeodäsie* (pp. 167-212). Berlin: Springer.
- KUHLMANN, H. & KLINGBEIL, L. (2017): Mobile Multisensorsysteme. In: W. SCHWARZ, *Ingenieurgeodäsie* (pp. 93–129). Berlin: Springer.
- KUKKO, A. (2013): Mobile laser scanning - system development, performance and applications. Dissertation, Aalto University School of Engineering (Espoo, Finland), Department of Real Estate, Planning and Geoinformatics.

- KUKKO, A., JAAKKOLA, A., LEHTOMAKI, M., KAARTINEN, H. & CHEN, Y. (2009): Mobile mapping system and computing methods for modelling of road environment. Joint Urban Remote Sensing Event (Shanghai, 20-22 May 2009), pp. 1-6.
- KUKKO, A., KAARTINEN, H., HYYPPÄ, J. & CHEN, Y. W. (2012): Multiplatform mobile laser scanning: usability and performance. *Sensors*, 12(9), pp. 11712-11733.
- LAFFERTY, J. D., MCCALLUM, A. & PEREIRA, F. C. (2001): Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. *Proceedings of the Eighteenth International Conference on Machine Learning ICML 2001* (pp. 282-289). San Francisco, CA, USA: Morgan Kaufmann Publishers.
- LAKSHMINARAYANAN, B., PRITZEL, A. & BLUNDELL, C. (2017): Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles. In: I. GUYON, U. VON LUXBURG, S. BENGIO, H. WALLACH, R. FERGUS, S. VISHWANATHAN & R. GARNETT [Hrsg.], *Advances in Neural Information Processing Systems 30 (NIPS)*, (pp. 6403-6414). Long Beach, CA, USA.
- LAM, N. S.-N. (1983): Spatial Interpolation Methods: A Review. *The American Cartographer*, 10(2), pp. 129-150.
- LANDAU, H., VOLLATH, U. & CHEN, X. (2002): Virtual Reference Station Systems. *Journal of Global Positioning Systems*, 1(2), pp. 137-143.
- LANDESAMT FÜR DIGITALISIERUNG, BREITBAND UND VERMESSUNG (2022): Satellitenpositionierungsdienst der deutschen Landesvermessung. Abgerufen am 11. Januar 2023 von <https://www.ldbv.bayern.de/produkte/dienste/sapos/allgemeines.html>
- LANGE, H. (2022): *Handbuch für Beleuchtung*. Heidelberg: C. H. Beck.
- LARA-ALVAREZ, C., ROJAS, A. & BAYRO-CORROCHANO, E. (2014): Geometric Indexing for Recognition of Places. In: E. BAYRO-CORROCHANO, E. HANCOCK & R. EDWIN [Hrsg.], *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications. 19th Iberoamerican Congress, CIARP 2014, Puerto Vallarta, Mexico, November 2-5, 2014, Proceedings* (pp. 548-555). Springer International Publishing.
- LAURENT, J. (2010): LCMS – Laser Crack Measurement System. Abgerufen am 22. Februar 2023 von https://www.vtti.vt.edu/PDFs/PE_2010/Laurent%20-%203D%20laser%20profiling%20sensors%20for%20the%20automated%20measurement%20of%20road%20surface%20conditions.pdf

- LAURENT, J., HÉBERT, J. F., LEFEBVRE, D. & SAVARD, Y. (2012): Using 3D laser profiling sensors for the automated measurement of road surface conditions. RILEM International Conference on Cracking in Pavement, pp. 157-167.
- LEHMANN+PARTNER GMBH (2022): Technologie. Abgerufen am 22. Februar 2023 von <https://www.lehmann-partner.de/messsysteme-mobile-erfassung/>
- LEHMANN+PARTNER POLSKA (2022): pomiarydrogowe. Abgerufen am 22. Februar 2023 von <https://www.pomiarydrogowe.pl/>
- LEICA GEOSYSTEMS AG (2017): Leica GS18 T - Datenblatt. Abgerufen am 25. Januar 2023 von <https://leica-geosystems.com/-/media/files/leicageosystems/products/datasheets/leica%20gs18%20t%20ds%20866431%200720%20de%20lr.ashx?la=de-de&hash=3764B56781964884CBCD66037293E548>
- LEICA GEOSYSTEMS AG (2021): Leica Pegasus:Two Ultimate - Mobile reality capture. Abgerufen am 26. Oktober 2021 von https://leica-geosystems.com/de-de/products/mobile-sensor-platforms/capture-platforms/leica-pegasus_two-ultimate
- LI, J. & HEAP, A. (2008): A Review of Spatial Interpolation Methods for Environmental Scientists. Geoscience Australia, Record 2008/23.
- LI, J. & HEAP, A. (2014): Spatial interpolation methods applied in the environmental sciences: A review. Environmental Modelling & Software, 53, pp. 173-189.
- LI, J. & HEAP, A. D. (2011): Review of Comparative Studies of Spatial Interpolation Methods in Environmental Sciences: Performance and Impact Factors. Ecological Informatics, 6, 228-241.
- LI, S., DRAGICEVIC, S., CASTRO, F. A., SESTER, M., WINTER, S., COLTEKIN, A., PETTIT, C., JIANG, B., HAWORTH, J. & STEIN, A. (2016): Geospatial big data handling theory and methods: a review and research challenges. ISPRS Journal of Photogrammetry and Remote Sensing, 115, pp. 119-133.
- LI, W. & HSU, C.-Y. (2022): GeoAI for Large-Scale Image Analysis and Machine Vision: Recent Progress of Artificial Intelligence in Geography. ISPRS International Journal of Geo-Information, 11(7), pp. 1-44.

- LIN, T.-Y., MAIRE, M., BELONGIE, S., HAYS, J., PERONA, P., RAMANAN, D., DOLLÁR, P., ZITNICK, C. L., FLEET, D., PAJDLA, T., SCHIELE, B. & TUYTELAARS, T. (2014): Microsoft COCO: Common Objects in Context. Computer Vision – ECCV 2014, pp. 740-755.
- LIPKOWSKI, S. & METTENLEITER, M. (2019): Genauigkeit, Terrestrische Laserscanner - Im Fokus der Genauigkeit. In: J.-A. PAFFENHOLZ [Hrsg.], Terrestrisches Laserscanning 2019 (TLS 2019). Beiträge zum 184. DVW-Seminar am 2./3.12.2019 (Fulda) (pp. 75-87). Augsburg: Wißner-Verlag.
- LOFFE, S. & SZEGEDY, C. (2015): Batch normalization: accelerating deep network training by reducing internal covariate shift. International Conference on Machine Learning, pp. 448-456.
- LONG, J., SHELHAMER, E. & DARRELL, T. (2015): Fully Convolutional Networks for Semantic Segmentation. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), (pp. 3431-3440). Boston, MA, USA .
- LUHMANN, T. (2018): Nahbereichsphotogrammetrie – Grundlagen, Methoden, Beispiele. Berlin/Offenbach: Wichmann.
- LUNDBERG, T., ANDRÉN, P., WAHLMAN, T., ERIKSSON, O., SJÖGREN, L. & EKDAHL, P. (2018): New technology for road surface measurement: Transverse profile and rut depth. VTI rapport 961A. Linköping: Swedish National Road and Transport Research Institute (VTI).
- LUO, W., PINGEL, T., HEO, J., HOWARD, A. & JUNG, J. (2015): A progressive black top hat transformation algorithm for estimating valley volumes on Mars. Computers & Geosciences, 75, pp. 17-23.
- LUO, X., SCHAUFLE, S. & RICHTER, B. (2018): Leica GS18 T – World's Fastest GNSS RTK Rover. White Paper. Abgerufen am 23. Februar 2023 von <https://leica-geosystems.com/-/media/files/leicageosystems/products/white-papers/leica-gs18-t-a4-white%20paper-lr.ashx?la=en&hash=653AB15C78163860AFD0DCC306611A15>
- MACMILLAN, R. A. & SHARY, P. A. (2009): Landforms and Landform Elements in Geomorphometry. In: T. HENGL & H. I. REUTER [Hrsg.], Geomorphometry - Concepts, Software, Applications. Developments in Soil Science (Vol. 33, pp. 227-254). Amsterdam: Elsevier.

- MAJCHER, J., PLETS, R. & QUINN, R. (2020): Residual relief modelling: digital elevation enhancement for shipwreck site characterisation. *Archaeological and Anthropological Sciences*, 12(122), pp. 1-13.
- MANSFELD, W. (2010): *Satellitenortung und Navigation*. Wiesbaden: Vieweg Teubner.
- MASCHEK, U. (2018): *Sicherung des Schienenverkehrs - Grundlagen und Planung der Leit- und Sicherungstechnik*. Wiesbaden: Springer Vieweg.
- MAYER, G., VAN DER SLUIS, S. & STEINAUER, B. (2002): Griffigkeit von Fahrbahnoberflächen. Teil 2: Signifikante Griffigkeitskennwerte aus Verkehrssicherheitssicht. *Forschung Straßenbau und Straßenverkehrstechnik* (Heft 841).
- MIRALIAKBARI, A. (2017): *Road Condition Mapping by Integration of Laser Scanning, RGB Imaging and Spectrometry*. Dissertation, Technische Universität Dresden, Fakultät Umweltwissenschaften, Dresden.
- MITASOVA, H. & MITAS, L. (1993): Interpolation by regularized spline with tension: I. Theory and implementation. *Mathematical Geology*, 25(6), pp. 641-655.
- MOHAN, A. & POOBAL, S. (2018): Crack detection using image processing: A critical review and analysis. *Alexandria Engineering Journal*, 57(2), pp. 787-798.
- MOKARRAM, M. & SATHYAMOORTHY, D. (2016): Relationship between landform classification and vegetation (case study: southwest of Fars province, Iran). *Open Geosciences*, 8, pp. 302-309.
- MONETT, D., LEWIS, C. W. & THÓRISSON, K. R. (2020): Introduction to the JAGI Special Issue "On Defining Artificial Intelligence" - Commentaries and Author's Response. *Journal of Artificial General Intelligence*, 11(2), pp. 1-4.
- MONTEALEGRE, A. L., LAMELAS, M. T. & RIVA, J. D. (2015): Interpolation Routines Assessment in ALS-Derived Digital Elevation Models for Forestry Applications. *Remote Sensing*, 7(7), pp. 8631-8654.
- MONTENBRUCK, O., STEIGENBERGER, P. & HAUSCHILD, A. (2015): Broadcast versus precise ephemerides: a multi-GNSS perspective. *GPS Solutions*, 19, pp. 321-333.
- MUNAWAR, H. S., HAMMAD, A. W., HADDAD, A., SOARES, C. A. & WALLER, S. T. (2021): Image-Based Crack Detection Methods: A Review. *Infrastructures*, 6(115), pp. 1-20.

- MUSTAFA, G., SHAH, A. A. & ASIM, M. R. (2008): Constrained Shepard Method for Modeling and Visualization of Scattered Data. *Journal of WSCG*, 16(1-3), pp. 49-56.
- NABER, T., SESSELMANN, M., STRICKER, R., SCHELLER, S., STELLING, N., KLOSE, D. & GROßMANN, A. (2021): Künstliche Intelligenz in der Straßenraumerfassung. *Straßenverkehrstechnik*, 65(12), pp. 896-902.
- NASTRO, L. (2008): Position and orientation data requirements for precise autonomous vehicle navigation. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXVI, 1237-1242.
- NEUBECK, J. & WIESEBROCK, A. (2015): Überprüfung der Signalverarbeitungsverfahren nach dem Prinzip der Mehrfachabtastung (HRM). Bergisch-Gladbach: Bundesanstalt für Straßenwesen (BASt).
- NIEVELT LABOR GMBH (2022): Zustandserhebung / Messtechnik. Abgerufen am 10. August 2022 von <https://nievelt.at/de/services/measurement>
- NOACK, D. (2019): GIS-gestützte Analyse zum Wölbäckervorkommen in der Prignitz. *AGIT – Journal für Angewandte Geoinformatik*, 5, pp. 60-72.
- NOTEBAERT, B., VERSTRAETEN, G., GOVERS, G. & POESEN, J. (2009): Qualitative and quantitative applications of LiDAR imagery in fluvial geomorphology. *Earth Surface Processes and Landforms*, 34, pp. 217-231.
- OLIVEIRA, G. L. (2019): Encoder-Decoder Methods for Semantic Segmentation: Efficiency and Robustness Aspects. Dissertation, Albert-Ludwigs-Universität Freiburg, Technische Fakultät.
- OLIVEIRA, H. & CORREIA, P. L. (2014): CrackIT - An image processing toolbox for crack detection and characterization. *IEEE International Conference on Image Processing (ICIP)*, pp. 798-802.
- OLIVEIRA, H. & CORREIA, P. L. (2017): Road surface crack detection: Improved segmentation with pixel-based refinement. *25th European Signal Processing Conference (EUSIPCO)*, pp. 2026-2030.
- OPENSHAW, S. & OPENSHAW, C. (1997): *Artificial Intelligence in Geography*. New York: John Wiley & Sons.

- PAULY, L., HOGG, D., FUENTES, R. & PEEL, H. (2017): Deeper networks for pavement crack detection. Proceedings of the 34th International Symposium in Automation and Robotics in Construction (ISARC), pp. 479-485.
- PAVEMETRICS (2018): Laser Crack Measurement System. Abgerufen am 30. Januar 2023 von https://www.pavemetrics.com/wp-content/uploads/2018/12/LCMS-2_Flyer_LR.pdf
- PENG, L., CHAO, W., SHUANGMIAO, L. & BAOCAI, F. (2015): Research on crack detection method of airport runway based on twice-threshold segmentation. International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC), pp. 1716-1720.
- PETITCLERC, B., LAURENT, J. & HABEL, R. (2018): Optimising 3D surface characteristics data collection by re-using the data for project level road design. 8th Symposium on Pavement Surface Characteristics: SURF 2018 – Vehicle to Road Connectivity. Brisbane, Queensland.
- PETRIE, G. (2011): Airborne Topographic Laser Scanners: Current Developments in the Technology. GeoInformatics, 14(1), pp. 35-44.
- PIKE, R. J., EVANS, I. S. & HENGL, T. (2009): Geomorphometry: A Brief Guide. In: T. HENGL & H. I. REUTER [Hrsg.], Geomorphometry - Concepts, Software, Applications. Developments in Soil Science (Vol. 33, pp. 3-30). Amsterdam: Elsevier.
- PINGEL, J. T. & CLARKE, K. (2014): Perceptually Shaded Slope Maps for the Visualization of Digital Surface Models. Cartographica The International Journal for Geographic Information and Geovisualization, 49(4), pp. 225-240.
- POLSON, N. G. & SOKOLOV, V. O. (2017): Deep learning for short-term traffic flow prediction. Transportation Research Part C: Emerging Technologies, 79, pp. 1-17.
- POTÓ, V., CSEPINSZKY, A. & BARSÍ, Á. (2018): Representing road related laserscanned data in curved regular grid: a support to autonomous vehicles. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, XLII(2), pp. 917-921.
- POULINAKIS, K., DRIKAKIS, D., KOKKINAKIS, I. W. & SPOTTSWOOD, S. M. (2023): Machine-Learning Methods on Noisy and Sparse Data. Mathematics, 11(1), pp. 1-19.

- PRIMA, O. D., ECHIGO, A., YOKOYAMA, R. & YOSHIDA, T. (2006): Supervised landform classification of Northeast Honshu from DEM-derived thematic maps. *Geomorphology*, 78, pp. 373-386.
- PROJECT JUPYTER (2022): Project Jupyter - Interactive Computing. Abgerufen am 22. November 2022 von <https://github.com/jupyter/>
- PUENTE, I., GONZÁLEZ-JORGE, H., MARTÍNEZ-SÁNCHEZ, J. & ARIAS, P. (2013): Review of mobile mapping and surveying technologies. *Measurement*, 46(7), 2127-2145.
- QGIS.ORG (2022): QGIS Geographic Information System. QGIS Association. Abgerufen am 27. November 2022 von <http://www.qgis.org>
- QUINTANA, M., TORRES, J. & MENENDEZ, J. M. (2016): A Simplified Computer Vision System for Road Surface Inspection and Maintenance. *IEEE Transactions on Intelligent Transportation Systems*, 17(3), pp. 608-619.
- QUINTERO, C. A., ESCOBAR, I. P. & PONTE-NETO, C. F. (2012): Applications of Voronoi and Delaunay Diagrams in the solution of the geodetic boundary value problem. *Boletim de Ciências Geodésicas*, 18(3), pp. 378-396.
- RADKE, S. (2019): Verkehr in Zahlen 2019/2020. Abgerufen am 16. Juni 2020 von https://www.bmvi.de/SharedDocs/DE/Publikationen/G/verkehr-in-zahlen-2019-pdf.pdf?__blob=publicationFile
- RANYAL, E., SADHU, A. & JAIN, K. (2022): Road Condition Monitoring Using Smart Sensing and Artificial Intelligence: A Review. *Sensors*, 22, pp. 1-27.
- RASE, W.-D. (2021): Kartographische Oberflächen - Interpolation, Analyse, Visualisierung. Norderstedt: Books on Demand.
- RASMUSSEN, J. (1983): Skills, Rules and Knowledge: Signals, Signs and Symbols, and other distinctions in human performance models. *IEEE transactions on systems, man and cybernetics*, SMC-13(3), pp. 257-266.
- RAUH, J. & MÖSSNER-BEIGEL, M. (2008): Tyre simulation challenges. *Vehicle System Dynamics*, 46(1), 49-62.

- REITERER, A. (2013): Laser Scanning Technology for Pavement Monitoring – A New Approach. Presentation held at the first European road profile users group meeting, ERPUG.Copenhagen, Denmark 10-12 September 2013. Abgerufen am 24. August 2021 von www.erpug.org/media/files/forelasningar_2013/10_Alexander_Reiterer_Fraunhofer%20institute_Laser%20scanning%20technology.pdf
- REITERER, A., DAMBACHER, M., MAINDORFER, I., HÖFLER, H., EBERSBACH, D., FREY, C., SCHELLER, S. & KLOSE, D. (2013): Straßenzustandsüberwachung in Sub-Millimeter. Photogrammetrie – Laserscanning – Optische 3D-Messtechnik: Beiträge der Oldenburger 3D-Tage 2013, pp. 78-85.
- REITERER, A., WÄSCHLE, K., STÖRK, D., LEYDECKER, A. & GITZEN, N. (2020): Fully Automated Segmentation of 2D and 3D Mobile Mapping Data for Reliable Modeling of Surface Structures Using Deep Learning. Remote Sensing, 12, pp. 1-13.
- RENKA, R. J. (1988): Algorithm 660: QSHEP2D: Quadratic Shepard Method for Bivariate Interpolation of Scattered Data. ACM Transactions on Mathematical Software, 14(2), pp. 139-148.
- RIEGER, P., STUDNICKA, N., PFENNIGBAUER, M. & ZACH, G. (2010): Boresight alignment method for mobile laser scanning systems. Journal of applied Geodesy, 4(1), 13-21.
- RIEGL LASER MEASUREMENT SYSTEMS GMBH (2012): Compact Mobile Laser Scanning System - RIEGL VMX-250. Abgerufen am 22. Februar 2023 von www.riegl.com/uploads/tx_pxpriegldownloads/10_DataSheet_VMX-250_20-09-2012.pdf
- RODRIGUEZ, F., MAIRE, E., COURJAULT-RADÉ, P. & DARROZES, J. (2002): The Black Top Hat function applied to a DEM: A tool to estimate recent incision in a mountainous watershed. Geophysical Research Letters, 29(6), pp. 1-4.
- ROHLEDER, M., KUNZ, K., WASSER, B., PULLWITT, E., MÜLLER, I., RIPKE, O., ZÖLLER, M. & PÖPPEL-DECKER, M. (2010): Griffigkeitsprognose an offenporigen Asphalten (OPA). Bremerhaven: Wirtschaftsverlag NW, Verlag für neue Wissenschaft GmbH.
- RONNEBERGER, O., FISCHER, P. & BROX, T. (2015): U-net: Convolutional networks for biomedical image segmentation. In: N. NAVAB, J. HORNEGGER, W. WELLS & A. FRANGI [Hrsg.], Medical Image Computing and Computer-Assisted Intervention (MICCAI) (Vol. 9351, pp. 234-241). Cham: Springer.

- ROSAS-CHAVOYA, M., GALLARDO-SALAZAR, J. L., LÓPEZ-SERRANO, P. M., ALCÁNTARA-CONCEPTION, P. C. & LEÓN-MIRANDA, A. K. (2022): QGIS a constantly growing free and open-source geospatial software contributing to scientific development. *Cuadernos de Investigación Geográfica*, 48(1), 197-213.
- ROSCHER, R. & DREES, L. (2022): Einführung in die künstliche Intelligenz und das maschinelle Lernen. In: W. GRUNAU [Hrsg.], *Künstliche Intelligenz in Geodäsie und Geoinformatik* (pp. 15-23). Berlin: Wichmann.
- ROSSITER, D. G. (2016): Exercise: Thin plate spline interpolation. Abgerufen am 23. Februar 2023 von http://www.css.cornell.edu/faculty/dgr2/_static/files/R_PDF/exTPS.pdf
- ROUSSOS, G. & BAXTER, B. J. (2005): Rapid evaluation of radial basis functions. *Journal of Computational and Applied Mathematics*, 180(1), pp. 51-70.
- RUMELHART, D. E., DURBIN, R., GOLDEN, R. M. & CHAUVIN, Y. (1996): Backpropagation: The basic theory. In: P. SMOLENSKY, M. C. MOZER & D. E. RUMELHART [Hrsg.], *Mathematical Perspectives on Neural Networks* (pp. 533-566). Hillsdale (NJ): Erlbaum.
- SAGA GIS (2022): SAGA GIS Code. Abgerufen am 20. November 2022 von <https://sourceforge.net/p/saga-gis/code/ci/master/tree/>
- SAKOV, P. (2021): nn (Natural Neighbours interpolation). Abgerufen am 4. August 2021 von <https://github.com/sakov/nn-c/tree/master/nn>
- SALMAN, M., MATHAVAN, S., KAMAL, K. & RAHMAN, M. (2013): Pavement crack detection using the gabor filter. *Proceedings of 16th International IEEE Annual Conference on Intelligent Transportation Systems*, pp. 2039-2044.
- SAMBRIDGE, M., BRAUN, J. & MCQUEEN, H. (1995): Geophysical parametrization and interpolation of irregular data using natural neighbours. *Geophysical Journal International*, 122(3), pp. 837-857.
- SANTURKAR, S., TSIPRAS, D., ILYAS, A. & MADRY, A. (2018): How Does Batch Normalization Help Optimization? In: S. BENGIO, H. WALLACH, H. LAROCHELLE, K. GRAUMAN, N. CESA-BIANCHI & R. GARNETT [Hrsg.], *Advances in Neural Information Processing Systems* (Vol. 31, pp. 2483-2493). Red Hook: Curran Associates, Inc.

