

Verwendungsnachweis, Teil 2 Sachbericht

Zuwendungsempfänger: Hochschule für Technik Stuttgart

Förderkennzeichen: ZF4666502HB9

Bewilligungstermin: 28.06.2019

Projektlaufzeit: 06.05.2019 – 31.07.2021

1. Kurzbezeichnung des FuE-Projekts

(max. 250 Zeichen)

Entwicklung eines Mess- und Auswerteverfahrens zur automatisierten BIM-fähigen Objekterfassung in Tunnelbauwerken

2. Zusammenfassung des Projekts

(max. 1.200 Zeichen)

Im Rahmen des Projektes wurde ein Mess- und Auswerteverfahren zur präzisen Digitalisierung der Oberfläche eines Tunnels mit hohem Automationsgrad entwickelt. Das für die Bestandsaufnahme der Tunneloberfläche entwickelte System besteht aus Kameras, LED-Blitzleuchten, einer Synchronisationseinheit, mit der die Kameras untereinander und mit den LEDs synchronisiert werden, sowie Rechner und Datenspeicher. Für die Inspektion der Tunneloberfläche werden die Bilder und daraus generierte 3D Punktwolken analysiert, um verschiedene Arten von Verformungen, insbesondere Risse, Abplatzungen, Tiefensprünge und Rostfahnen sowohl geometrisch wie auch semantisch zu erfassen. Mit der automatisierten photogrammetrischen Auswertung des im Tunnel erfassten Bildverbandes (Bündelorientierung zur Georeferenzierung und Erzeugung dichter Punktwolken durch Bildzuordnung) konnte die angestrebte Koordinatengenauigkeit bezüglich Lage- und Tiefe von ca. 1 mm wie geplant erreicht werden. Lediglich Bereiche der Tunneloberflächen, die völlig texturlos waren, bleiben in den bildbasierten Verfahren unberücksichtigt. Die automatische Objekterfassung erkennt nicht nur Risse, Abplatzungen, etc. zur Schadensdokumentation über Deep-Learning-Methoden, sondern auch alle Ausstattungsfeatures, insbesondere Schilder, Markierungen, Fahrstreifensignale, Beleuchtungseinrichtungen, Hydranten, Lautsprecher, etc. mit hoher Qualität. Die konzipierte Erfassungsgeschwindigkeit von bis zu 65 km/h ist in den

Systemkomponenten (Belichtungszeit der Kameras, Pulsbreite des LED-Blitzes, Datenrate von Schnittstellen und Datenspeicher, etc.) vollumfänglich berücksichtigt. Der experimentelle Nachweis durch umfangreiche Tests konnte aber insbesondere aufgrund der Corona-Einschränkungen nicht im angestrebten Umfang durchgeführt werden. Die Untersuchungen mussten auf wenige Tunnelexperimente beschränkt werden.

3. Wurden die Arbeitspakete des Projekts wie in Anlage 5 des Antrags geplant bearbeitet?

(max. 2.000 Zeichen)

- ☐ Ja
- ☒ Nein

Wenn Arbeitspakete nicht wie geplant bearbeitet wurden – Erläuterung:

(Soll-Ist-Vergleich der Arbeitspakete bzgl. Personaleinsatz, kurze Begründung, Erläuterung zu Änderungen)

Es konnten letztlich alle im Antrag geplanten Arbeiten durchgeführt werden. Allerdings hat sich die Beschaffung von Komponenten der Messhardware als teilweise äußerst langwierig herausgestellt. Insbesondere die Anforderungen an die LED-Beleuchtung haben eine intensive Marktrecherche notwendig gemacht, die sich aufgrund der Corona-bedingt schlechten Erreichbarkeit der Lieferanten stark in die Länge zog. Die verzögerte Lieferung sorgte für eine Verschiebung der Messungen und somit des ganzen Projektes, da nachfolgende Arbeiten, wie die Erzeugung der Annotationsdaten für die Deep-Learning-Methoden, von der Bilderfassung vor Ort abhängig waren. Vom Projektträger wurde eine 3-monatige Verlängerung gewährt, so dass das Projekt trotzdem erfolgreich zu Ende geführt werden konnte.