- SCALABEL (2022): ScaLabel. Abgerufen am 21. November 2022 von <https://github.com/scalabel/scalabel>
- SCHERZINGER, B. & HUTTON, J. (2023): Applanix IN-Fusion+: The Ultimate Solution for Enhanced Spatial Intelligence. White Paper. Abgerufen am 30. Januar 2023 von https://www.applanix.com/pdf/Applanix-In_Fusion-Plus-White-Paper.pdf
- SCHMID, M. (2019): Kommunalen Straßenbau und -unterhalt. In: B. K. PRÜFUNGSVERBAND [Hrsg.], Geschäftsbericht (pp. 54-116). München: Bayerischer Kommunalen Prüfungsverband (BKPV).
- SCHMUCK, A. (1987): Straßenerhaltung mit System – Grundlagen des Managements. Bonn: Kirschbaum Verlag.
- SCHREIBER, C. (2020): Schlaglöcher per App bekämpfen. ADAC Motorwelt (Regionalmagazin Baden-Württemberg), pp. 16-17.
- SESSELMANN, M. (2018): Automatische Detektion von Substanzmerkmalen auf 3D-Fahrbahnoberflächen. AGIT – Journal für angewandte Geoinformatik, 4, pp. 65-74.
- SESSELMANN, M., NABER, T., SCHELLER, S., STELLING, N. & GROSSMANN, A. (2022): Mobiles Laserscanning und Deep Learning für die automatische Kartierung des Straßenzustandes im kommunalen Straßennetz. Photogrammetrie – Laserscanning – Optische 3D-Messtechnik: Beiträge der Oldenburger 3D-Tage 2022, pp. 214-227.
- SESSELMANN, M., NABER, T., STRICKER, R. & SCHELLER, S. (2023): Mobiles Laserscanning und Deep Learning zur automatischen Erfassung des Oberflächenzustandes kommunaler Straßen. avn - Allgemeine Vermessungsnachrichten, 130, pp. 21-34.
- SESSELMANN, M., STRICKER, R. & EISENBACH, M. (2019): Einsatz von Deep Learning zur automatischen Detektion und Klassifikation von Fahrbahnschäden aus mobilen LiDAR-Daten. AGIT-Journal für Angewandte Geoinformatik, 5, pp. 100-114.
- SESSELMANN, M., STRICKER, R., NABER, T. & SCHELLER, S. (2020): Automatische Detektion und objektscharfe Georeferenzierung von Fahrbahnschäden aus Bilddaten eines Mobile-Mapping-Systems mithilfe von Deep Learning. gis.Science, 1, pp. 18-30.
- SHEPARD, D. (1968): A two-dimensional interpolation function for irregularly-spaced data. Proceedings of the 23rd ACM National Conference, pp. 517-524.

- SIBSON, R. (1981): A Brief Description of Natural Neighbor Interpolation. In: V. BARNETT [Hrsg.], *Interpreting Multivariate Data* (pp. 21-36). New York: John Wiley & Sons.
- SIMONYAN, K. & ZISSERMAN, A. (2015): Very Deep Convolutional Networks for Large-Scale Image Recognition. *International Conference on Learning Representations*.
- SITECO (2022): Siteco Informatica. Abgerufen am 10. Oktober 2022 von <https://www.sitecoinf.it/en/>
- SKAKUJ, M., BALCK, H., BECKEDAH, H. J., SCHRÖDTER, T., KOPPERS, S., RAMADAN, E., BRZEZINSKA, P., SMET, W. & WOPENBERG, M. (2019): *Temperatureinfluss auf die Sichtbarkeit von Substanzmerkmalen (Oberfläche)*. Bremen: Fachverlag NW in der Carl Ed. Schünemann KG.
- SKALECKI, P., SESSELMANN, M., RECHKEMMER, S., BRITZ, T., GROßMANN, A., GARRECHT, H. & SAWODNY, O. (2021): Process Evaluation for Smart Concrete Road Construction: Road Surface and Thickness Evaluation Using High-Speed LiDAR Technology. *Automation*, 2(1), 31-47.
- SMIRNOV, E., TIMOSHENKO, D. & ANDRIANOV, S. (2014): Comparison of Regularization Methods for ImageNet Classification with Deep Convolutional Neural Networks. 2nd AASRI Conference on Computational Intelligence and Bioinformatics, 6, pp. 89-94.
- SMITH, T. R. (1984): Artificial intelligence and its applicability to geographical problem solving. *The Professional Geographer*, 36(2), pp. 147-158.
- SOILLE, P. (2004): *Morphological Image Analysis - Principles and Applications*. New York: Springer-Verlag.
- SRIVASTAVA, N., HINTON, G., KRIZHEVSKY, A., SUTSKEVER, I. & SALAKHUTDINOV, R. (2014): Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, 15, pp. 1929-1958.
- STASCHKIEWICZ, M. & RADENBERG, M. (2022): Prüfverfahren zur Beurteilung des Kälteverhaltens von Bitumen. *Straße und Autobahn*, 73(4), pp. 286-292.
- STEINAUER, B., UECKERMANN, A. & MAERSCHALK, G. (2006): *Analyse vorliegender messtechnischer Zustandsdaten und Erweiterung der Bewertungsparameter fuer Innerortsstrassen*. Bremerhaven: Wirtschaftsverlag NW, Verlag für neue Wissenschaft GmbH.

- STEINIGER, S. & BOCHER, E. (2009): An overview on current free and open source desktop GIS developments. *International Journal of Geographical Information Science*, 23(10), 1345-1370.
- STILL, P. B. & JORDAN, P. G. (1980): Evaluation of the TRRL High-Speed Profilometer (Report LR922). Crowthorne: Transportation and Road Research Laboratory.
- STÖCKERT, U., SCHMERBECK, R. & LAU, F. (2019): Erhaltungsmanagement der Bundesfernstraßen. *Straße und Autobahn*, 5, pp. 399-409.
- STRICKER, R., AGANIAN, D., LEMBACH, J., STEPHAN, B., SESSELMANN, M., NABER, T., SPIELHOFER, R., HAHN, M. & HAUTZ, A. (2021a): Automatisierte Schadstellenerkennung für unterschiedliche Fahrbahnbeläge mittels Deep-Learning-Techniken (FFG Ergebnisbericht "ASFaLT"). Abgerufen am 10. August 2022 von https://projekte.ffg.at/anhang/61b8b0e7527cc_ASFaLT_Ergebnisbericht.pdf
- STRICKER, R., AGANIAN, D., SESSELMANN, M., SEICHTER, D., ENGELHARDT, M., SPIELHOFER, R., HAHN, M., HAUTZ, A., DEBES, K. & GROSS, H.-M. (2021b): Road Surface Segmentation - Pixel Perfect Distress and Object Detection for Road Assessment. *IEEE International Conference on Automation Science and Engineering (CASE)*, pp. 1789-1796.
- STRICKER, R., EISENBACH, E., SESSELMANN, M., DEBES, K. & GROSS, H.-M. (2019): Improving Visual Road Condition Assessment by Extensive Experiments on the Extended GAPS Dataset. *IEEE International Joint Conference on Neural Networks (IJCNN)*, 1-8.
- SU, J., MIYAZAKI, R., TAMAKI, T. & KANEDA, K. (2019): High-Resolution Representation for Mobile Mapping Data in Curved Regular Grid Model. *Sensors*, 19(24), pp. 1-17.
- SZYPULA, B. (2016): Geomorphometric comparison of DEMs built by different interpolation methods. *Landform Analysis*, 32, pp. 45-58.
- TAGIL, S. & JENNESS, J. S. (2008): GIS-Based Automated Landform Classification and Topographic, Landcover and Geologic Attributes of Landforms Around the Yazoren Polje, Turkey. *Journal of Applied Sciences*, 8(6), pp. 910-921.
- TELEDYNE OPTTECH INCORPORATED (2021): Lynx HS600 Mobile Survey System. Abgerufen am 26. Oktober 2021 von <https://www.teledyneoptech.com/en/products/mobile-survey/lynx-hs600/>

- TEMME, A. J., HEUVELINK, G. B., SCHOORL, J. M. & CLAESSENS, L. (2009): Geostatistical Simulation and Error Propagation in Geomorphometry. In: T. HENGL & H. I. REUTER [Hrsg.], Geomorphometry - Concepts, Software, Applications. Developments in Soil Science (Vol. 33, pp. 121-140). Amsterdam: Elsevier.
- TENSORBOARD (2022): TensorBoard: TensorFlow's visualization toolkit. Abgerufen am 22. November 2022 von https://www.tensorflow.org/tensorboard/get_started
- TENSORFLOW (2022): TensorFlow. Abgerufen am 22. November 2022 von <https://github.com/tensorflow/tensorflow/graphs/contributors>
- THACKER, W. I., ZHANG, J., WATSON, L. T., BIRCH, J. B., IYER, M. A. & BERRY, M. W. (2010): Algorithm 905: SHEPPACK: Modified Shepard Algorithm for Interpolation of Scattered Multivariate Data. ACM Transactions on Mathematical Software, 37(3), pp. 1-20.
- THRUN, S., BURGARD, W. & FOX, D. (2005): Probabilistic Robotics. Cambridge: MIT Press.
- TOBLER, W. (1970): A Computer Movie Simulating Urban Growth in the Detroit Region. Economic Geography, 46, pp. 234-240.
- TOPCON CORPORATION (2021): IP-S3 HD1 - Compact, High-Density 3D Mobile Mapping System. Abgerufen am 26. Oktober 2021 von https://www.topconpositioning.com/sites/default/files/product_files/ip-s3_broch_7010_2175_revc_sm_0.pdf
- TRIMBLE INC (2021): Trimble MX9. Abgerufen am 06. Oktober 2021 von <https://de.geospatial.trimble.com/products-and-solutions/trimble-mx9>
- TSAI, V. J. (1993): Delaunay triangulations in TIN creation: an overview and a linear-time algorithm. International Journal of Geographical Information Systems, 7(6), pp. 501-524.
- TSAI, Y. J. (2018): Pavement Maintenance Programming Using 3D Laser Technology. Abgerufen am 22. Februar 2023 von <https://onlinepubs.trb.org/onlinepubs/webinars/180326.pdf>
- TSIDAEV, A. (2016): Parallel algorithm for natural neighbor interpolation. CEUR Workshop Proceedings, 1729, 78-83.

- UECKERMANN, A. & OESER, M. (2015): Approaches for a 3D assessment of pavement evenness data based on 3D vehicle models. *Journal of Traffic and Transportation Engineering (English Edition)*, 2(2), pp. 68-80.
- VAN DEN EECKHAUT, M., POESEN, J., VERSTRAETEN, G., VANACKER, V., MOEYERSONS, J., NYSSSEN, J. & VAN BEEK, L. P. (2005): The effectiveness of hillshade maps and expert knowledge in mapping old deep-seated landslides. *Geomorphology*, 67(3-4), pp. 351-363.
- VAN KREVELD, M. (1997): Digital elevation models and TIN algorithms. In: M. VAN KREVELD, J. NIEVERGELT, T. ROOS & P. WIDMAYER [Hrsg.], *Algorithmic Foundations of Geographic Information Systems*. CISM School 1996. Lecture Notes in Computer Science (Vol. 1340, pp. 37-78). Berlin, Heidelberg: Springer.
- VARADHARAJAN, S., JOSE, S., SHARMA, K., WANDER, L. & MERTZ, C. (2014): Vision for road inspection. *IEEE Winter Conference on Applications of Computer Vision*, pp. 115-122.
- VOPHAM, T., HART, J. E., LADEN, F. & CHIANG, Y.-Y. (2018): Emerging trends in geospatial artificial intelligence (geoAI): potential applications for environmental epidemiology. *Environmental Health*, 17, pp. 1-6.
- WACKER, T. (2021): Präzisierung der Zustandsbewertung als Beitrag zur Verbesserung der Erhaltungsstrategie für Bundesautobahnen (Dissertation). Siegen: Naturwissenschaftlich-Technische Fakultät der Universität Siegen.
- WAGNER, W., ULLRICH, A. & BRIESE, C. (2003): Der Laserstrahl und seine Interaktion mit der Erdoberfläche. *Österreichische Zeitschrift für Vermessung und Geoinformation*, 91(4), 223-235.
- WANG, B., ZHAO, W., GAO, P., ZHANG, Y. & WANG, Z. (2018): Crack Damage Detection Method via Multiple Visual Features and Efficient Multi-Task Learning Model. *Sensors*, 6, pp. 1-18.
- WANG, C., WEN, C., DAI, Y., YU, S. & LIU, M. (2020): Urban 3D modeling using mobile laser scanning: a review. *Virtual Reality & Intelligent Hardware*, 2(3), pp. 175-212.
- WANG, C.-F. (2018): A Basic Introduction to Separable Convolutions. Abgerufen am 27. Januar 2023 von <https://towardsdatascience.com/a-basic-introduction-to-separable-convolutions-b99ec3102728>

- WANG, P. (2008): What Do You Mean by “AI”? In: P. WANG, B. GOERTZEL & S. FRANKLIN [Hrsg.], Artificial General Intelligence 2008. Proceedings of the First AGI Conference, Frontiers in Artificial Intelligence and Applications (Vol. 171, pp. 362-373). Amsterdam: IOS Press.
- WANG, X. & HU, Z. (2017): Grid-based pavement crack analysis using deep learning. 4th International Conference on Transportation Information and Safety (ICTIS), pp. 917-924.
- WANG, Y., CHEN, Q., ZHU, Q., LIU, L., LI, C. & ZHENG, D. (2019): A Survey of Mobile Laser Scanning Applications and Key Techniques over Urban Areas. Remote Sensing, 11(13), 2072-4292.
- WANG, Z., WANG, E. & ZHU, Y. (2020): Image segmentation evaluation: a survey of methods. Artificial Intelligence Review, 53, pp. 5637-5674.
- WATSON, D. F. (1995): Contouring: A Guide to the Analysis and Display of Spatial Data. Oxford: Pergamon Press.
- WATSON, D. F. & PHILIP, G. M. (1985): A Refinement of Inverse Distance Weighted Interpolation. Geoprocessing, 2, pp. 315-327.
- WEBER, G., DETTMERING, D. & GEBHARD, H. (2005): Networked Transport of RTCM via Internet Protocol (NTRIP). In: F. SANSÒ [Hrsg.], A Window on the Future of Geodesy. (pp. 60-64). Berlin: Springer.
- WEHR, A. & LOHR, U. (1999): Airborne Laser Scanning – An Introduction and Overview. ISPRS Journal of Photogrammetry and Remote Sensing, 54(2-3), pp. 68-82.
- WEISS, A. (2001): Topographic position and landforms analysis. San Diego: ESRI User Conference.
- WEISS, K., KHOSHGOFTAAR, T. M. & WANG, D. (2016): A survey of transfer learning. Journal of Big Data, 3, pp. 1-40.
- WIELAND, M. & SESSELMANN, M. (2018): Plattenspezifische 3D-Oberflächenanalyse im Kontext mit der rechnerischen Dimensionierung und Restsubstanzbewertung von Betonfahrbahndecken. Straße und Autobahn, 6, pp. 447-458.

- WILD-PFEIFFER, F. (2015): Das Potential von MEMS-Inertialsensoren zur Anwendung in der Geodäsie und Navigation. Reihe B - Angewandte Geodäsie - Heft Nr. 319. München: Verlag der Bayerischen Akademie der Wissenschaften.
- WILSON, J. P. & GALLANT, J. C. (2000): Terrain Analysis: Principles and Applications. New York: Wiley.
- WISE, S. (2011): Cross-validation as a means of investigating DEM interpolation error. Computers and Geosciences, 37(8), pp. 978-991.
- WISE, S. M. (1998): The effect of GIS interpolation errors on the use of digital elevation models in geomorphology. In: S. N. LANE, K. S. RICHARDS & J. H. CHANDLER [Hrsg.], Landform Monitoring, Modelling and Analysis (pp. 139–164). Hoboken: Wiley.
- WITTE, B., SPARLA, P. & BLANKENBACH, J. (2020): Vermessungskunde für das Bauwesen mit Grundlagen des Building Information Modeling (BIM) und der Statistik. Berlin: Wichmann.
- WOOD, J. (2008): Digital Elevation Model (DEM). In: K. K. KEMP & K. K. KEMP [Hrsg.], Encyclopedia of Geographic Information Science (pp. 107–110). Thousand Oaks: SAGE Publications, Inc. .
- WOOD, S. N. (2003): Thin plate regression splines. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 65(1), pp. 95-114.
- WOODMAN, O. (2007): An introduction to inertial navigation. Cambridge: University of Cambridge Computer Laboratory.
- WROBEL, S., JOACHIMS, T. & MORIK, K. (2012): Maschinelles Lernen und Data Mining. In: G. GÖRZ, J. SCHNEEBERGER & U. SCHMID [Hrsg.], Handbuch der Künstlichen Intelligenz (pp. 405-471). München: Oldenbourg.
- WU, Q., DENG, C. & CHEN, Z. (2016): Automated delineation of karst sinkholes from LiDAR-derived digital elevation models. Geomorphology, 266, pp. 1-10.
- XIE, Y., CAI, J., BHOJWANI, R., SHEKHAR, S. & KNIGHT, J. (2020): A locally-constrained yolo framework for detecting small and densely-distributed building footprints. International Journal of Geographical Information Science, 34(4), pp. 777-801.
- YAMASHITA, R., NISHIO, M., DO, R. K. & TOGASHI, K. (2018): Convolutional neural networks: an overview and application in radiology. Insights Imaging, 9, pp. 611-629.

- YOKOYAMA, R., SHIRASAWA, M. & PIKE, R. J. (2002): Visualizing topography by openness: a new application of image processing to digital elevation models. *Photogrammetric engineering and remote sensing*, 68(3), pp. 257-266.
- YUTZLER, J. (2020): #GeoPackageDay 2020 - what is GeoPackage? Abgerufen am 27. November 2022 von <https://www.ogc.org/blog/3148>
- Z+F (2021): Z+F PROFILER 9020. Abgerufen am 17. Januar 2023 von <https://www.positionpartners.com.au/wp-content/uploads/2021/07/Datasheet-PROFILER-9020-en.pdf>
- ZEILER, M. D. & FERGUS, R. (2014): Visualizing and Understanding Convolutional Networks. In: D. FLEET, T. PAJDLA, B. SCHIELE & T. TUYTELAARS [Hrsg.], *Computer Vision – ECCV 2014* (pp. 818-833). Cham: Springer.
- ZHANG, L., YANG, F., ZHANG, Y. D. & ZHU, Y. J. (2016): Road crack detection using deep convolutional neural network. *IEEE International Conference on Image Processing (ICIP)*, pp. 3708-3712.
- ZHANG, Q., NIU, X. & SHI, C. (2020): Impact Assessment of Various IMU Error Sources on the Relative Accuracy of the GNSS/INS Systems. *IEEE Sensors Journal*, 20(9), pp. 5026-5038.
- ZHANG, Y. & DAVISON, B. D. (2020): Impact of ImageNet Model Selection on Domain Adaptation. *IEEE Winter Applications of Computer Vision Workshops (WACVW)*, pp. 173-182.
- ZHAO, H., SHI, J., QI, X., WANG, X. & JIA, J. (2017): Pyramid Scene Parsing Network. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6230-6239.
- ZHOU, W., HOU, J., LIU, L., SUN, T. & LIU, J. (2017): Design and Simulation of the Integrated Navigation System based on Extended Kalman Filter. *Open Physics*, 15(1), pp. 182-187.
- ZHU, D., CHENG, X., ZHANG, F., YAO, X., GAO, Y. & LIU, Y. (2020): Spatial interpolation using conditional generative adversarial neural networks. *International Journal of Geographical Information Science*, pp. 735-758.

Anhang I

Kalibrierschein für das Stahllineal, der zur Kalibrierung des Laserscanners genutzt wurde.



akkreditiert durch die / accredited by the

Deutsche Akkreditierungsstelle GmbH



als Kalibrierlaboratorium im / as calibration laboratory in the

Deutschen Kalibrierdienst



Kalibrierschein
Calibration certificate

Kalibrierzeichen
Calibration mark

10138
D-K- 17059-02-00
2018-06

Gegenstand Object	Prüflineal aus Stahl
Hersteller Manufacturer	--
Typ Type	5000 mm
Fabrikat/Serien-Nr. Serial number	173122-1
Auftraggeber Customer	Lehmann & Partner GmbH Sachsenallee 24 01723 Wilsdruff
Auftragsnummer Order No.	AN180345
Anzahl der Seiten des Kalibrierscheines Number of pages of the certificate	4
Datum der Kalibrierung Date of calibration	13.06.2018

Dieser Kalibrierschein dokumentiert die Rückführung auf nationale Normale zur Darstellung der Einheiten in Übereinstimmung mit dem internationalen Einheitensystem (SI).

Die DAkKS ist Unterzeichner der multilateralen Übereinkommen der European co-operation for Accreditation (EA) und der International Laboratory Accreditation Cooperation (ILAC) zur gegenseitigen Anerkennung der Kalibrierscheine. Für die Einhaltung einer angemessenen Frist zur Wiederholung der Kalibrierung ist der Benutzer verantwortlich.

This calibration certificate documents the traceability to national standards, which realize the units of measurement according to the International System of Units (SI).

The DAkKS is signatory to the multilateral agreements of the European co-operation for Accreditation (EA) and of the International Laboratory Accreditation Cooperation (ILAC) for the mutual recognition of calibration certificates.

The user is obliged to have the object recalibrated at appropriate intervals.

Dieser Kalibrierschein darf nur vollständig und unverändert weiterverbreitet werden. Auszüge oder Änderungen bedürfen der Genehmigung sowohl der Deutschen Akkreditierungsstelle GmbH als auch des ausstellenden Kalibrierlaboratoriums. Kalibrierscheine ohne Unterschrift haben keine Gültigkeit.

This calibration certificate may not be reproduced other than in full except with the permission of both the Deutsche Akkreditierungsstelle GmbH and the issuing laboratory. Calibration certificates without signature are not valid.

Datum Date	Leiter des Kalibrierlaboratoriums Head of the calibration laboratory	Bearbeiter Person in charge
15.06.2018	 Dr.-Ing. Will	 Lajdych

Seite
Page 2

10138
D-K- 17059-02-00
2018-06

1. Kalibriergegenstand

Prüflineal aus Stahl

Identnummer: 173122-1
Abmessung: 5000 mm
Typ / Standard: Flachlineale nach DIN 874-1:November 2003

2. Kalibrierverfahren

Die Kalibrierung erfolgt auf der Grundlage der L&W Richtlinie LW-K-Geradheit-1411-18.
Es wird ein Differenzmessverfahren mit zwei elektronischen Neigungsmessgeräten (Wyler BlueLevel) angewendet.

Der Anschluss an das nationale Normal erfolgte durch folgende Messgeräte und Normale:

Wyler Messsystem BlueLEVEL Funk
Serial-Nr.: 00158 / 00159 / 00717, Kalibrierschein Nr. SCS-044-1709134

3. Ort der Kalibrierung

Werkhalle der Lehmann & Partner GmbH;

4. Messbedingungen

Die Basislänge des Messgerätes beträgt 150 mm. Es wird eine Schrittweite von 125 mm gewählt. Vor der Kalibrierung werden folgende Parameter sichergestellt:

Die Führungsschiene muss sich im unbelasteten Zustand befinden und auf den vorgesehenen Auflagepunkten gelagert sein.

Der Temperaturunterschied im Prüfling darf 2 K nicht überschreiten.

Die Führungsschiene wird in 40 Abschnitte eingeteilt. Diese Einteilung erfolgt so, dass sich die einzelnen Messstrecken überlappen.

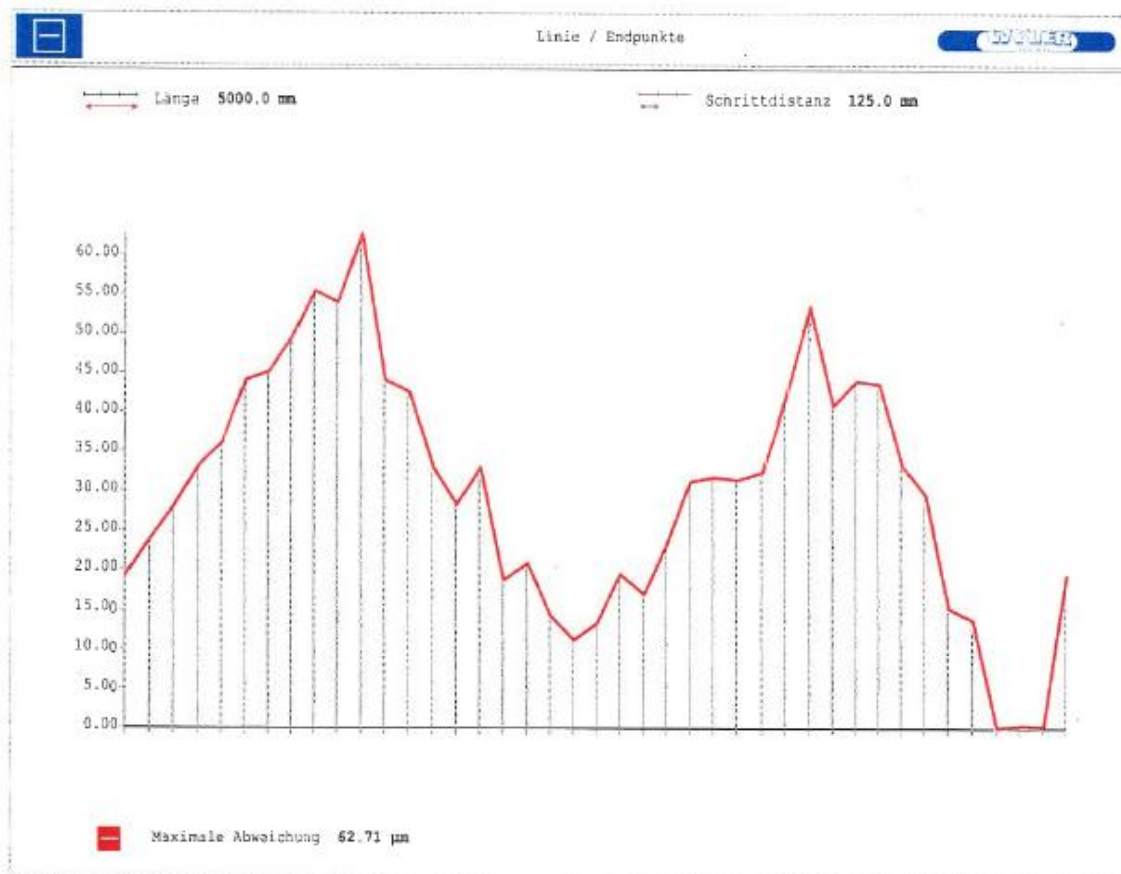
5. Umgebungsbedingungen

Temperatur und relative Luftfeuchte während der Kalibrierung liegen im Laboratorium auf Datenträger vor und können bei Bedarf eingesehen werden.

Die Temperatur betrug bei der Messung $(22,7 \pm 0,5) ^\circ\text{C}$.

6. Messergebnisse

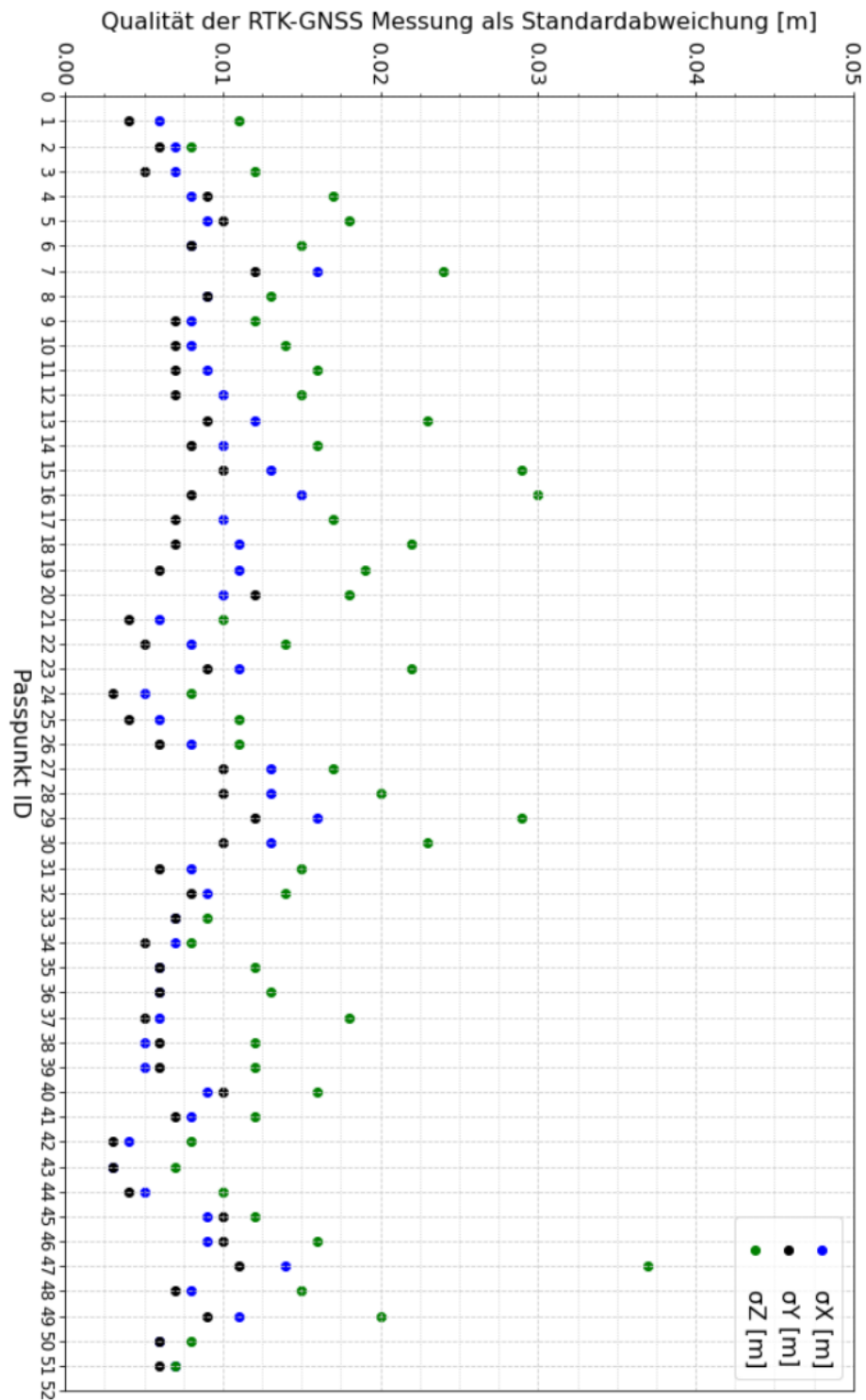
	Auftraggeber	Lehmann & Partner GmbH
	Prüfgegenstand	Prüflineal aus Stahl; Länge 5000mm
	Typ	DIN 874-1:2008
	Ident.-Nr.	173122-1
	Datum	13.06.2018
	Temperatur bei Beginn der Prüfung	22,7 °C
	Temperatur bei Ende der Prüfung	22,7 °C



Detail: C:\Users\Nutzer\Desktop\Arbeit\Wylar-Daten\173122-1_A_wy1. Datum: 13.06.2018 11:55

Anhang II

Qualitätsmaß, welches vom GNSS-RTK Rover Leica GS18 T mit dem Bedienelement Leica CS 20 für jeden gemessenen Passpunkt angegeben wird.



Anhang III

Zertifikat des LCMS-Prüfkörpers, der zur Bestimmung der Wiederholgenauigkeit verwendet wurde.



Traveler Sheet / Certificate of compliance
LCMS - Calibration object

INO-7816-FOR-005
 Edition 00 Revision 02

Done by :	Dominique St-Gelais	Date :	2019-10-29	Serial # :	170
------------------	---------------------	---------------	------------	-------------------	-----

Step #	Description	Measure (inch)	Tolerance (inch)	Pass / Fail	Comments
1	Assemble calibration object as per drawing # 7752-MAS002	n/a	n/a	Pass	
2	Assemble and screw the 2 plates 7752-MPC071 to have an angle of 90°	n/a	n/a	Pass	
3	Measure the height of the assembly of the 2 plates 7752-MPC071 - near the level	3,96	Min: 3,877 Max: 3,997	Pass	
4	Measure the height of the assembly of the 2 plates 7752-MPC071 - near the leg	3,95	Min: 3,877 Max: 3,997	Pass	
5	Measure the width of the assembly of the 2 plates 7752-MPC071 - near the level	7,883	Min: 7,808 Max: 7,928	Pass	
6	Measure the width of the assembly of the 2 plates 7752-MPC071 - near the leg	7,87	Min: 7,808 Max: 7,928	Pass	
7	Assemble the 2 tubes 7752-MPC070 and check the parallelism with plates 7752-MPC071	n/a	n/a	n/a	

ECME	Description: Inventory # : 1148279 Calibration due date : 2020-01-30	Micrometer
-------------	---	-------------------

Assembly

Done by : Dominique St-Gelais

Date : 2019-10-29

Verification (data results and torques)

Done by : David Hayes

Date : 2019-10-30

Anhang IV

Zertifikat der Stahleindmaße, die zur Bestimmung der Spaltbreiten des Prüfkörpers genutzt wurden.

Certificate of Calibration

Set No: 87 Serial No: LS6014 Material: steel
 Grade : 1 Standard: ISO3650-98 Date: 2016-01-10
 Temperature: (20±1.0) °C Coefficient of Thermal Expansion: $(11.5 \pm 1) \times 10^{-6} / ^\circ\text{C}$

Page: 1

Nominal Size mm	Deviation μm	fo μm	+	fu μm	=	fs μm	Dev. max μm	Grade	Serial number
1.001	0.06	0.00		0.02		0.02	0.06	0	154616
1.002	-0.05	0.00		0.05		0.05	0.10	0	134331
1.003	0.00	0.00		0.04		0.04	0.04	0	137909
1.004	0.08	0.00		0.04		0.04	0.08	0	146936
1.005	0.15	0.00		0.07		0.07	0.15	1	154528
1.006	0.15	0.00		0.05		0.05	0.15	1	151329
1.007	0.05	0.00		0.04		0.04	0.05	0	153233
1.008	0.00	0.00		0.07		0.07	0.07	0	137461
1.009	0.00	0.00		0.05		0.05	0.05	0	130294
1.01	0.02	0.00		0.05		0.05	0.03	0	142011
1.02	0.10	0.00		0.03		0.03	0.10	0	151206
1.03	0.10	0.00		0.04		0.04	0.10	0	142652
1.04	0.04	0.00		0.04		0.04	0.04	0	141740
1.05	-0.05	0.00		0.03		0.03	0.08	0	149825
1.06	0.12	0.00		0.04		0.04	0.12	0	147267
1.07	0.14	0.00		0.07		0.07	0.14	1	159318
1.08	-0.03	0.00		0.04		0.04	0.07	0	148474
1.09	0.06	0.00		0.02		0.02	0.06	0	156432
1.1	0.00	0.00		0.07		0.07	0.07	0	142500
1.11	0.08	0.00		0.04		0.04	0.08	0	149173
1.12	-0.02	0.00		0.04		0.04	0.06	0	145042
1.13	-0.05	0.00		0.03		0.03	0.08	0	157958
1.14	-0.05	0.00		0.05		0.05	0.10	0	147677
1.15	0.00	0.00		0.06		0.06	0.06	0	131785
1.16	0.00	0.00		0.05		0.05	0.05	0	155496
1.17	0.02	0.00		0.06		0.06	0.04	0	153931
1.18	0.04	0.00		0.04		0.04	0.04	0	152097
1.19	0.05	0.00		0.02		0.02	0.05	0	150826
1.2	0.12	0.00		0.03		0.03	0.12	0	150783
1.21	0.12	0.00		0.07		0.07	0.12	0	140961
1.22	0.16	0.00		0.06		0.06	0.16	1	152228
1.23	-0.08	0.00		0.05		0.05	0.13	1	151818
1.24	0.15	0.00		0.07		0.07	0.15	1	147655
1.25	0.02	0.00		0.05		0.05	0.03	0	152717
1.26	-0.08	0.00		0.06		0.06	0.14	1	122938
1.27	-0.02	0.00		0.03		0.03	0.05	0	157043
1.28	0.12	0.00		0.07		0.07	0.12	0	151399
1.29	-0.02	0.00		0.05		0.05	0.07	0	156913
1.3	0.08	0.00		0.08		0.08	0.08	0	130033
1.31	0.00	0.00		0.04		0.04	0.04	0	142091
1.32	-0.01	0.00		0.03		0.03	0.04	0	139616
1.33	0.15	0.00		0.05		0.05	0.15	1	152324
1.34	0.06	0.00		0.10		0.10	0.06	0	120147
1.35	0.15	0.00		0.10		0.10	0.15	1	132218
1.36	-0.08	0.00		0.05		0.05	0.13	1	148238
1.37	-0.08	0.00		0.04		0.04	0.12	0	147004
1.38	-0.04	0.00		0.06		0.06	0.10	0	158559
1.39	-0.08	0.00		0.03		0.03	0.11	0	156349
1.4	0.00	0.00		0.03		0.03	0.03	0	152384

Serial No: LS6014

Page: 2

Nominal Size mm	Deviation μm	fo μm	+	fu μm	=	fs μm	Dev. max μm	Grade	Serial number
1.41	-0.06	0.00		0.04		0.04	0.10	0	136345
1.42	0.15	0.00		0.10		0.10	0.15	1	157568
1.43	0.06	0.00		0.02		0.02	0.06	0	144455
1.44	-0.04	0.00		0.06		0.06	0.10	0	132281
1.45	0.10	0.00		0.10		0.10	0.10	0	135465
1.46	0.03	0.00		0.05		0.05	0.03	0	150422
1.47	0.10	0.00		0.05		0.05	0.10	0	151297
1.48	0.12	0.00		0.05		0.05	0.12	0	150726
1.49	0.06	0.00		0.04		0.04	0.06	0	137132
0.5	0.00	0.00		0.10		0.10	0.10	0	154171
1.0	0.11	0.00		0.03		0.03	0.11	0	156836
1.5	-0.08	0.00		0.05		0.05	0.13	1	153124
2.0	-0.06	0.00		0.05		0.05	0.11	0	157262
2.5	-0.05	0.00		0.04		0.04	0.09	0	152845
3.0	0.10	0.00		0.07		0.07	0.10	0	155927
3.5	0.15	0.00		0.08		0.08	0.15	1	154536
4.0	0.10	0.00		0.04		0.04	0.10	0	146449
4.5	-0.04	0.00		0.06		0.06	0.10	0	157724
5.0	0.07	0.00		0.02		0.02	0.07	0	150396
5.5	-0.05	0.00		0.03		0.03	0.08	0	156914
6.0	0.16	0.00		0.06		0.06	0.16	1	151004
6.5	0.12	0.00		0.06		0.06	0.12	0	150867
7.0	0.10	0.00		0.07		0.07	0.10	0	152568
7.5	-0.04	0.00		0.06		0.06	0.10	0	158144
8.0	-0.04	0.00		0.02		0.02	0.06	0	152382
8.5	0.15	0.00		0.07		0.07	0.15	1	141429
9.0	-0.02	0.00		0.03		0.03	0.05	0	159549
9.5	-0.05	0.00		0.05		0.05	0.10	0	150518
10	0.10	0.00		0.04		0.04	0.10	0	152934
20	0.20	0.00		0.03		0.03	0.20	1	155131
30	0.10	0.00		0.10		0.10	0.10	0	154701
40	0.15	0.00		0.05		0.05	0.15	0	158952
50	0.16	0.06		0.00		0.06	0.22	1	151569
60	0.05	0.05		0.05		0.10	0.10	0	158491
70	-0.10	0.05		0.05		0.10	0.15	0	150750
80	-0.20	0.10		0.00		0.10	0.20	0	151534
90	0.02	0.05		0.00		0.05	0.07	0	150204
100	-0.10	0.10		0.00		0.10	0.10	0	157570

INSPECTOR:

PT06

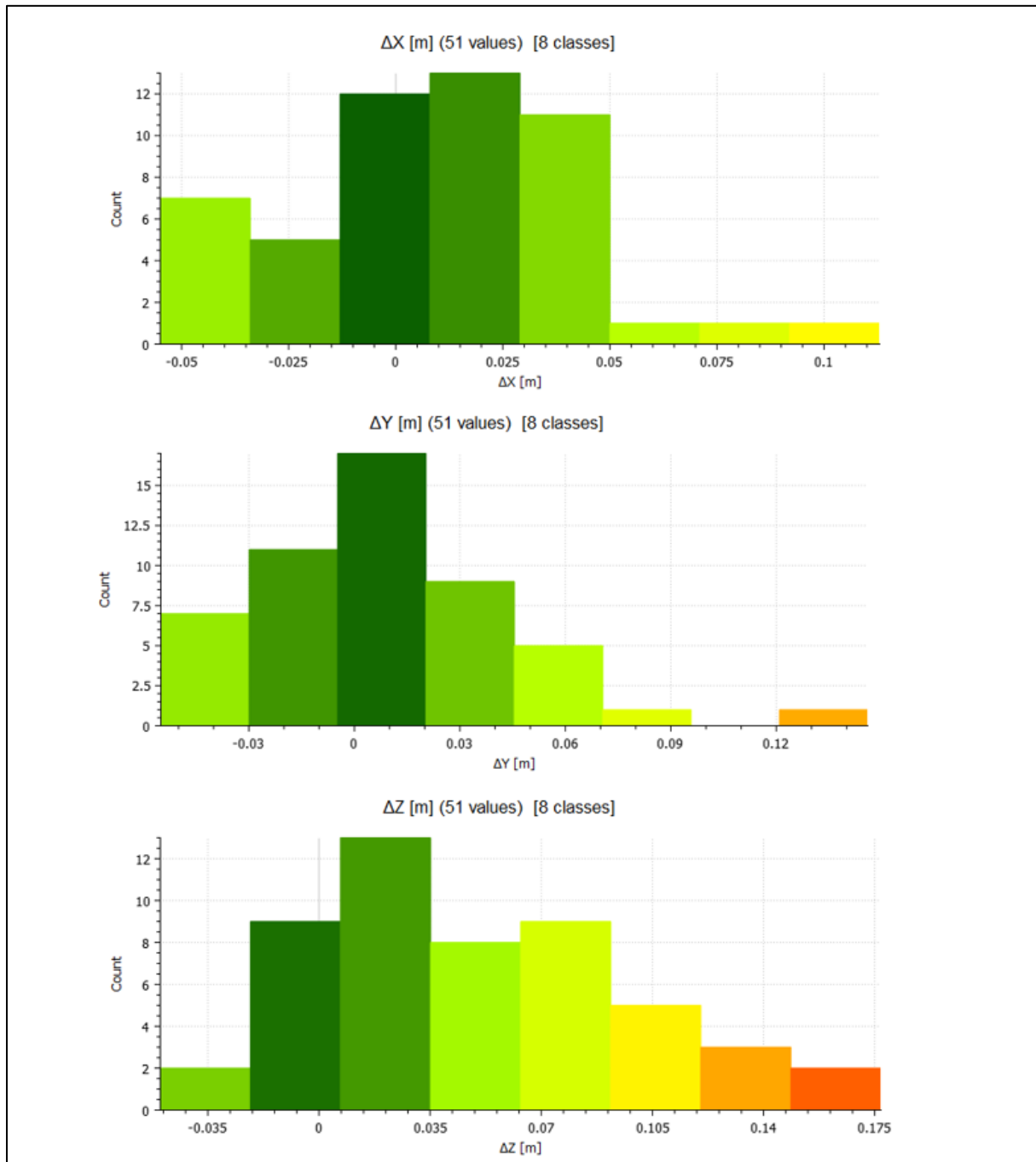
Anhang V

Statistik der Trajektoriengenauigkeit für einzelne Messungen nach Post-Processing in der Applanix Software POSPac MMS. Es wurden ein SmartBase-Netz aus DGNSS-Referenzstationen für die Optimierung genutzt.

Messung Nummer	Länge [km]	RMSE Northing [m]					RMSE Easting [m]					RMSE Height [m]					Geschw. [km/h]
		Max.	Min.	Ø	σ		Max.	Min.	Ø	σ		Max.	Min.	Ø	σ		
1	3,0	0,141	0,016	0,033	0,036		0,132	0,014	0,030	0,033		0,147	0,048	0,063	0,029		46
2	3,9	0,016	0,015	0,016	0,000		0,016	0,014	0,015	0,000		0,049	0,048	0,049	0,000		44
3	2,4	0,180	0,014	0,072	0,057		0,148	0,014	0,059	0,045		0,202	0,046	0,098	0,051		43
4	4,0	0,021	0,016	0,017	0,001		0,019	0,015	0,016	0,001		0,056	0,052	0,053	0,001		42
5	1,3	0,020	0,017	0,017	0,001		0,019	0,015	0,016	0,001		0,056	0,054	0,054	0,000		42
6	7,3	0,337	0,016	0,095	0,092		0,278	0,016	0,086	0,079		0,409	0,053	0,129	0,101		47
7	8,0	0,086	0,016	0,026	0,017		0,083	0,015	0,026	0,016		0,105	0,055	0,063	0,012		47
8	7,8	0,121	0,017	0,025	0,020		0,106	0,015	0,024	0,018		0,137	0,055	0,062	0,016		48

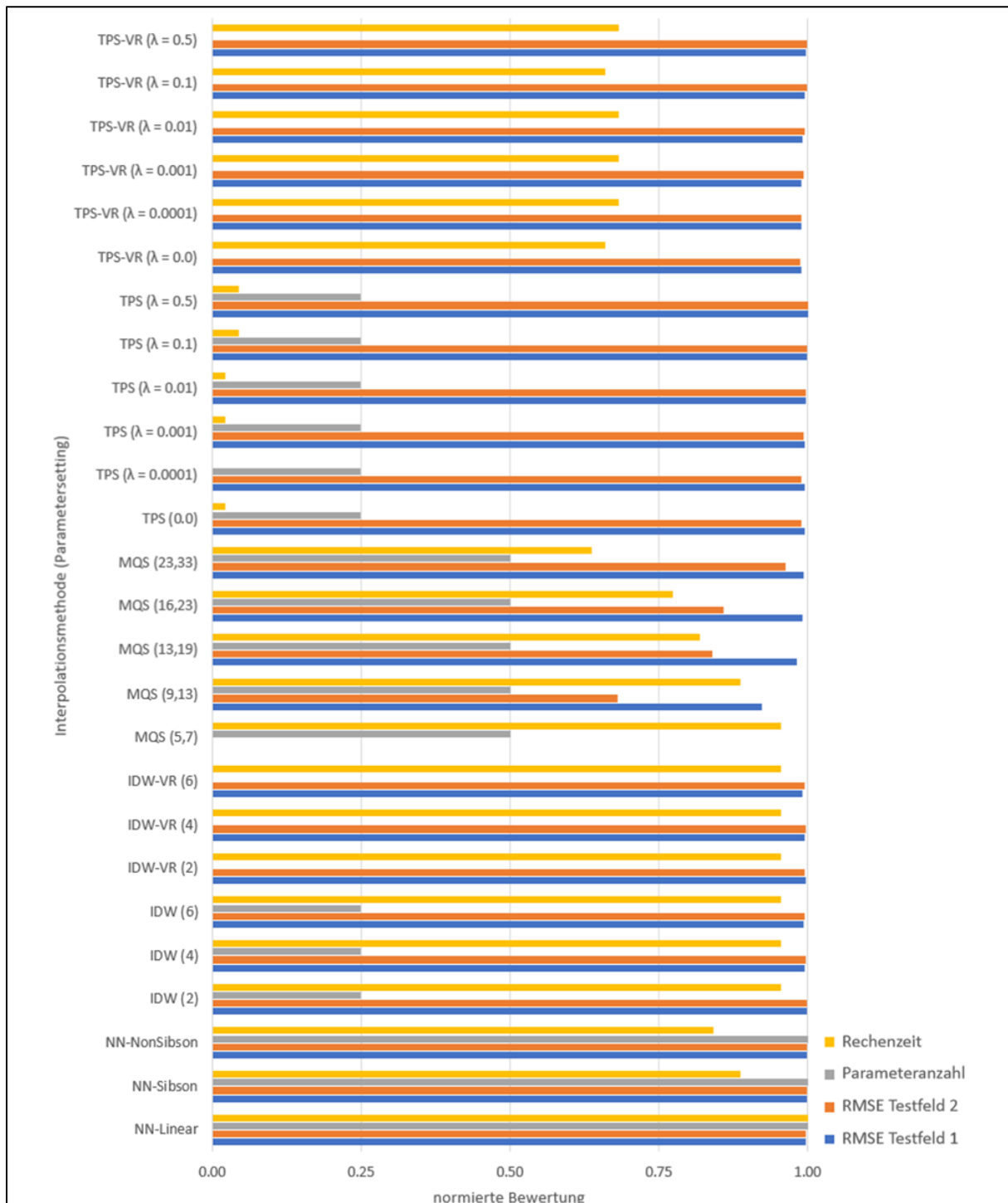
Anhang VI

Histogramme der an den 51 Passpunkten festgestellten Abweichungen für die drei Raumkomponenten X, Y und Z.