4. Ergebnisse des Projektes einschließlich eines Vergleichs der angestrebten und erreichten technischen Parameter mit Bezug auf Meilensteine, technische Zielstellung, Begründung für Nichterreichung

(max. 50.000 Zeichen)

AP 1 Konzeptpräzisierung

Zu Beginn des Projekts wurden aktuelle Studien und Technologien recherchiert, um Informationen über die neuesten Systeme für die Tunnelinspektion bei hohen Geschwindigkeiten zu erhalten und die wichtigsten Komponenten und erforderlichen Konfigurationen zu ermitteln.

Bezüglich der Hardwarekomponenten wurden die zum Teil veröffentlichten Spezifikationen und verwendeten Systemkomponenten der beiden führenden kommerziellen Tunnelmesssysteme der Firmen Terra Vermessungen AG und Dibit Messtechnik GmbH eingehend analysiert. Die Analyse der am Markt verfügbaren Systemkomponenten haben dann zur finalen Spezifikation und den Entwurf des endgültigen Systems geführt. Eine besondere Herausforderung stellte die Synchronisierung der Machine Vision Kameras [2] und der LEDs dar. Die Lösung, mit den GPIO-Anschlüssen (General Purpose Input/Output, Abbildung 1) den Stroboskopausgang einer Kamera als Eingang für andere Kameras zu verwenden, hat sich am geeignetsten erwiesen. Die synchronisierte Erfassung benötigt (1.) die Verkabelung der Kameras und LEDs, mit Verbindungen zwischen den entsprechenden GPIO-Pins und (2.) die Konfiguration der Ein- und Ausgänge für jede Kamera mit der Kamerasoftware.



Diagram	Color	Pin	Function	Description
	Black	1	IO	Opto-isolated input (default Trigger in)
	White	2	O1	Opto-isolated output
	Red	3	IO2	Input/Output/serial transmit (TX)
	Green	4	IO3	Input/Output/serial receive (RX)
	Brown	5	GND	Ground for bi-directional IO, V _{EXT} , +3.3 V pins
	Blue	6	OPTO_GND	Ground for opto-isolated IO pins
	Orange	7	V _{EXT}	Allows the camera to be powered externally
	Yellow	8	+3.3 V	Power external circuitry up to 150 mA

Abbildung 1: GPIO-Anschlüsse für die Synchronisation von Kameras und LEDs

Überlegungen, auch den HFT-Laserscanner bei der Synchronisation zu berücksichtigen, wurden verworfen, als klar wurde, dass dies den Zeitrahmen des Projekts gesprengt hätte.

Bezüglich der Convolutional Neural Networks (CNNs) wurden bei der Präzisierung des Gesamtkonzepts eine Reihe vortrainierter CNNs wie Unet, Link-Net, Seg-Net, FCN, PSP-Net und Deeplab in Betracht gezogen, die für die Bildsegmentierung verwendet werden können. Zunächst wurden verschiedene Netze anhand eines frei verfügbaren Datensatzes von Rissbildern und annotierten Karten [1] trainiert. Wie schon im Zwischenbericht dargestellt, haben U-Net [4] und Deeplab [5] eine bessere Performance erzielt und wurden für dieses Projekt sowohl für die Schadenserkenkung als auch für die Objekterkennung ausgewählt.

AP 2 Entwicklung und Aufbau der Hardware

Die Anforderungen an das Kamerasystem sind entsprechend den Projektanforderungen spezifiziert und festgelegt worden (Tabelle 1). Für die Entwicklung und den Aufbau des Systems wurde auf die an der HFT Stuttgart vorhandenen Kameras zurückgegriffen und diese, soweit nötig und möglich, an das Projektdesign angepasst. Die Kameras verfügen über ein 6 mm Objektiv, einen flächenbasierten 5 MP CMOS Sensor und einen globalen Shutter, so dass Aufnahmen auch bei schneller Bewegung möglich sind. Das bei der Entwicklung verwendete Setup besteht aus 3 Kameras in ringförmiger Anordnung, wobei ein Winkelunterschied (α) zwischen den Blickrichtungen zweier benachbarter Kameras von 45 Grad gewählt wurde. Die Überlappung zwischen Aufnahmen zweier im Ring benachbarten Kameras (im Folgenden als vertikale Überlappung bezeichnet und auf den Öffnungswinkel bezogen) und die Überlappung zwischen zwei aufeinanderfolgend in Bewegungsrichtung erfassten Bildern (im Folgenden als horizontale Überlappung bezeichnet und auf das Bildformat bezogen) und werden durch die Gleichungen 1 und 2 bestimmt.