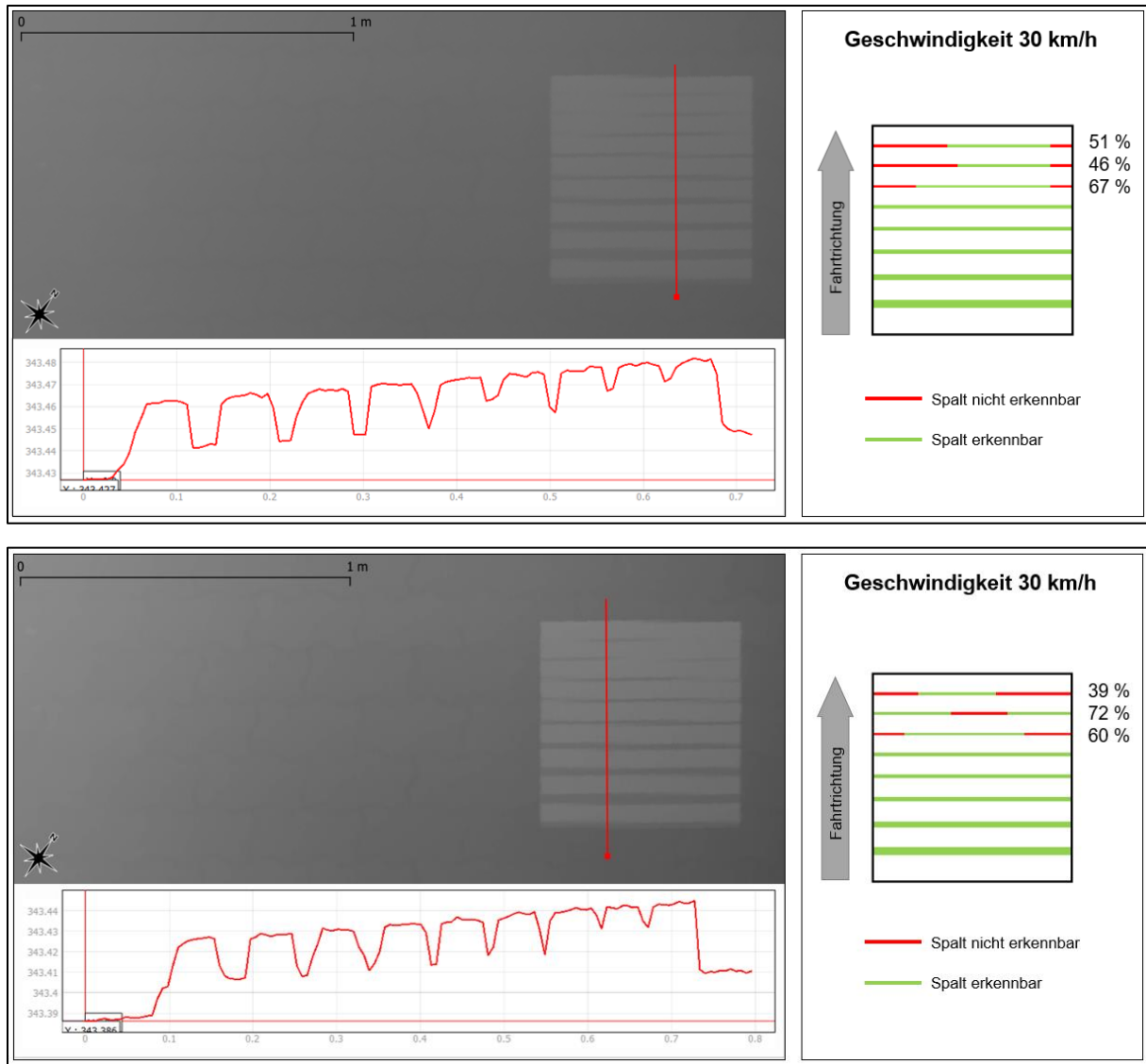
Anhang VII

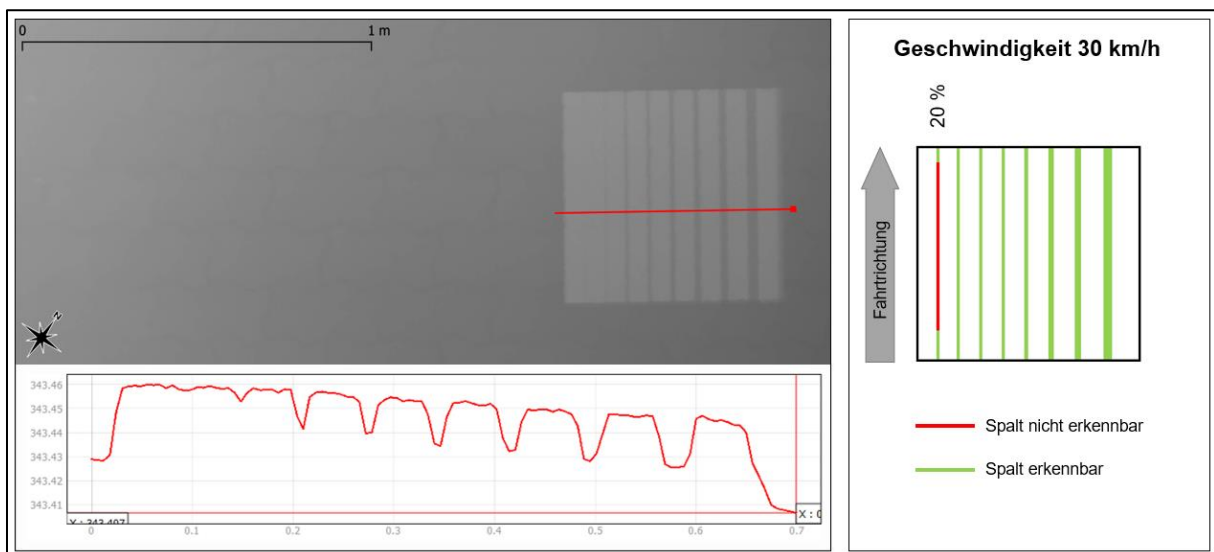
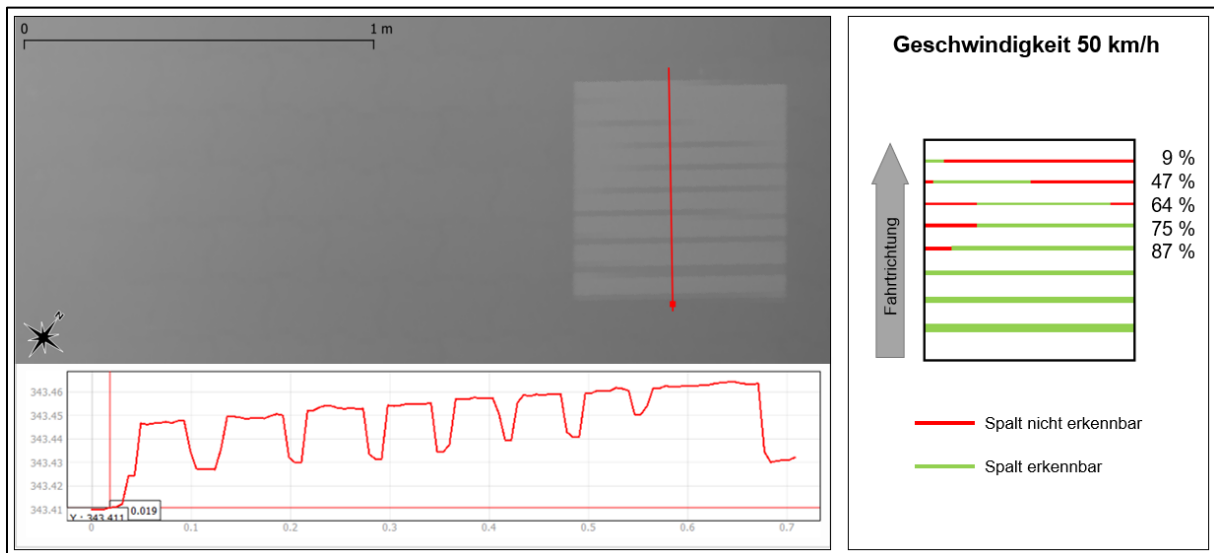
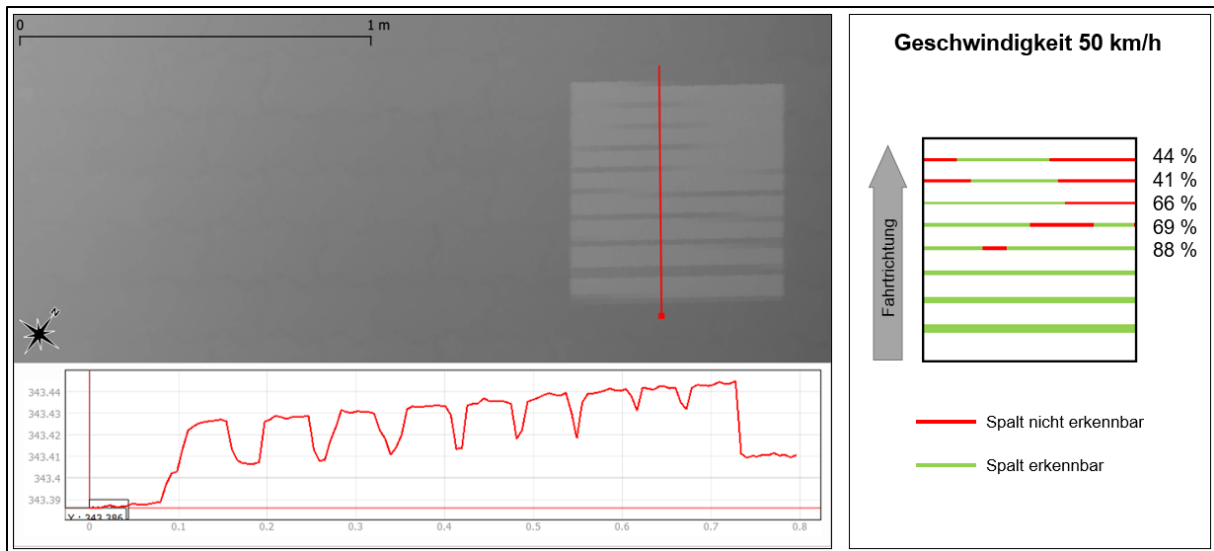
Balkendiagramm der normierten Bewertungskriterien für jede untersuchte Interpolationsmethode. Je mehr der Balken zum Wert 1,0 tendiert, desto besser ist die Bewertung für das jeweilige Kriterium. Das Fehlen eines Balkens verweist auf die im Vergleich schlechteste Merkmalsausprägung.

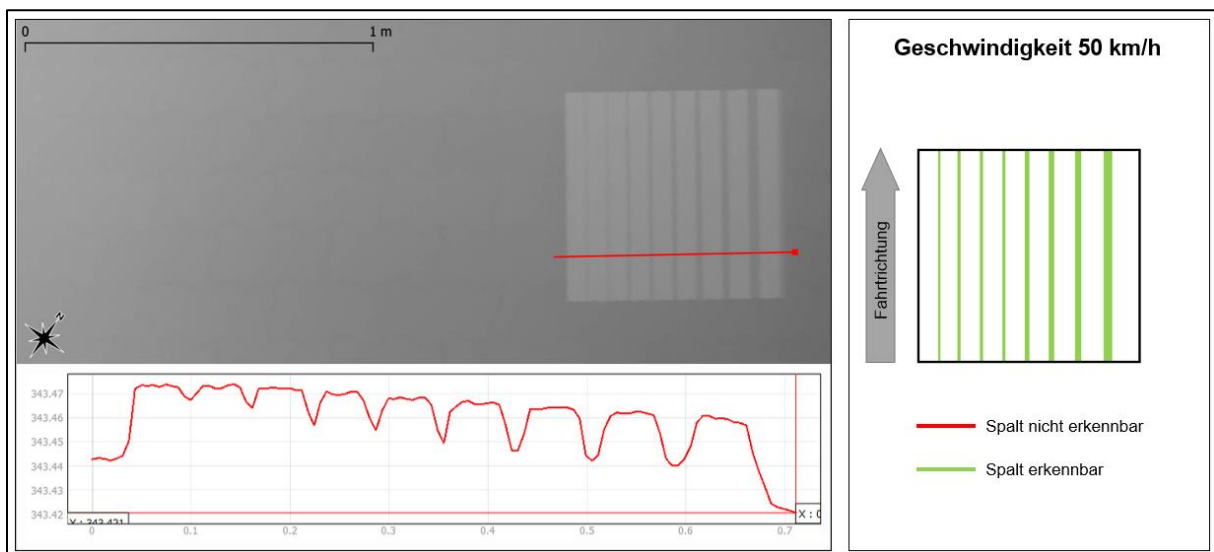
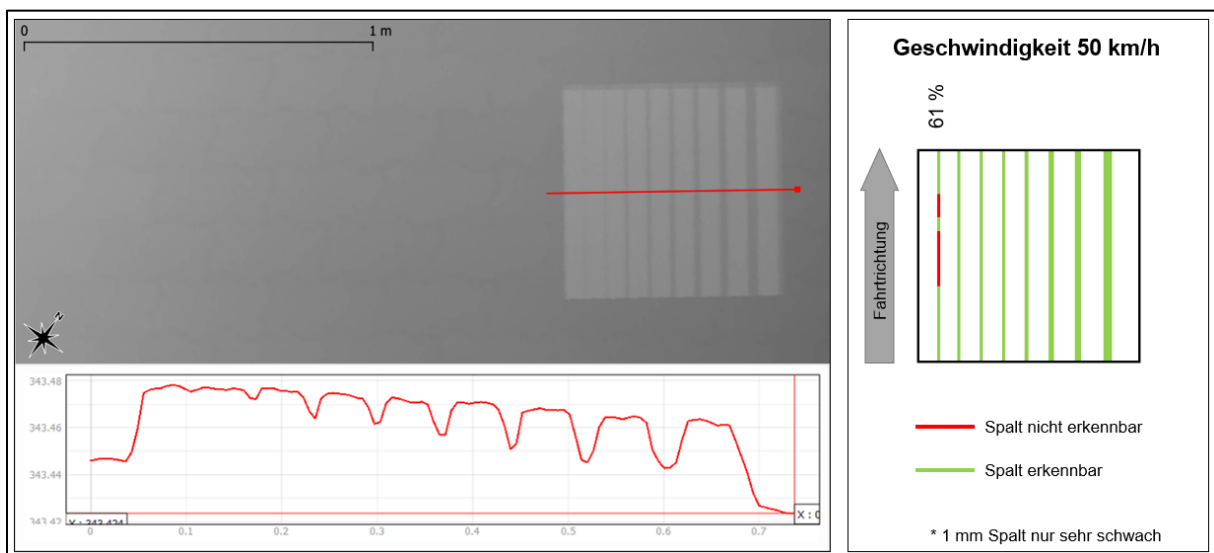
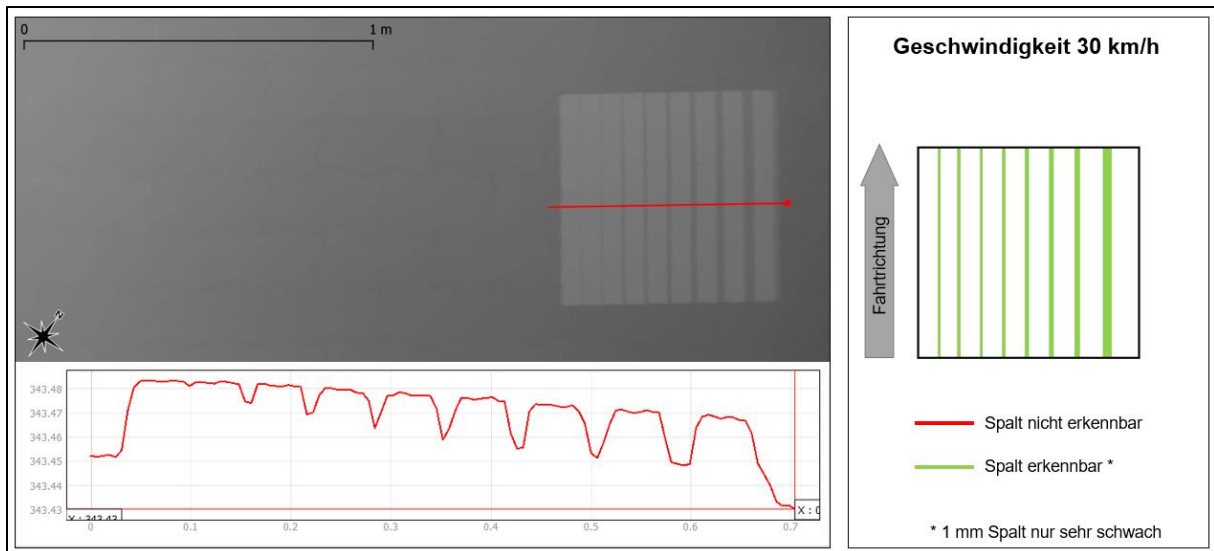


Anhang VIII

Erkennbarkeit von Spalten im Höhenmodell des Prüfkörpers bei verschiedenen Szenarien in wiederholten Messungen. Variiert wurden die Geschwindigkeit sowie die Ausrichtung des Prüfkörpers relativ zur Befahrungsrichtung.







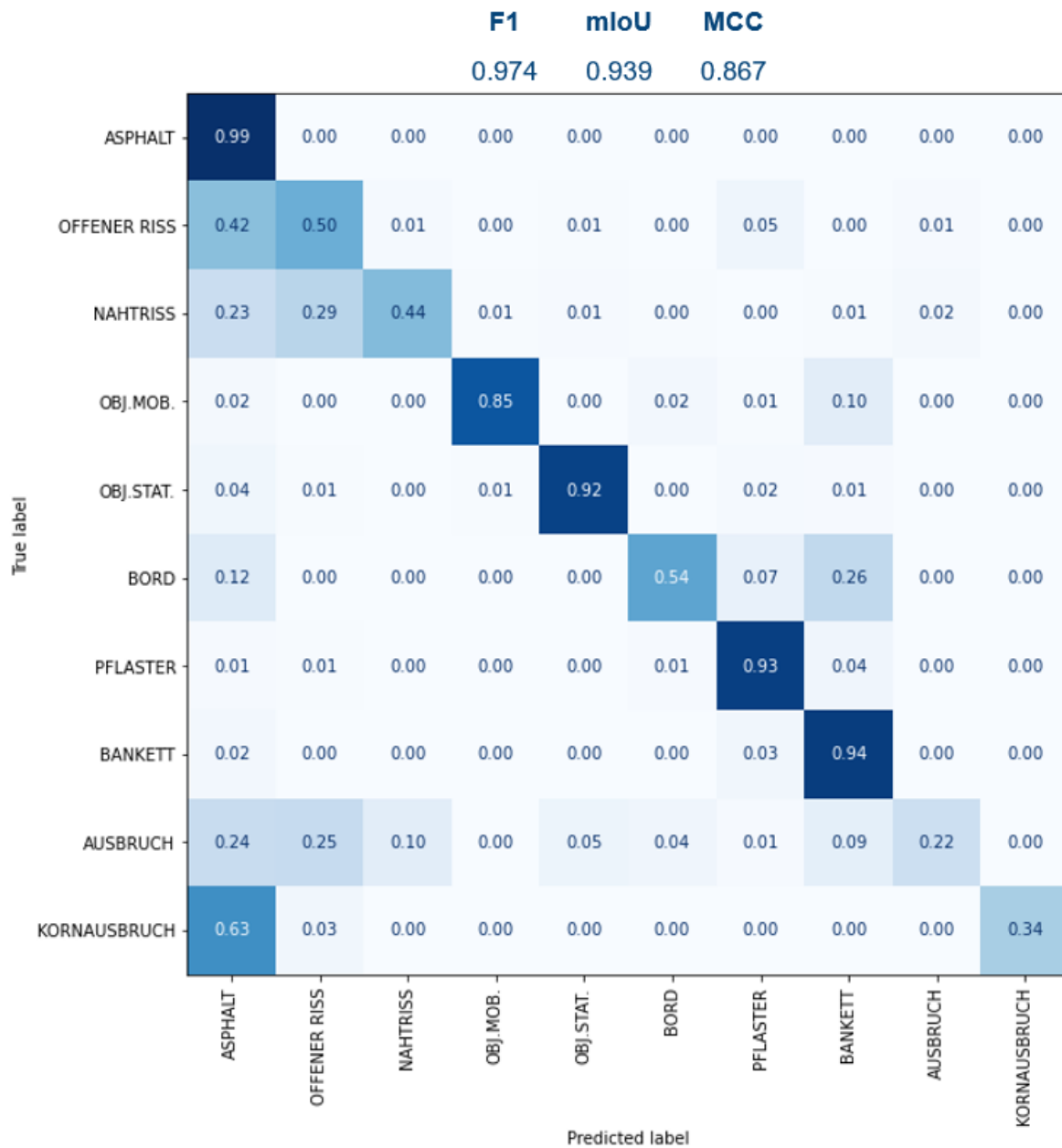
Anhang IX

Konfusionsmatrix des Modells, dass auf Basis von Hillshade-Rastern des reduzierten Z-A/B-Datensatzes für 150k Iterationen trainiert wurde.



Anhang X

Konfusionsmatrix des Modells, dass auf Basis von Topographic-Position-Index-Rastern des reduzierten Z-A/B-Datensatzes für 150k Iterationen trainiert wurde.