$$VerticalOverlap [\%] = \left(1 - \frac{\alpha[deg]}{FOV_w[deg]}\right) \times 100 \quad (1)$$

$$HorizontalOverlap [\%] = \left(1 - \frac{V[m/s] \times f[mm]}{FPS[Hz] \times h[mm] \times D[m]}\right) \times 100 \quad (2)$$

w und h sind die Breite und Höhe des Bildsensors, FOV_w und FOV_h die dazugehörigen Öffnungswinkel der Kamera, V ist die Geschwindigkeit des Fahrzeugs, f die Brennweite der Kamera, FPS die Anzahl der pro Sekunde aufgenommenen Bilder und D der Abstand zwischen Kamera und abgebildeter Oberfläche. (Anmerkung: Werden die Kameras beim Einbau um 90 Grad verdreht, gehen w und h vertauscht in Berechnung der Überlappung ein). Ausgehend von den Kameraparametern und Tunnelabmessungen in Tabelle 1 beträgt die vertikale Überlappung bei drei Kameras 35,7 % des entsprechenden Sichtfelds (Abbildung 2).

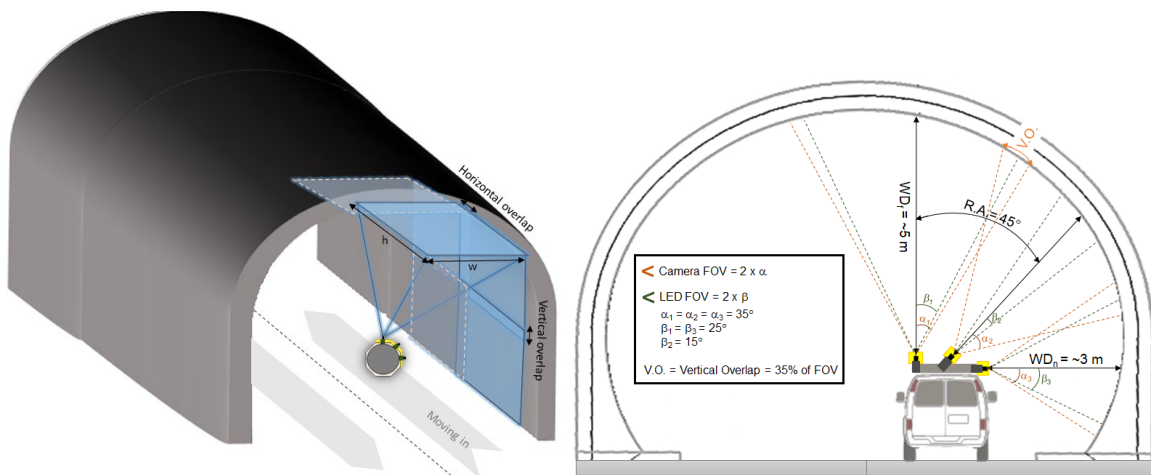


Abbildung 2: Aufbau des Multikamerasystems mit drei Kameras

Zur Beleuchtung im Tunnel wurden LEDs der Firmen Lumimax [6] und Metaphase [3] in die engere Wahl genommen. Da Dauerlicht LEDs bei kurzen Belichtungszeiten die Tunneloberfläche nicht genügend hell ausstrahlen, fiel schon bei der Konzeption die Wahl auf LED-Stroboskope (LED-Blitzleuchten), die sich mit den Kameras hochgenau synchronisieren lassen. Unsere Erwartungen an die LED-Balkenbeleuchtung von Lumimax (Abbildung 3) haben sich bei zunehmendem Objektstand (größer 1 m) nicht erfüllt and auch die Synchronisation zwischen LED und Kamera war problematisch (Abbildung 7,b).