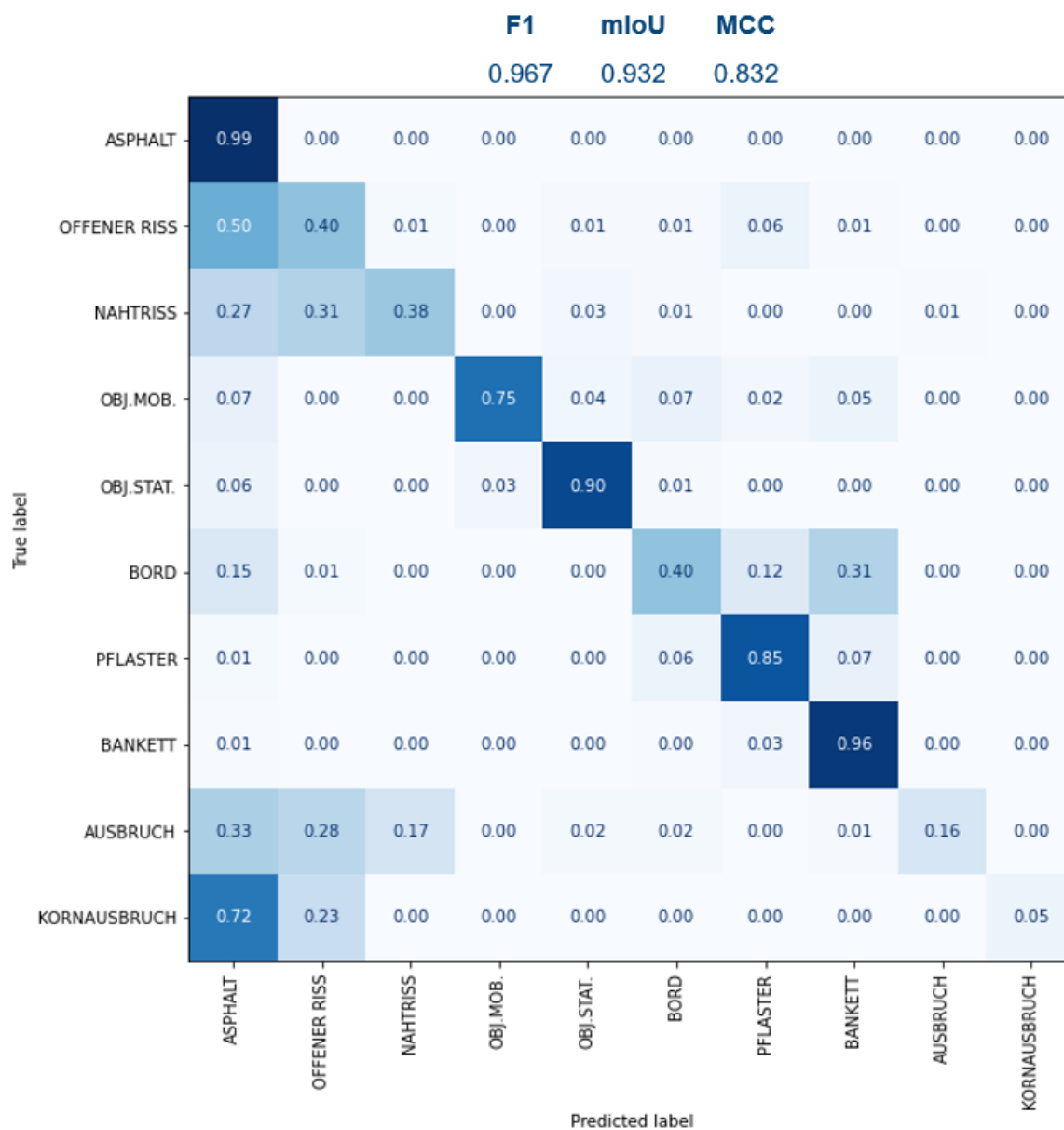


Anhang XI

Konfusionsmatrix des Modells, dass auf Basis von Positive-Openness-Rastern des reduzierten Z-A/B-Datensatzes für 150k Iterationen trainiert wurde.



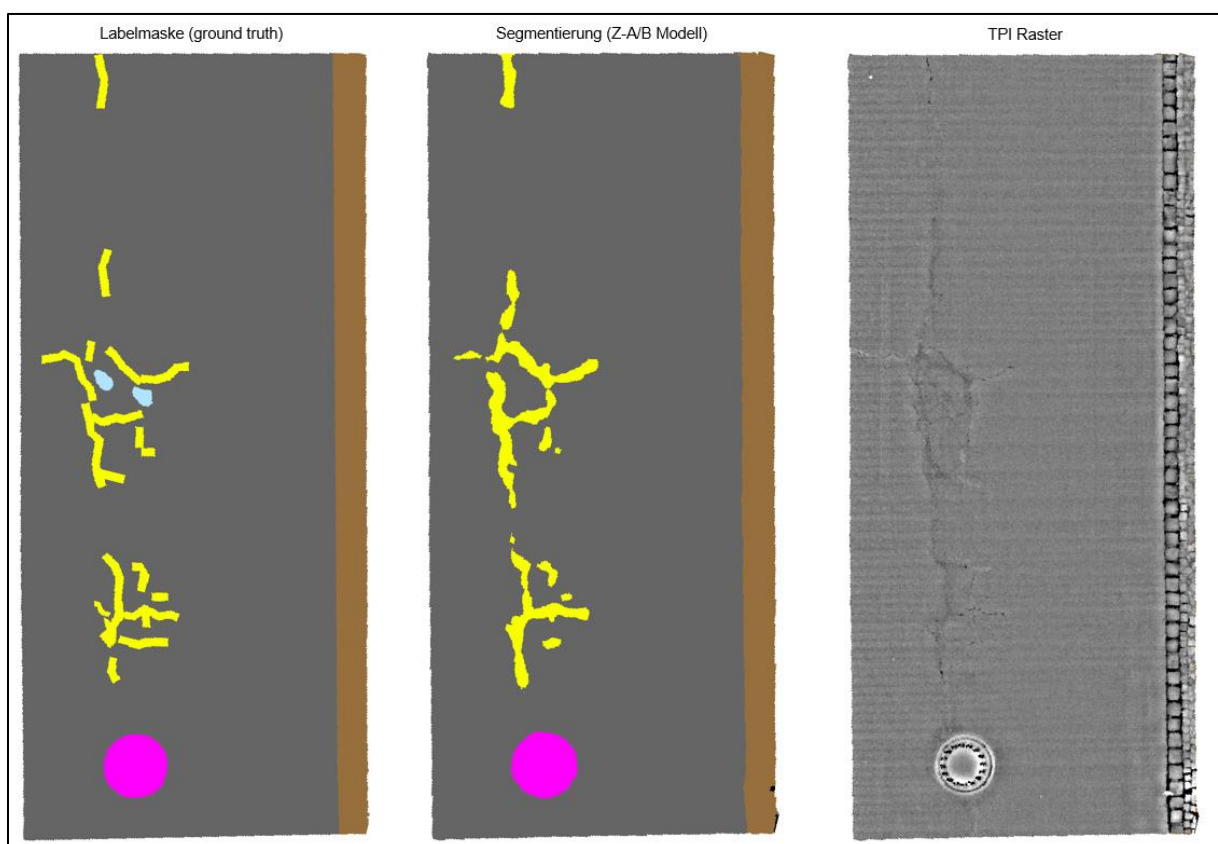
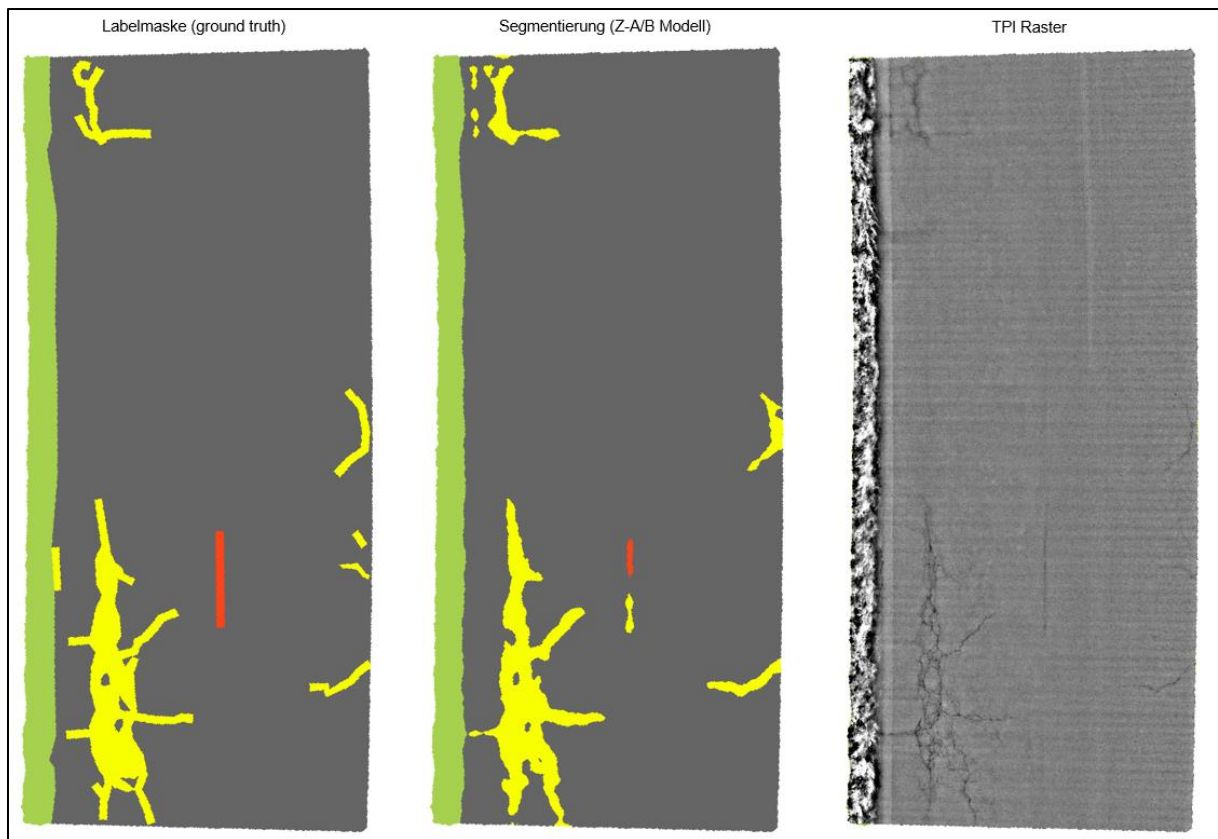
Konfusionsmatrix des Modells, dass auf Basis von Black-Top-Hat-Transform-Rastern des reduzierten Z-A/B-Datensatzes für 150k Iterationen trainiert wurde.

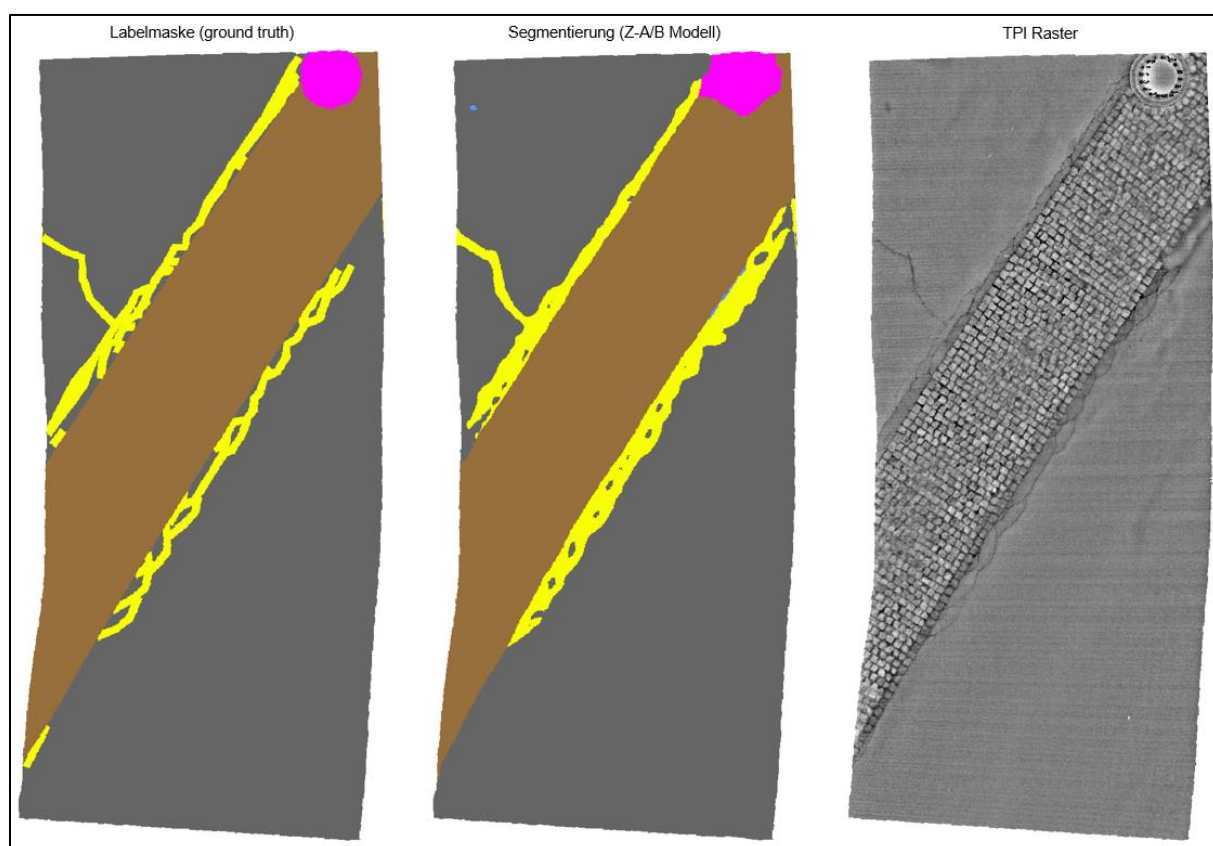
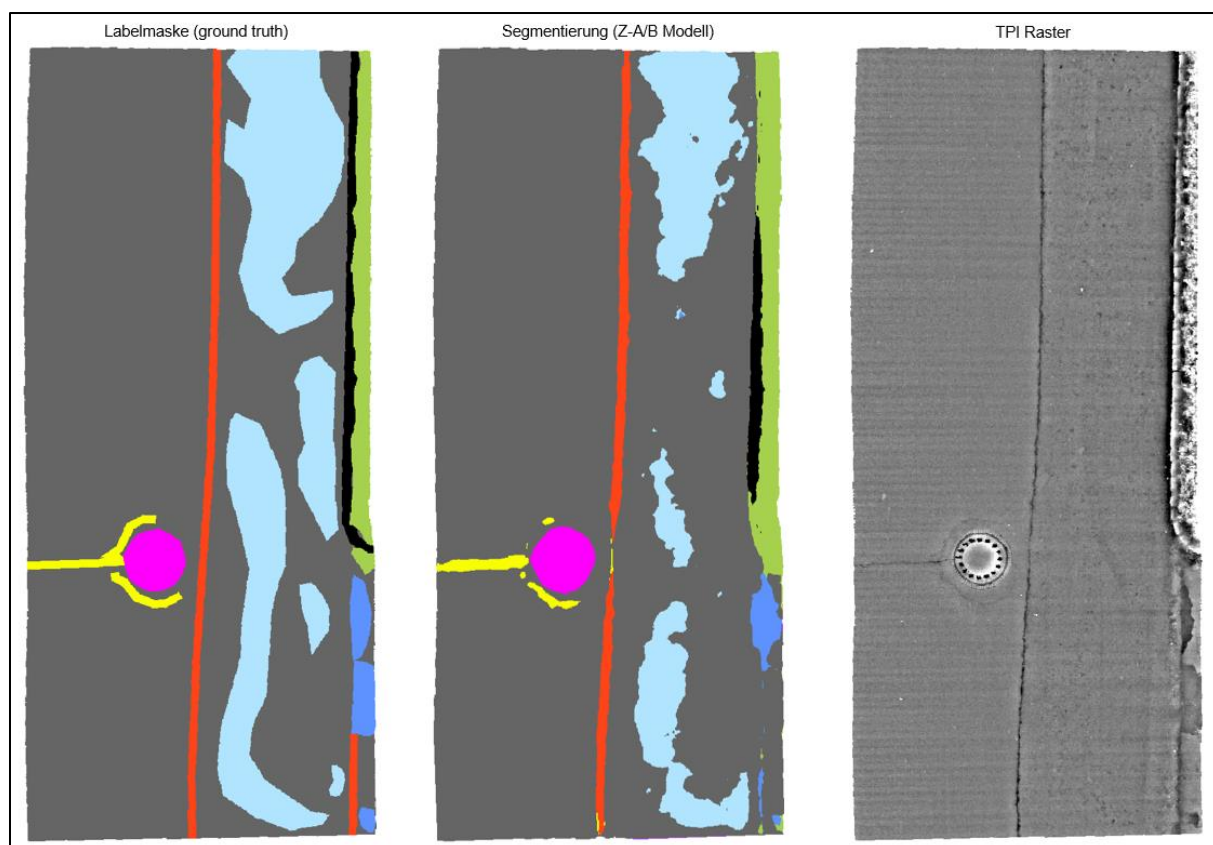


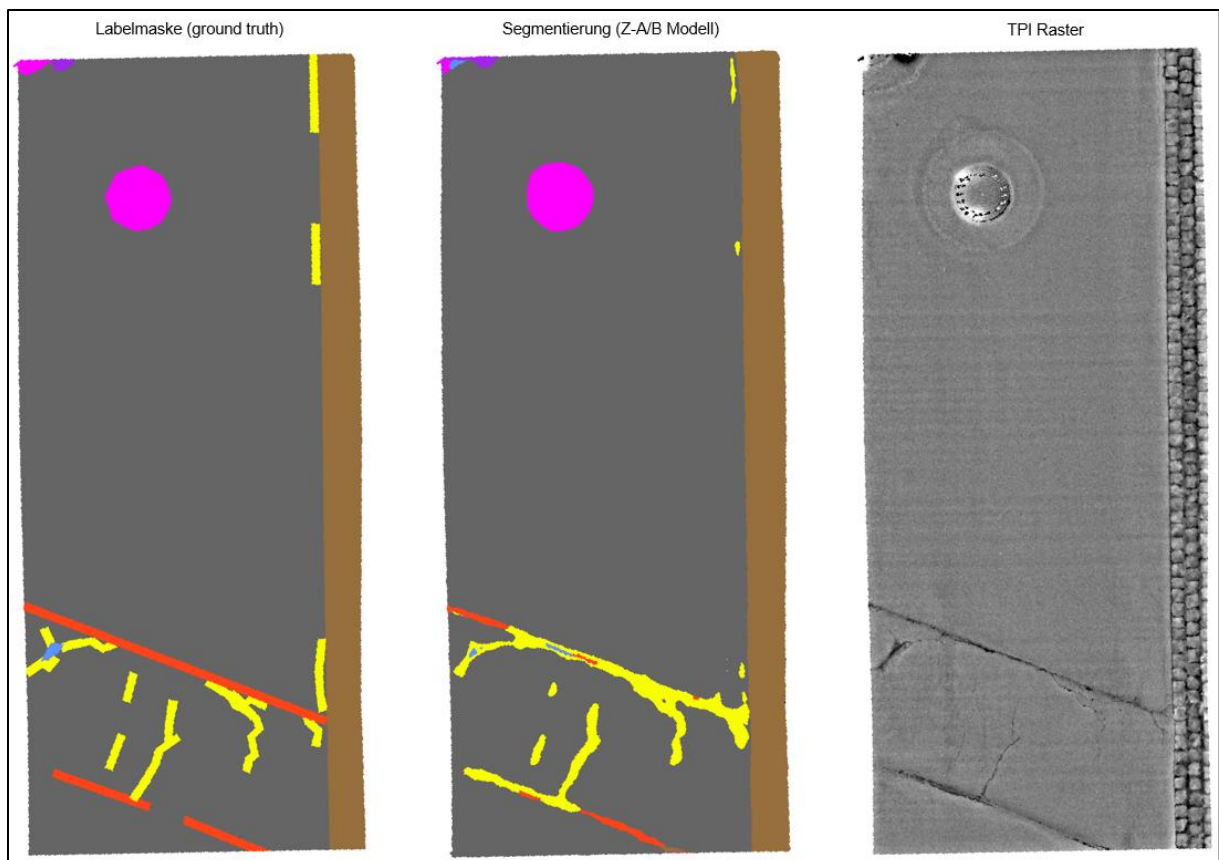
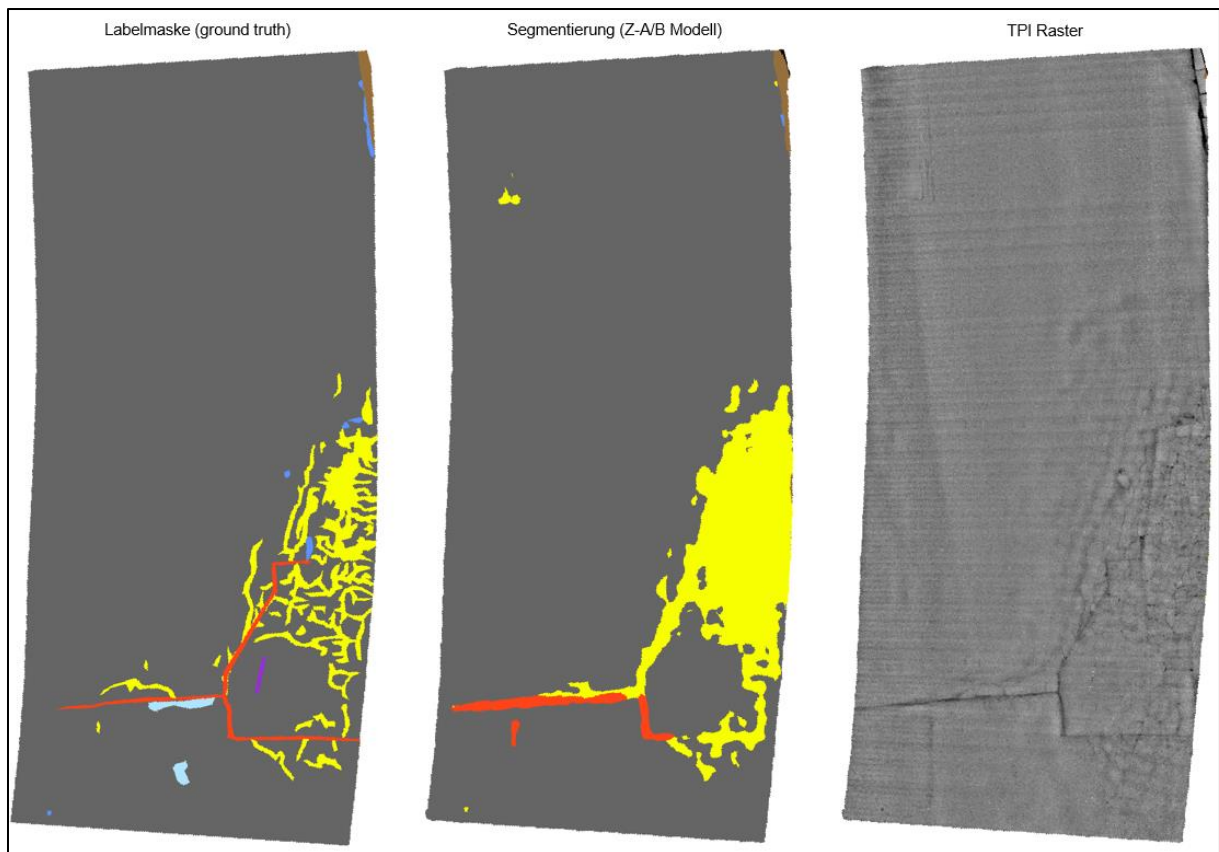
Anhang XIII