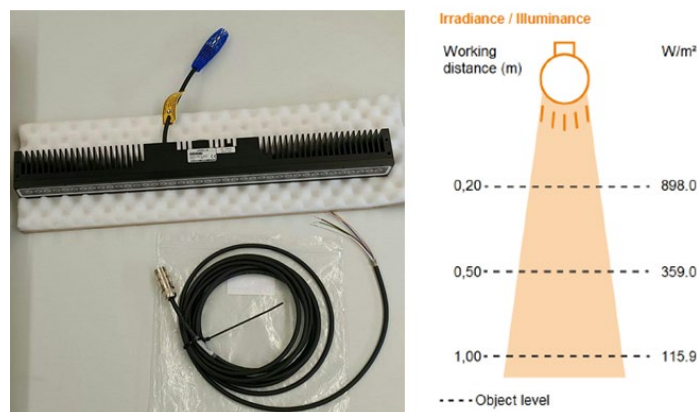


Abbildung 3: Lumimax LED Balkenleuchte

Da die Beleuchtung und Synchronisation mit den von der Firma Metaphase [3] angebotenen MetaBright LEDs erfolgreich war (Abbildung 7,c) wurden im Projekt entsprechende LEDs und Controller beschafft. Der Arbeitsabstand der vorgeschlagenen LEDs beträgt 3 – 8 m. Die LED kann mehr als 50 kHz Blitzlicht mit einer Impulsbreite im Bereich von 20 μ s bis 60 ms liefern. Je nach Entfernung bzw. beleuchteter Fläche können unterschiedlich große Linsen eingesetzt werden, um das Licht effizient zu projizieren und rechteckige Flächen perfekt zu beleuchten. Je größer der Öffnungswinkel der Linse, desto mehr Fläche wird bei dann abnehmender

Lichtintensität beleuchtet. In den Untersuchungen haben wir Linsen mit einem Winkel von 30° und 50° für größere bzw. kleinere Entfernungen verwendet (Tabelle 1). Abbildung 4 zeigt eine LED der Firma Metaphase und den dazugehörigen Universal LED Controller (ULC-2).

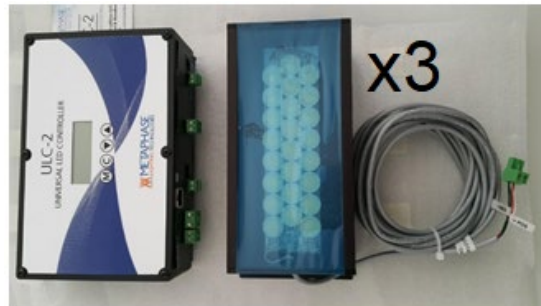


Abbildung 4: LED Strobe mit ULC-2 Controller der Fa Metaphase

Die Integration einer Thermografiekamera durch Synchronisation mit den Machine Vision (MV) Kameras war technisch nicht möglich, da keine Thermografiekamera am Markt ist, welche die Standards der MV-Kamerasynchronisation erfüllt. Wärmebildaufnahmen müssen daher durch eine getrennte Erfassung in den Projekten erfolgen. Da die primär bearbeiteten Schäden und Artefakte in RGB-Bildern sichtbar sind, bestand keine Notwendigkeit, die Thermografiekamera in die Untersuchungen einzubeziehen. Deren Einsatz bleibt Spezialanwendungen vorbehalten, bei denen Temperaturanomalien für die Interpretation relevant sind. Da LED-Beleuchtungen im für den Menschen unsichtbaren Nahen Infrarot erhältlich sind, wurden Untersuchungen mit einer NIR Kamera vorgenommen. In Abbildung 5 haben wir die Aufnahmen einer Canon Rebel T1i/EOS 500D (Bänder: G, R, NIR) und einer RGB-Kamera gegenübergestellt. Das NIR Band enthält Potential für die automatische Schadensdokumentation. Da aber die Ausstattungsfeatures nicht nur über die Geometrie, sondern auch über die Farbe erkannt werden, wurde eine Beschränkung auf NIR Aufnahmen nicht weiter verfolgt.



RGB Bild



GRNIR Bild

Abbildung 5: Visueller Vergleich von RGB- und VNIR-Aufnahmen

Um alle Komponenten in ein Gesamtsystem zu integrieren, haben wir eine kostengünstige Synchronisationseinheit (Abbildung 6, a) selbst entwickelt, die auf dem GPIO-Anschluss der Kameras basiert. Der Ausgangspin der "primären" Kamera ist hierbei dafür verantwortlich, den

Trigger-Strobe an die Eingangspins der "sekundären" Kameras sowie an die Eingangspins der LEDs zu senden. Für die Synchronisierung werden die entsprechenden Pins verbunden, wie Abbildung 6, b, c gezeigt. In Abbildung 7,c ist ein Bild dargestellt, das die erfolgreiche Synchronisation der MV-Kameras und LEDs deutlich macht.

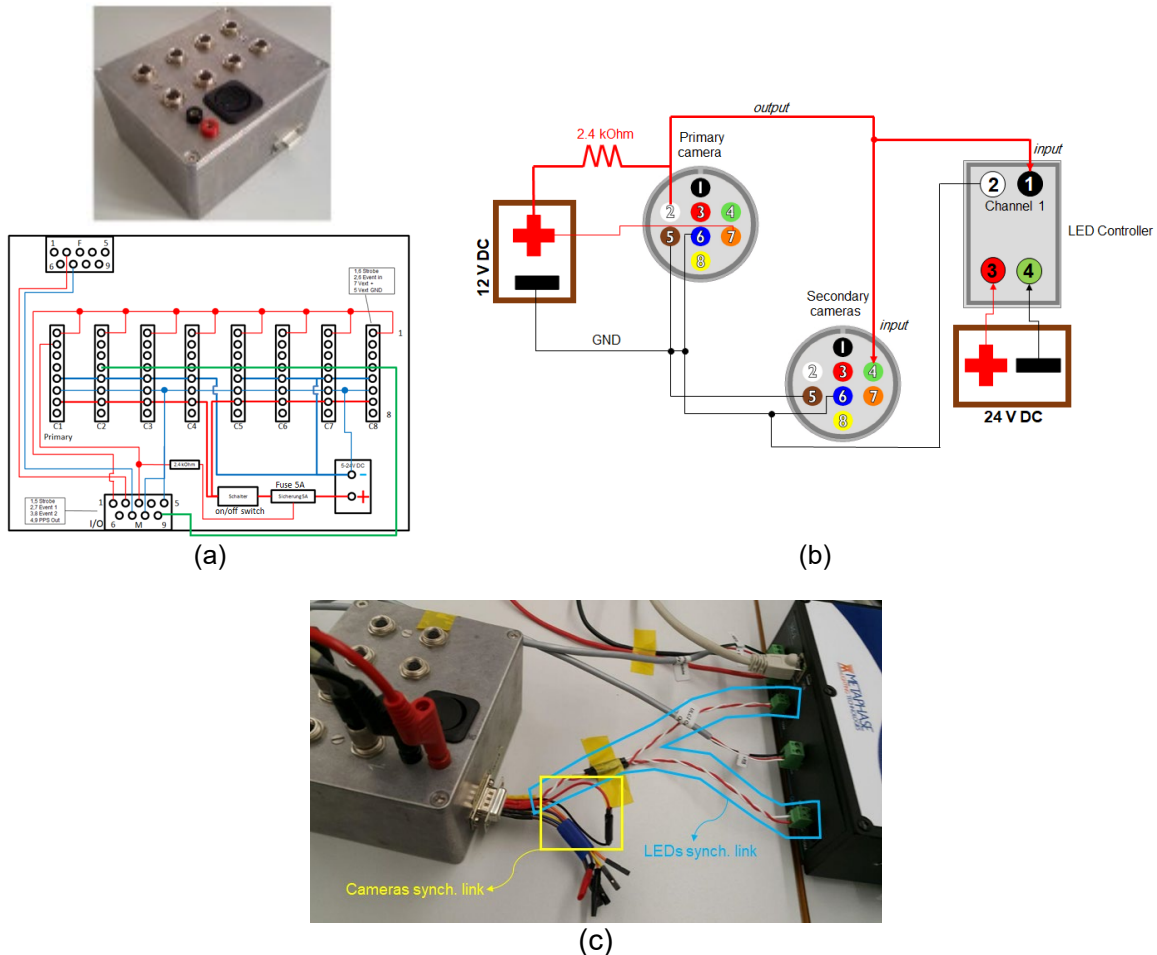


Abbildung 6: Anschluss der Komponenten an die Synchronisationseinheit