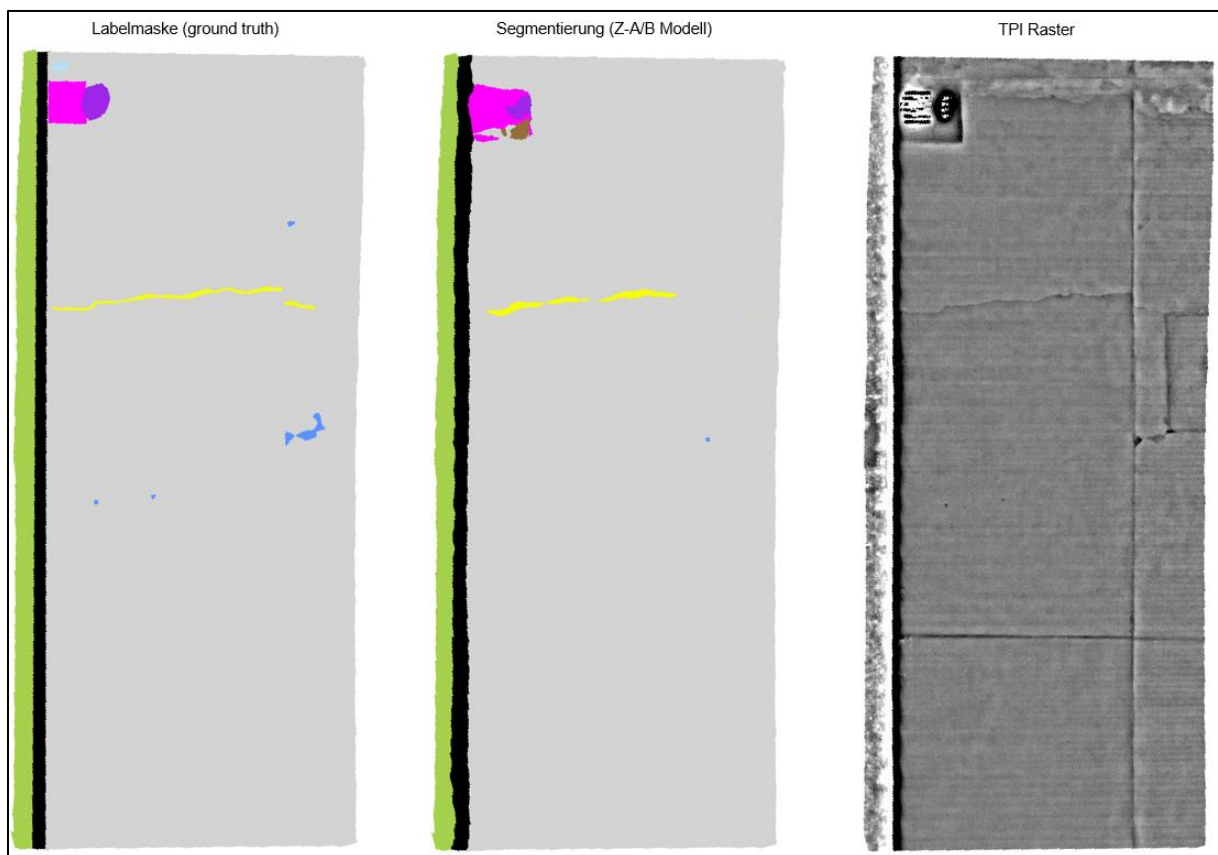
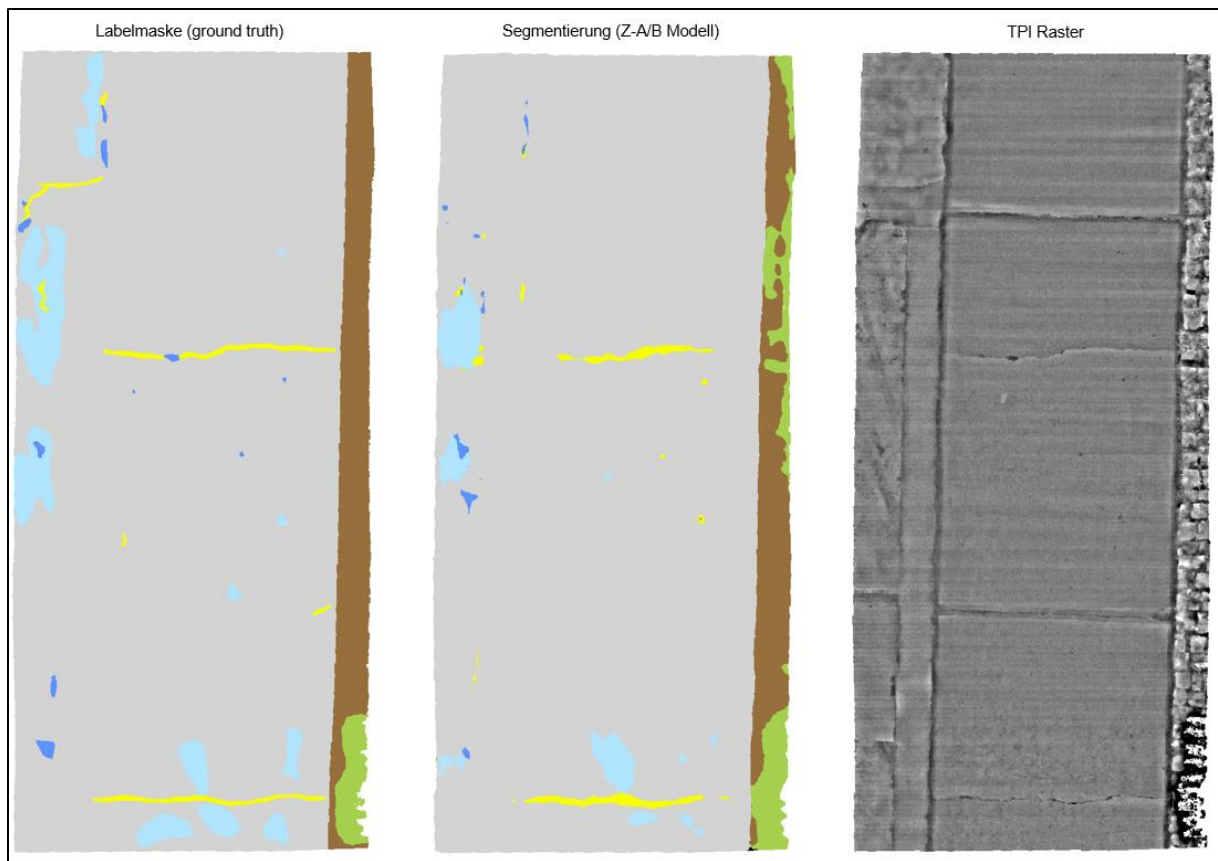
Auf den folgenden Seiten sind Beispiele des Testdatensatzes dargestellt. Die Abbildungen zeigen die manuell gelabelte Referenzmaske (links), die Prädiktion des jeweiligen finalen DeepLabv3+ Modells (Mitte) und das Eingabebild, welches der semantischen Segmentierung zugrundeliegt (rechts). Die nachstehend aufgeführte Legende gilt für alle Abbildungen von Anhang XIII.

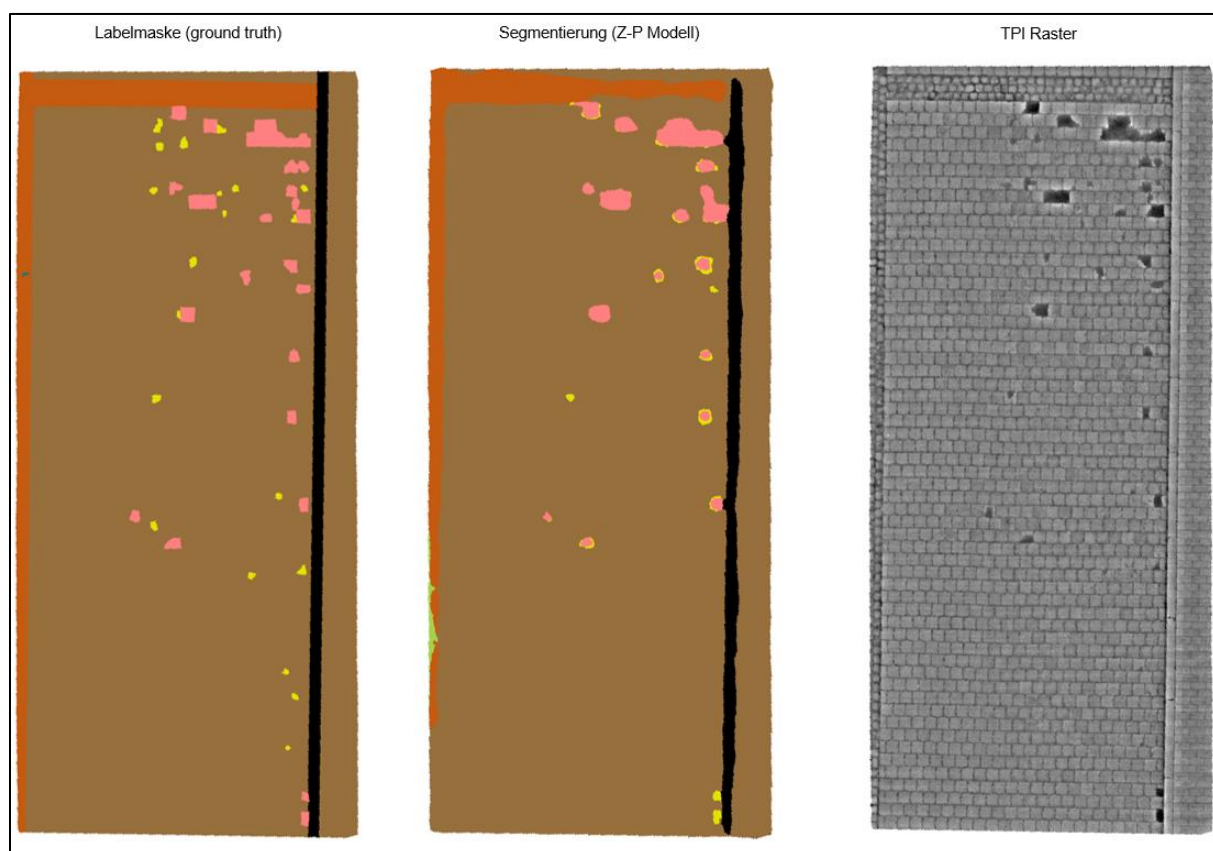
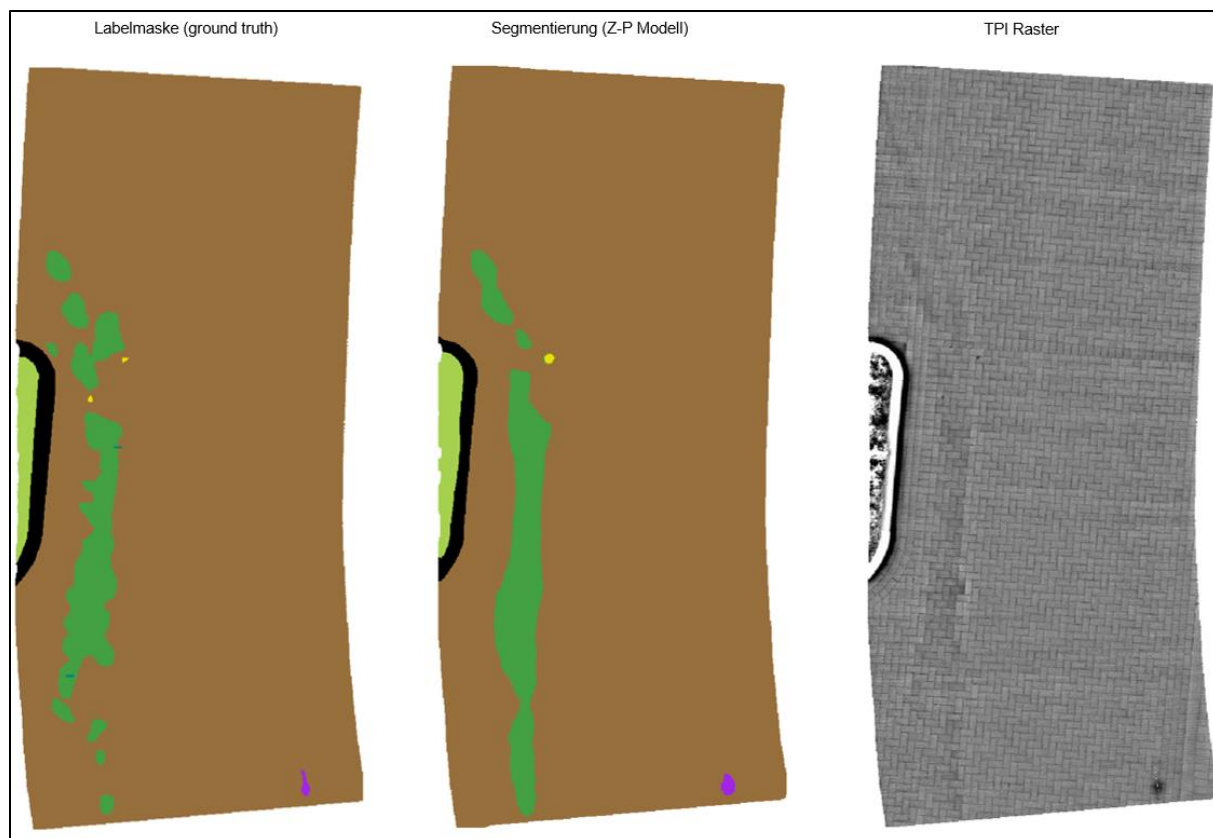


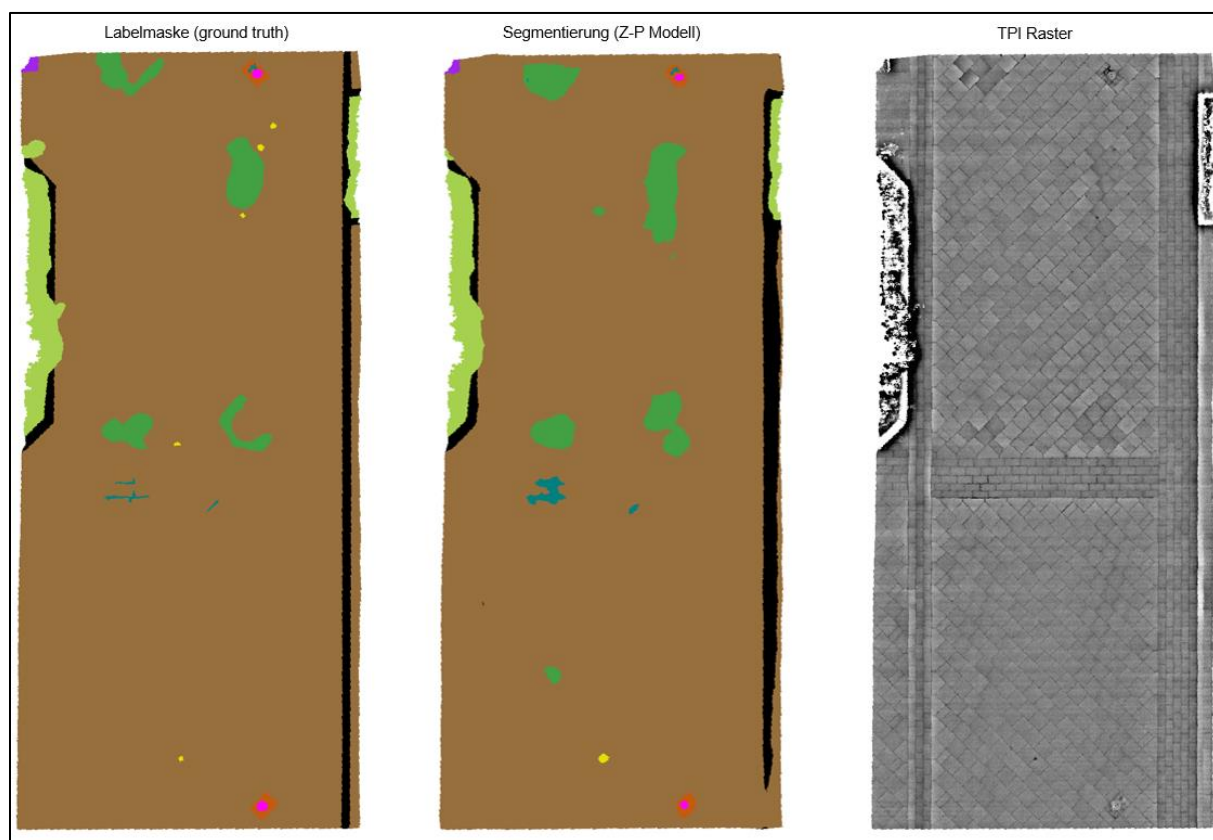
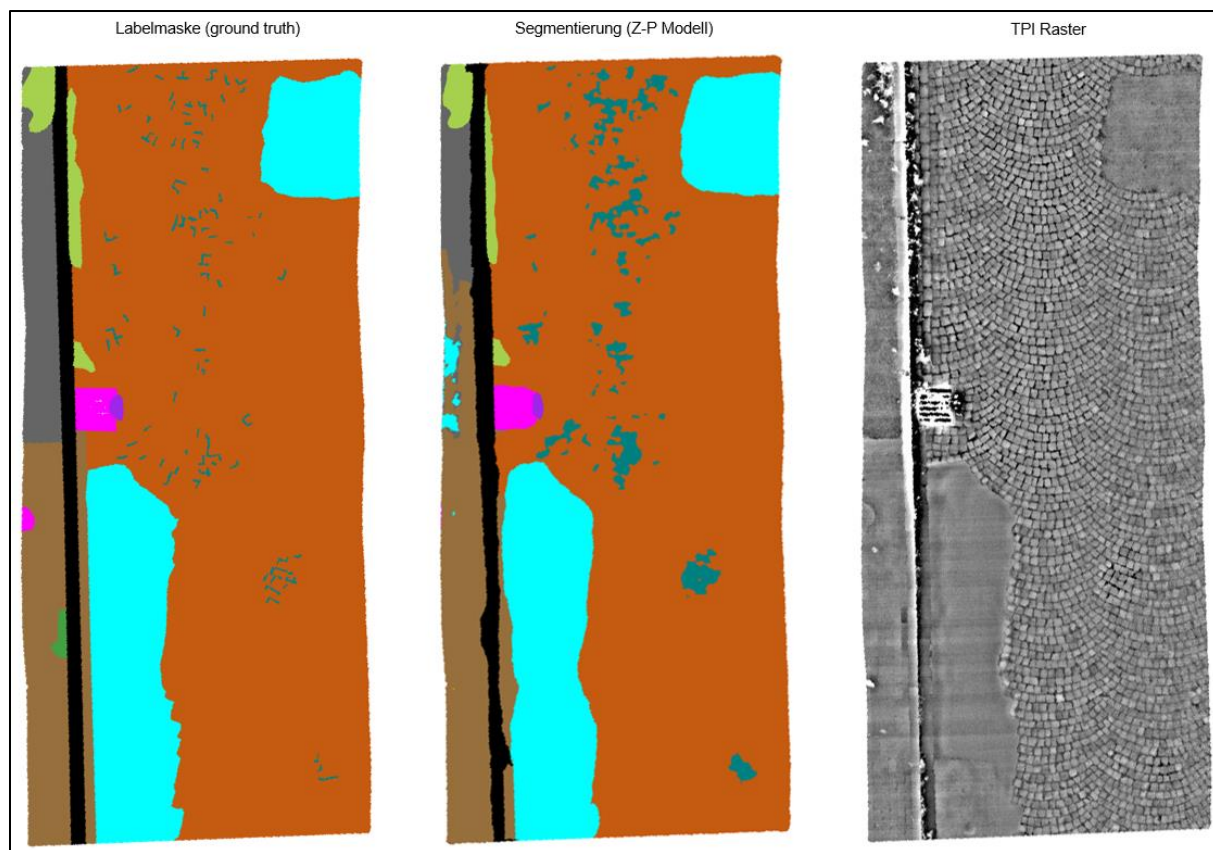


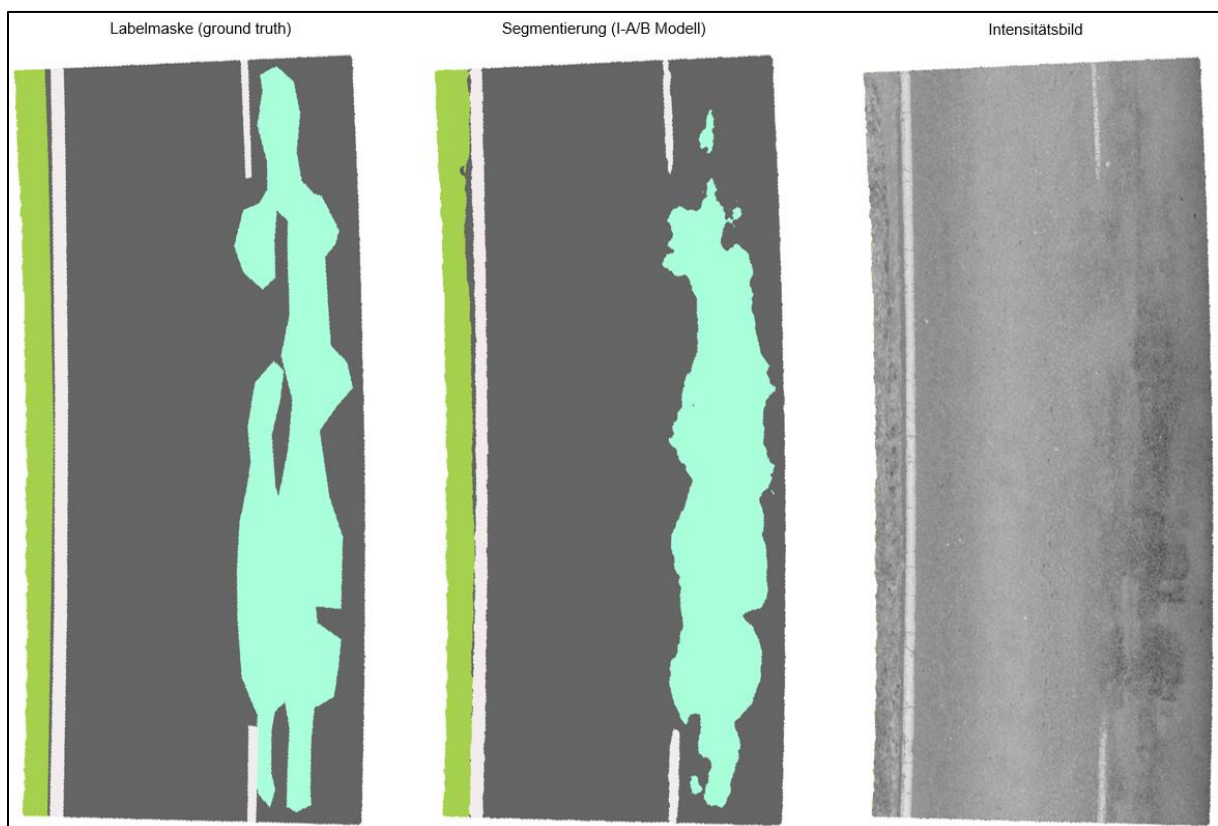
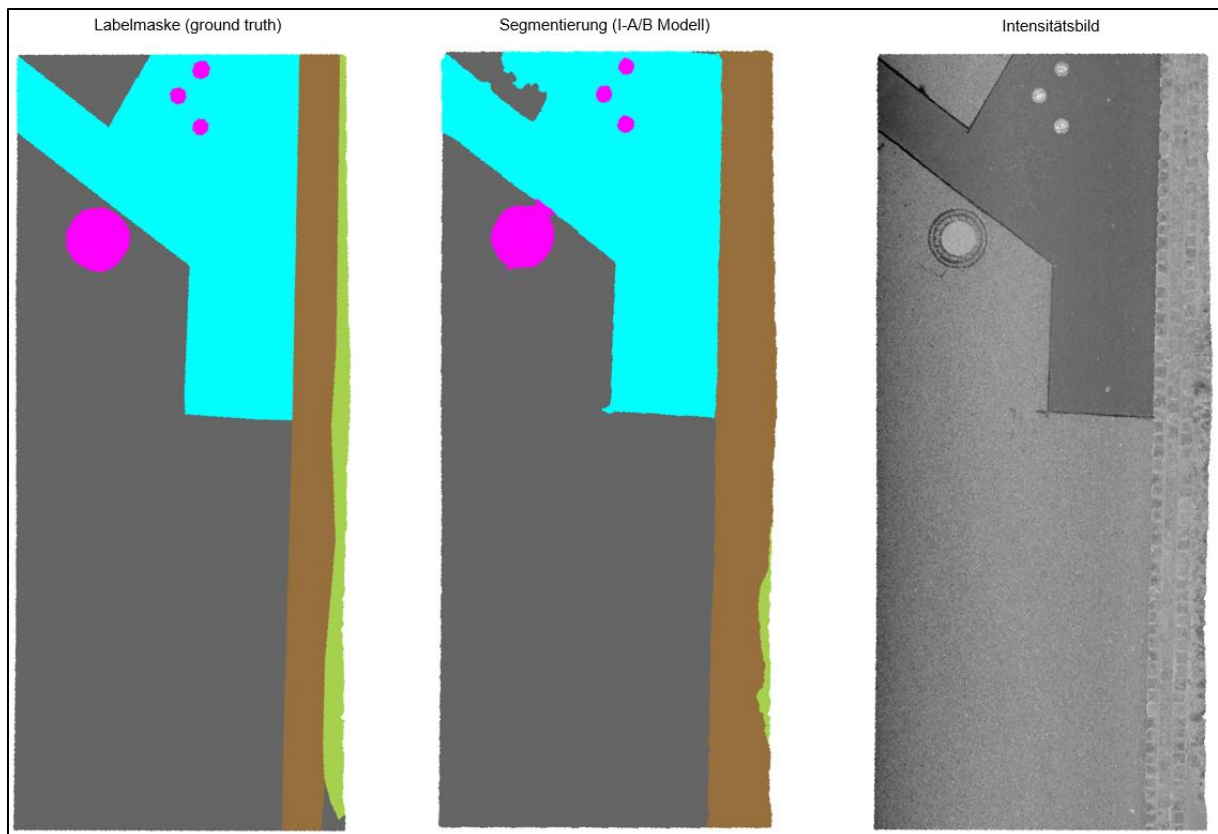


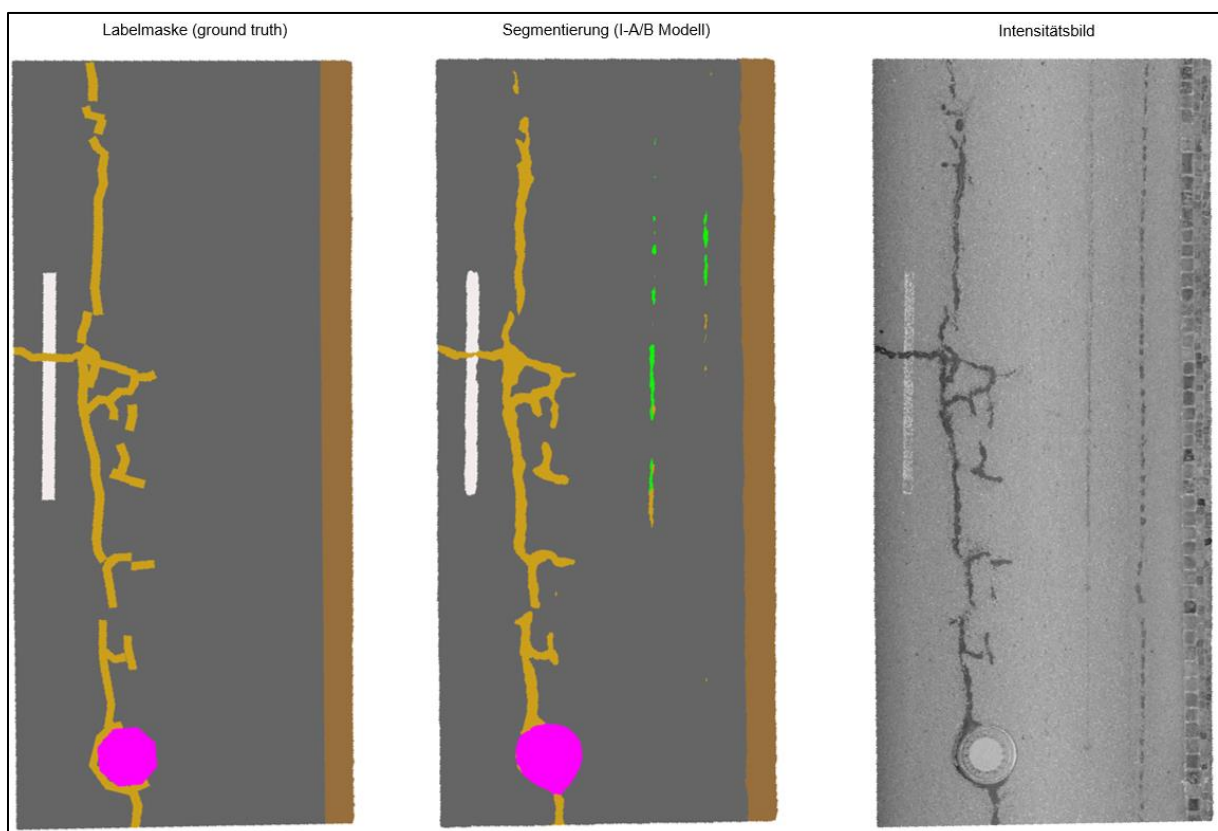
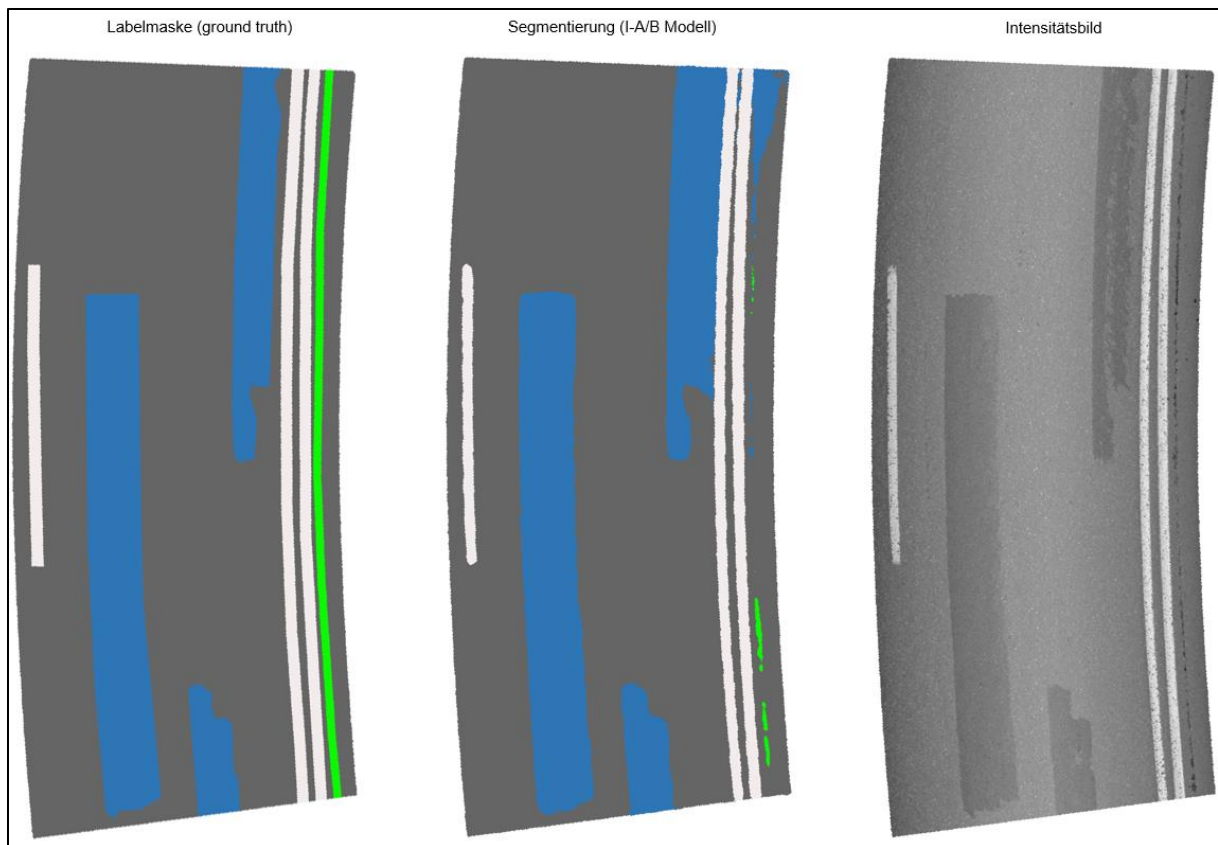


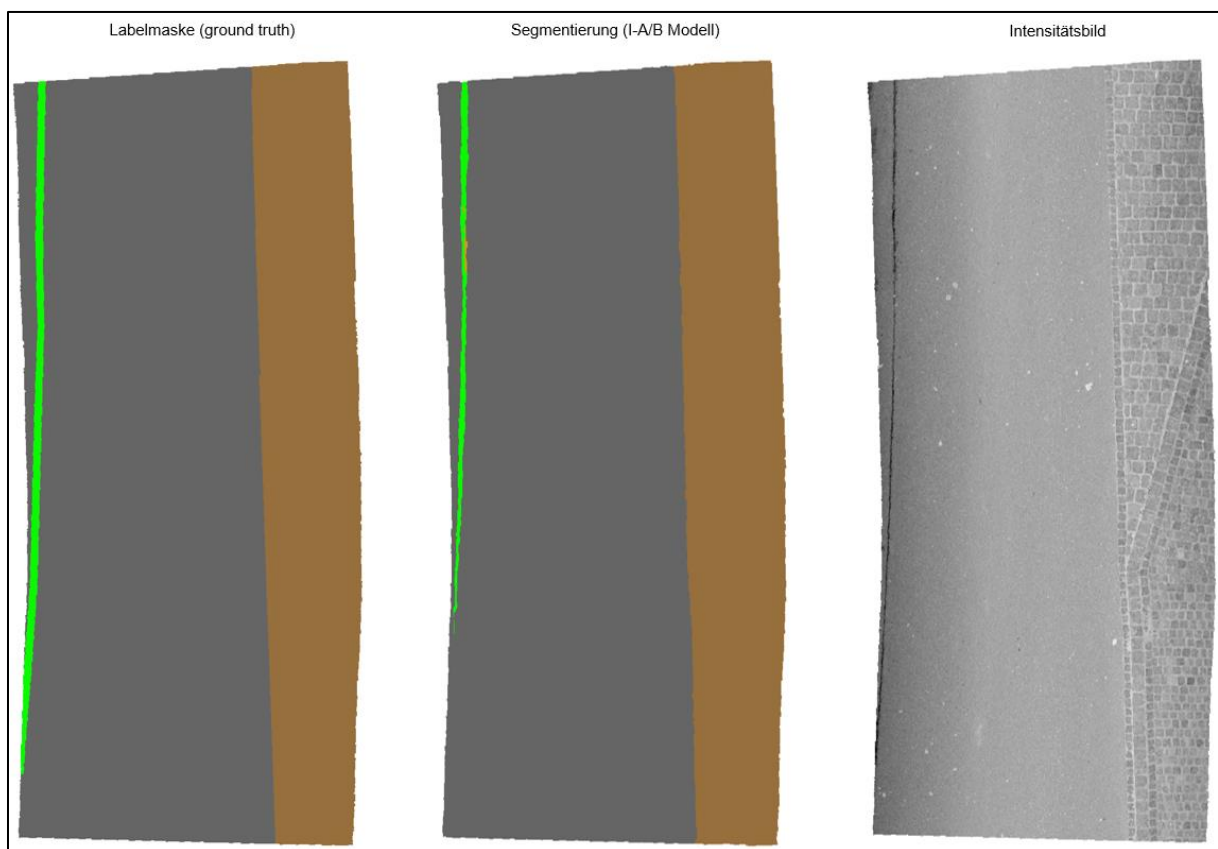
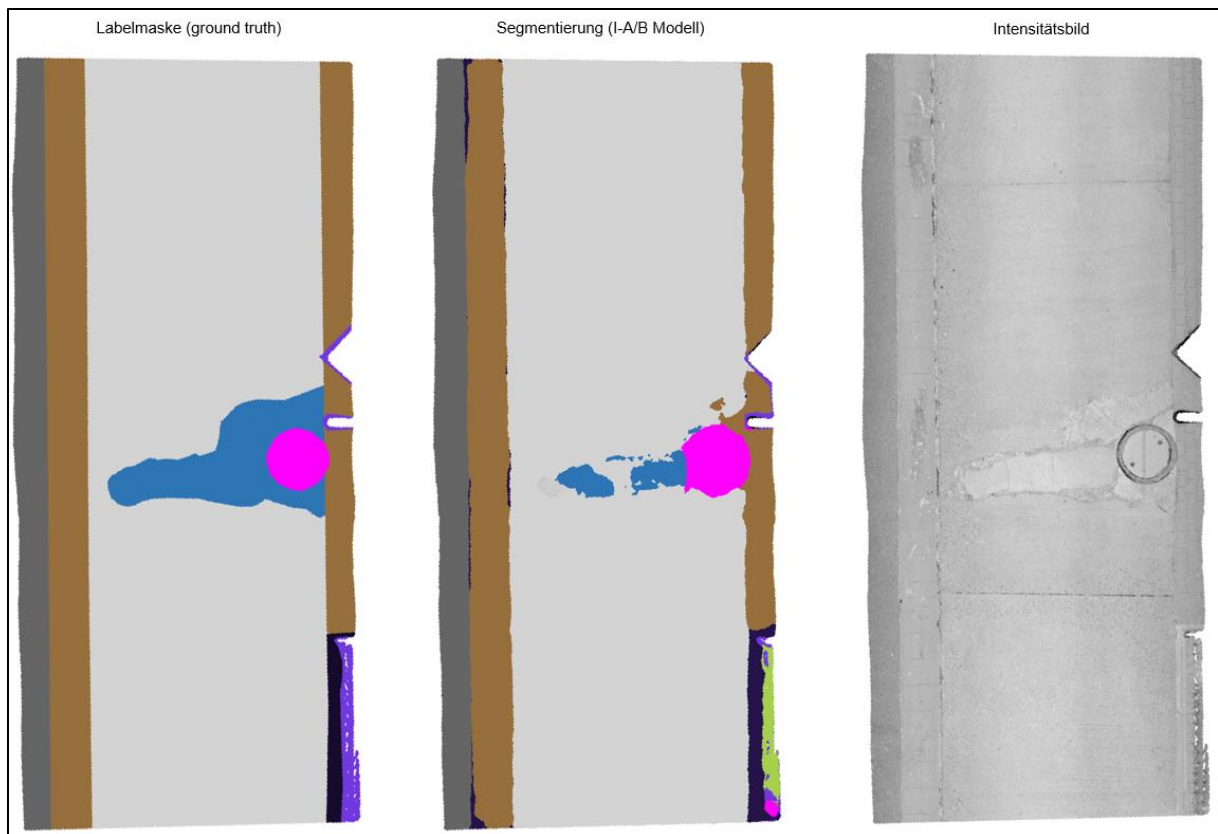






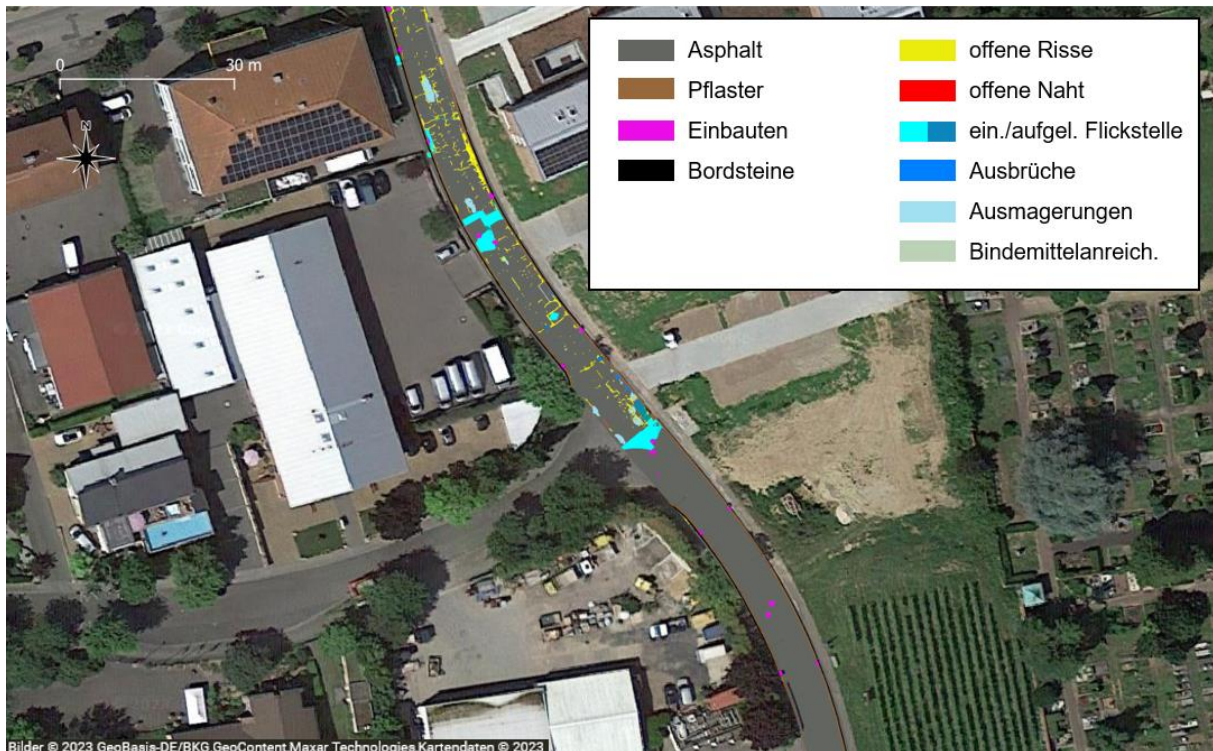


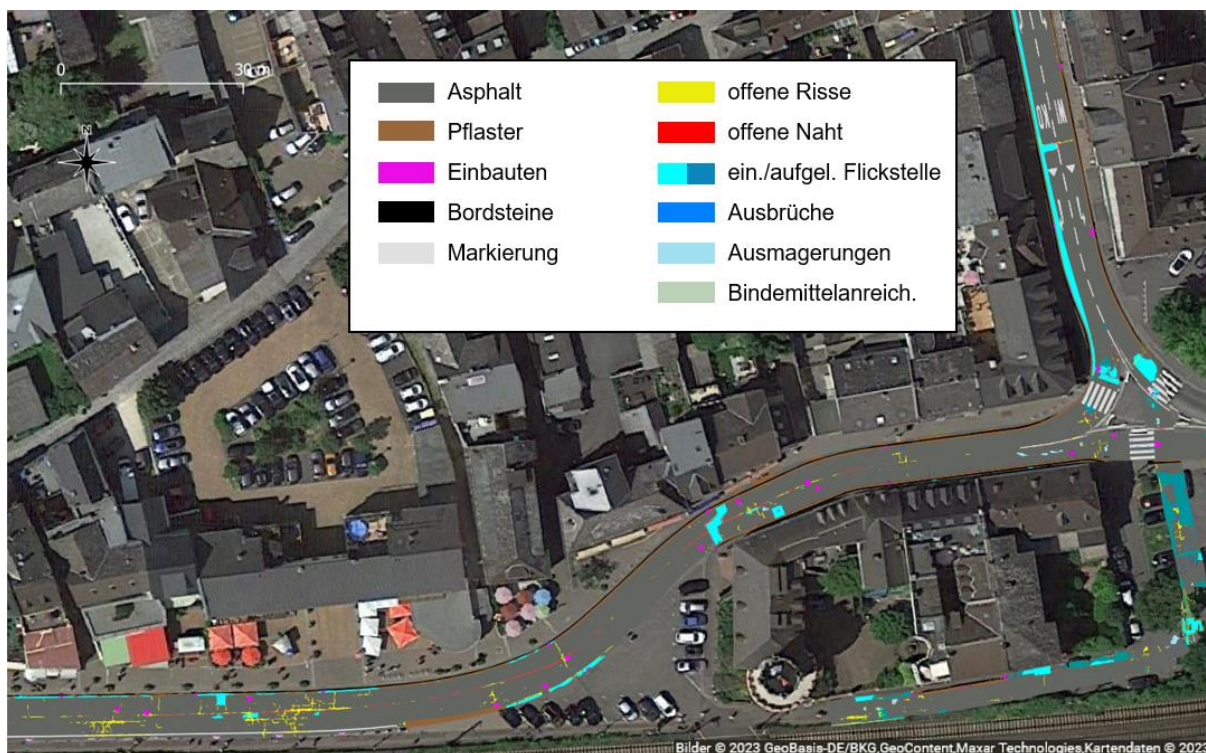




Anhang XIV

Beispiele für die Ergebnisse der automatischen Kartierung aus PPS-Oberflächenmodellen. Die obere Abbildung zeigt hier jeweils die konturscharfe Kartierung von Materialien, Objekten und Schäden. Die untere Abbildung zeigt die Überführung in eine Ergebnisdatei (10 m Abschnitte), wie sie in der kommunalen Straßenerhaltung typischerweise verwendet wird.







Anhang XV

Vergleich von manueller und automatischer Auswertung mithilfe der Eigenüberwachung gemäß ZTV ZEB-Stb (FGSV 2006). Die verwendeten Normierungsfunktionen nutzen die definierten Fixpunkte für Ortsdurchfahrten. Die zulässigen Toleranzen werden eingehalten.



Messkilometer	Erstmessung		Wiederholungsmessung		Differenzen	
	RISS-1	FLI-1	RISS-2	FLI-2	RISS	FLI
0.1	3.0	2.4	2.6	2.4	0.3	0.0
0.2	4.0	1.3	3.4	1.3	0.5	0.0
0.3	3.7	1.4	3.6	1.1	0.1	0.3
0.4	2.5	1.5	2.3	1.5	0.2	0.0
0.5	3.9	2.0	3.3	2.0	0.6	0.0
0.6	3.6	1.9	3.5	1.6	0.1	0.3
0.7	3.8	1.4	3.7	1.4	0.1	0.0
0.8	3.1	1.3	3.0	1.2	0.1	0.0
0.9	4.9	2.1	4.4	2.2	0.5	-0.1
1	5.0	3.0	4.9	2.8	0.1	0.2
1.1	2.2	5.0	2.1	4.9	0.1	0.1
1.2	2.4	3.8	2.2	3.9	0.1	-0.1
1.3	1.7	5.0	1.5	5.0	0.1	0.0
1.4	1.3	5.0	1.2	5.0	0.1	0.0
1.5	1.9	5.0	1.8	5.0	0.1	0.0
1.6	1.6	5.0	1.5	5.0	0.1	0.0
1.7	1.8	5.0	1.9	5.0	-0.1	0.0
1.8	2.9	4.8	3.1	4.6	-0.2	0.2
1.9	3.6	2.2	2.9	2.3	0.6	-0.1
2	3.9	3.6	3.3	3.7	0.7	-0.1

Anhang XVI

Parameter für die automatische Schadenserkennung in der Software LcmsRoadInspect auf Basis der LCMS2-Sensoren. Dargestellt sind nur die Abschnitte der XML-Datei, die die Modulparameter zur Erkennung der relevanten Schadensarten beinhalten (Cracking = offener Riss, SealedCrack = sanierter Riss, Patch = Flickstelle, Pothole = Ausbruch, Ravelling = Ausmagerung/Kornverluste, Bleeding = Bindemittelanreicherung).

```
<CrackingModule_Parameters>
<CrackingModule_AutomaticPeakDetection>1</CrackingModule_AutomaticPeakDetection>
<CrackingModule_UserDefinedPeakDetectionThresL_mm>2.00</CrackingModule_UserDefinedPeakDetectionThresL_mm>
<CrackingModule_UserDefinedPeakDetectionThresR_mm>2.00</CrackingModule_UserDefinedPeakDetectionThresR_mm>
<CrackingModule_AutoPavementTypeDetection>0</CrackingModule_AutoPavementTypeDetection>
<CrackingModule_UserDefinedPavementType>1</CrackingModule_UserDefinedPavementType>
<CrackingModule_EdgeCrackingEnable>1</CrackingModule_EdgeCrackingEnable>
<LCMS2_CrackingModule_MinCrackDepth_mm>0.75</LCMS2_CrackingModule_MinCrackDepth_mm>
<CrackingModule_MinCrackDepthConcrete_mm>0.75</CrackingModule_MinCrackDepthConcrete_mm>
<CrackingModule_MinCrackDepthTinning_mm>0.00</CrackingModule_MinCrackDepthTinning_mm>
<CrackingModule_MinCrackLength_mm>175</CrackingModule_MinCrackLength_mm>
<CrackingModule_MinCrackLengthConcrete_mm>175</CrackingModule_MinCrackLengthConcrete_mm>
<CrackingModule_ReportDetailedCrkStats>0</CrackingModule_ReportDetailedCrkStats>
<CrackingModule_EnableCrackFaulting>0</CrackingModule_EnableCrackFaulting>
<CrackingModule_CrackFaultingDistanceApart_mm>100</CrackingModule_CrackFaultingDistanceApart_mm>
</CrackingModule_Parameters>

<SealedCrackModule_Parameters>
<SealedCrackModule_MinBlobArea_m2>0.02</SealedCrackModule_MinBlobArea_m2>
<SealedCrackModule_MinDiff_IOSmoothness>0.20</SealedCrackModule_MinDiff_IOSmoothness>
<SealedCrackModule_MaxIntensity>255</SealedCrackModule_MaxIntensity>
<SealedCrackModule_EdgeCrackingEnable>0</SealedCrackModule_EdgeCrackingEnable>
<SealedCrackModule_OffsetFromLaneMark_mm>0</SealedCrackModule_OffsetFromLaneMark_mm>
</SealedCrackModule_Parameters>

<PatchModule_Parameters>
<PatchModule_MinBlobArea_m2>0.10</PatchModule_MinBlobArea_m2>
<PatchModule_MinDiff_IOSmoothness>0.50</PatchModule_MinDiff_IOSmoothness>
<PatchModule_MinimumConfidentScore>0.35</PatchModule_MinimumConfidentScore>
<PatchModule_MinDiff_IOAverageIntensity>20</PatchModule_MinDiff_IOAverageIntensity>
<PatchModule_EnableDetectionOnConcrete>0</PatchModule_EnableDetectionOnConcrete>
<PatchModule_PatchDetectOutsideLaneLimits>0</PatchModule_PatchDetectOutsideLaneLimits>
</PatchModule_Parameters>

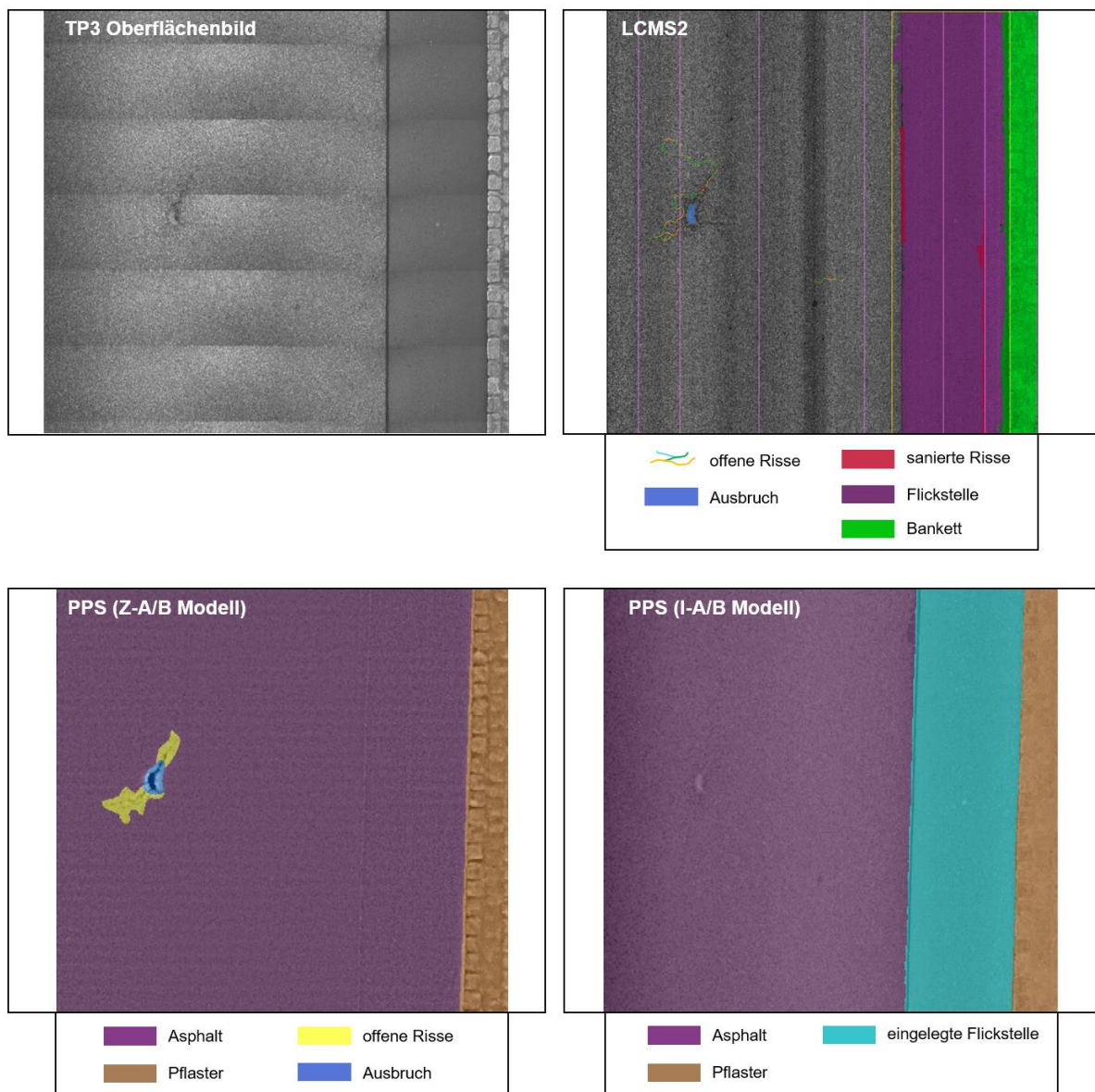
<PotholeModule_Parameters>
<PotholeModule_MinWidth_mm>150</PotholeModule_MinWidth_mm>
<PotholeModule_MaxWidth_mm>750</PotholeModule_MaxWidth_mm>
<PotholeModule_MinAvgDepth_mm>5.000000</PotholeModule_MinAvgDepth_mm>
<PotholeModule_MinArea_m2>0.005000</PotholeModule_MinArea_m2>
<PotholeModule_MaxAvgInt>150</PotholeModule_MaxAvgInt>
<PotholeModule_MinMinorAxisDiameter_mm>75</PotholeModule_MinMinorAxisDiameter_mm>
<PotholeModule_MaxRatioMaxMinAxes>4.000000</PotholeModule_MaxRatioMaxMinAxes>
</PotholeModule_Parameters>

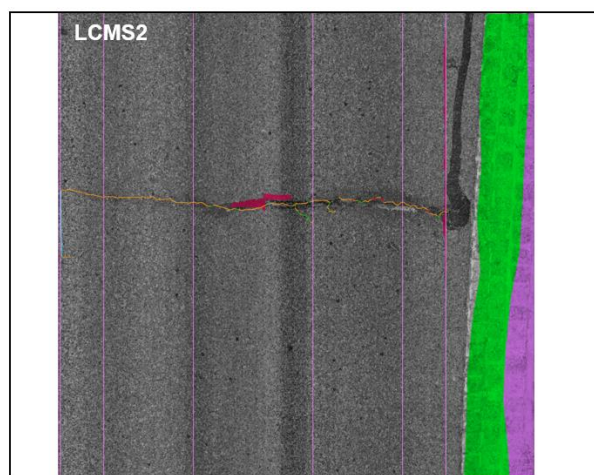
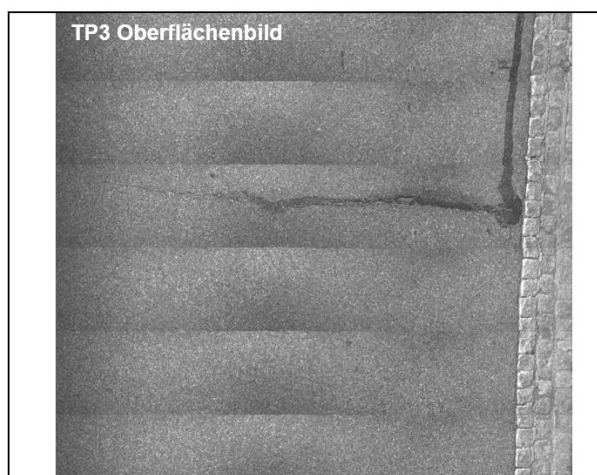
<RavelingModule_Parameters>
<RavelingModule_Threshold_cm3_m2>100</RavelingModule_Threshold_cm3_m2>
<RavelingModule_Zone_Width_mm>250</RavelingModule_Zone_Width_mm>
<RavelingModule_Zone_Height_mm>250</RavelingModule_Zone_Height_mm>
<RavelingModule_ExcludeCracksOnRav>0</RavelingModule_ExcludeCracksOnRav>
<RavelingModule_ChipSeal>0</RavelingModule_ChipSeal>
<RavelingModule_DisableOnConcrete>0</RavelingModule_DisableOnConcrete>
<RavelingModule_AutoPeakThresholdMode>0</RavelingModule_AutoPeakThresholdMode>
<RavelingModule_Peak_Threshold>2.500000</RavelingModule_Peak_Threshold>
</RavelingModule_Parameters>

<BleedingModule_Parameters>
<BleedingModule_AvgInterval_m>0.50</BleedingModule_AvgInterval_m>
<BleedingModule_SevereThreshold>2.00</BleedingModule_SevereThreshold>
<BleedingModule_MediumThreshold>1.75</BleedingModule_MediumThreshold>
<BleedingModule_LightThreshold>1.50</BleedingModule_LightThreshold>
</BleedingModule_Parameters>
```

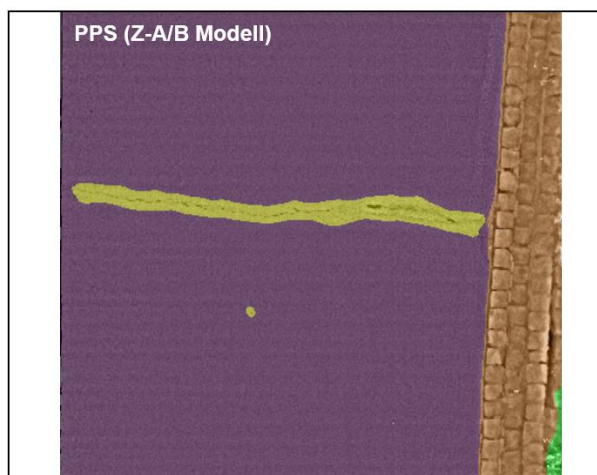

Anhang XVII

Beispiele für die Daten der beiden Referenzsysteme TP3-Oberflächenbild (Teilabbildung oben links) und LCMS2 (Teilabbildung oben rechts). Die untere Reihe der Teilabbildungen zeigt das Segmentierungsergebnis des Z-A/B-Modells (links) und des I-A/B-Modells (rechts) als farbkodierte Überlagerung über dem jeweiligen PPS-Oberflächenmodell. Die Abbildungen repräsentieren ca. 3 m x 4 m Straßenoberfläche.

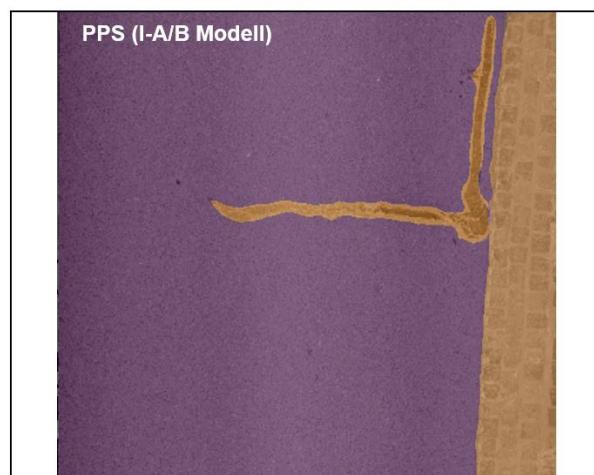




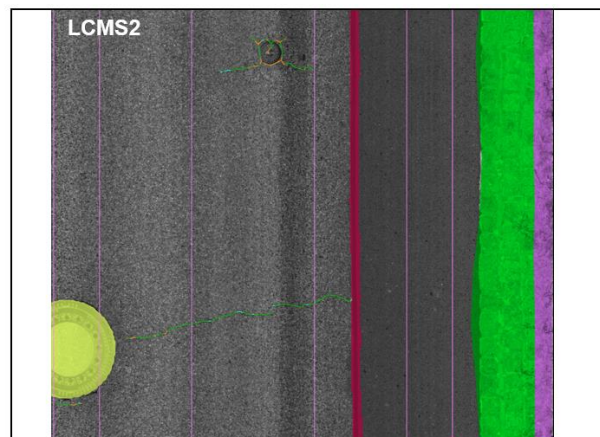
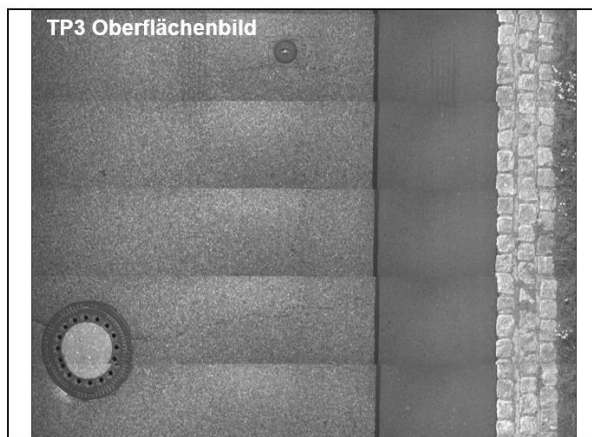
offene Risse sanierte Risse
Sonstiges



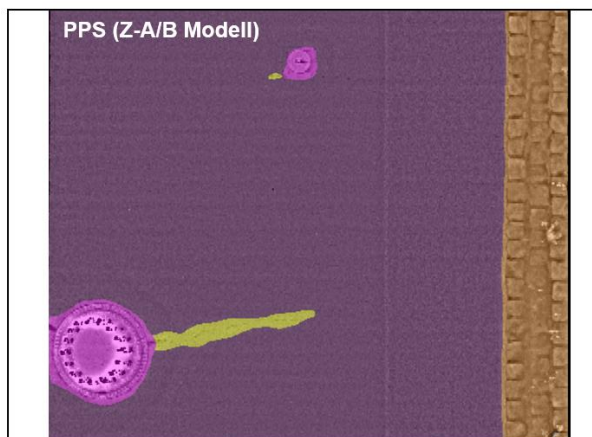
Asphalt offene Risse
Pflaster
Vegetation



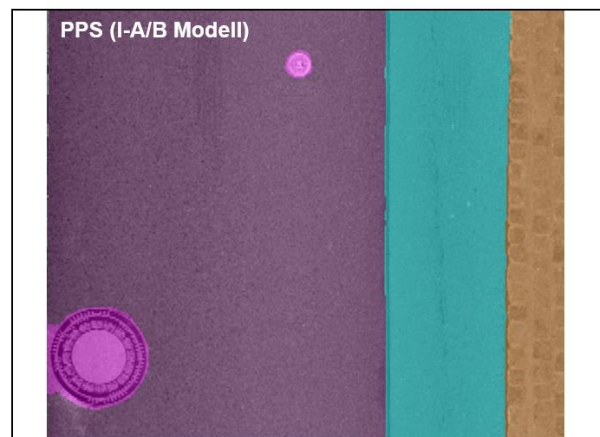
Asphalt sanierter Riss
Pflaster



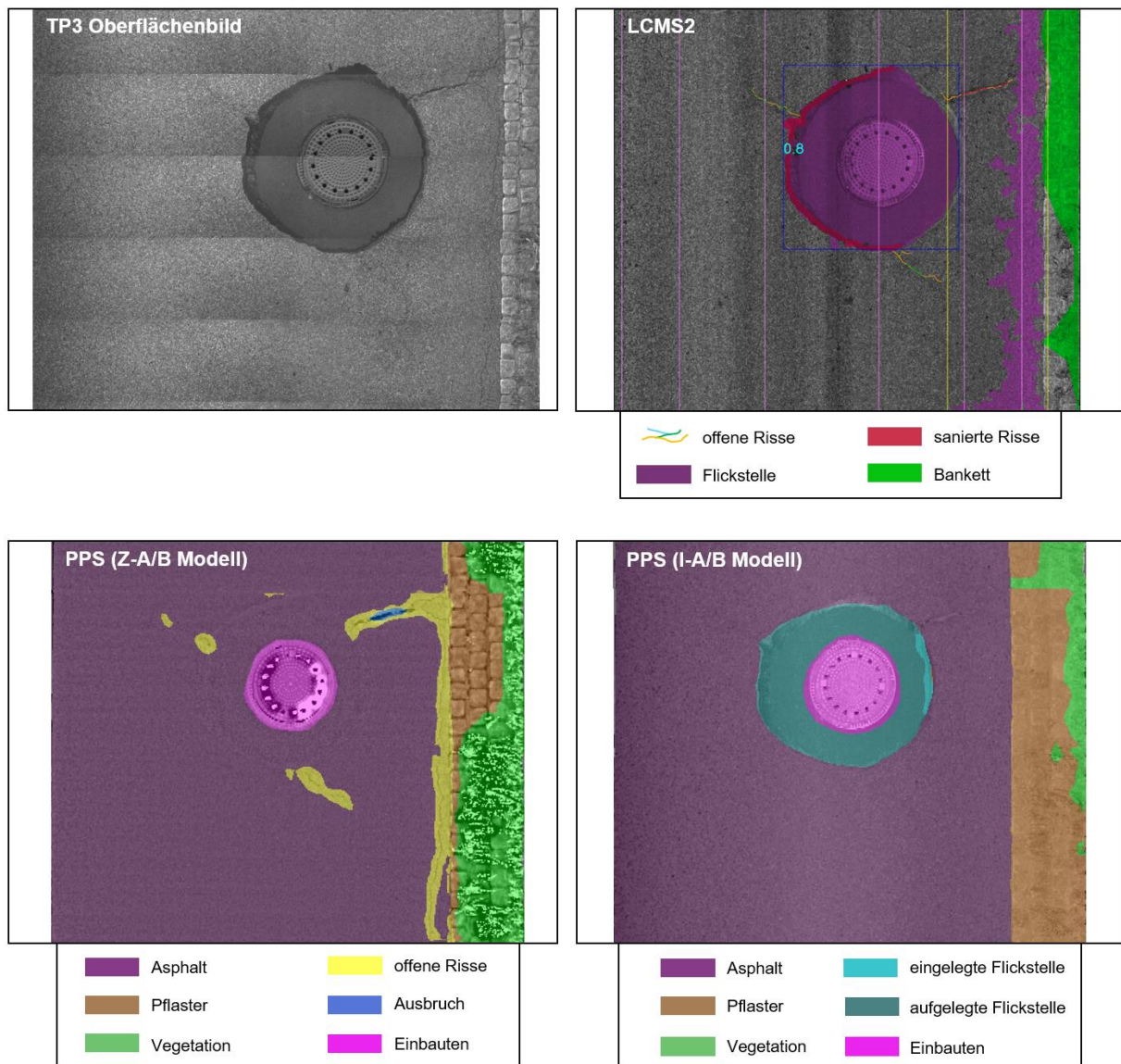
 offene Risse	 sanierte Risse
 Schachtdeckel	 Bankett/Bord

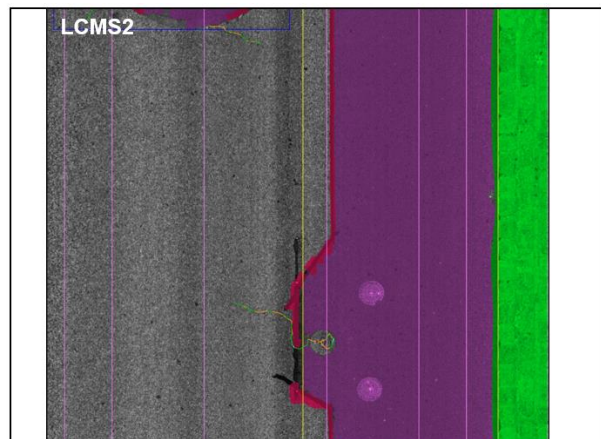
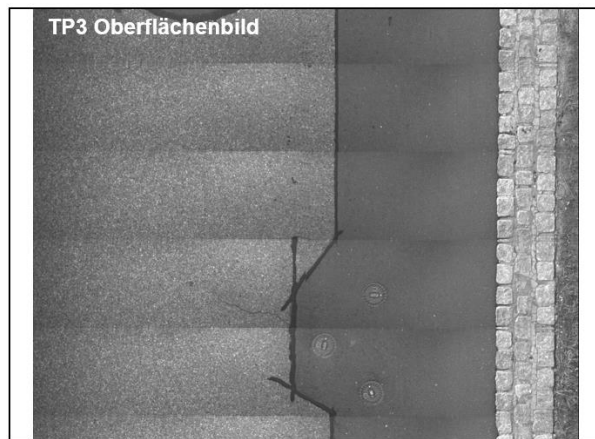


 Asphalt	 offene Risse
 Pflaster	 Einbauten

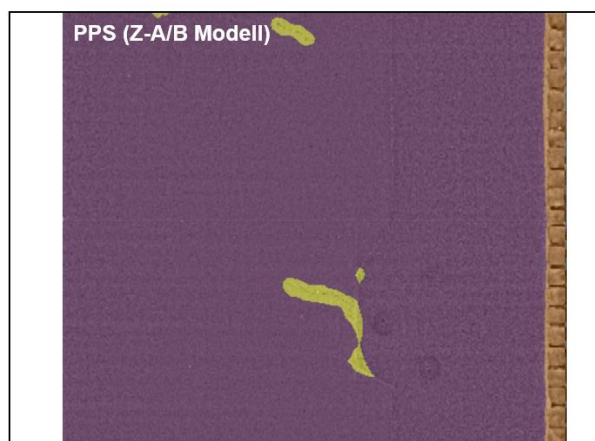


 Asphalt	 eingelegte Flickstelle
 Pflaster	 Einbauten

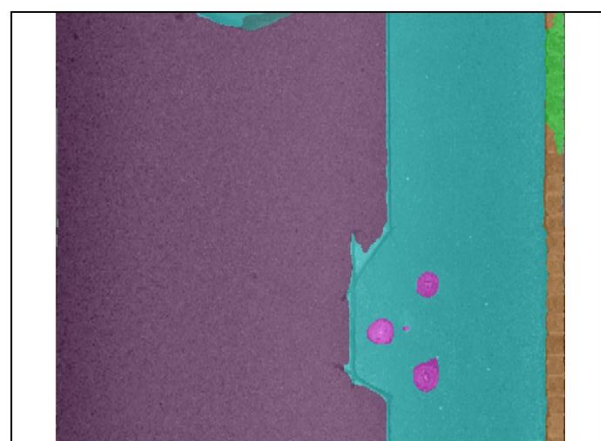




 offene Risse	 sanierte Risse
 Flickstelle	 Bankett



 Asphalt	 offene Risse
 Pflaster	 Ausbruch
 Vegetation	



 Asphalt	 eingelegte Flickstelle
 Pflaster	 aufgelegte Flickstelle
 Vegetation	 Einbauten