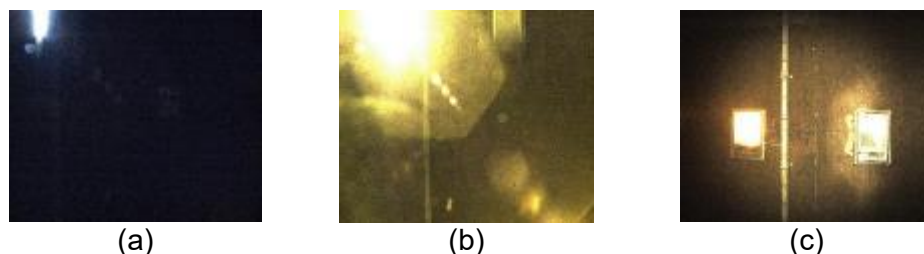


Abbildung 7: Synchronisierte Bildaufnahme (a) ohne Synchronisierung, (b) Lumimax-LED-Synchronisierung, (c) Metaphase-LED-Synchronisierung

Die Erfassung der synchronisiert aufgenommenen Bilder wird durch eine von FLIR entwickelte Open-Source-API-Bibliothek für Bildverarbeitungskameras gesteuert, die auf einem normalen Laptop installiert ist. Mit den richtigen Einstellungen für den Start- und Endzeitpunkt der

Erfassung, die Bildrate, das Bildformat und die Beleuchtungsstärke der Primär- und Sekundärkameras in der API (Abbildung 8) können die Daten für jede Fahrt automatisch auf der Festplatte gesammelt werden. Weiterführende Skripte werden nicht benötigt.

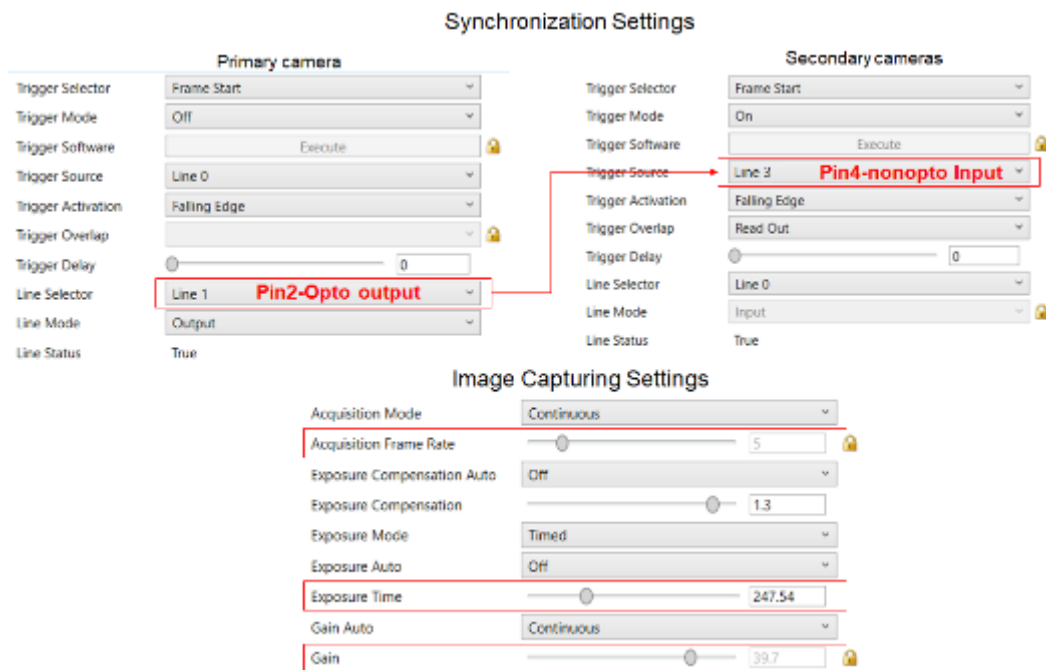
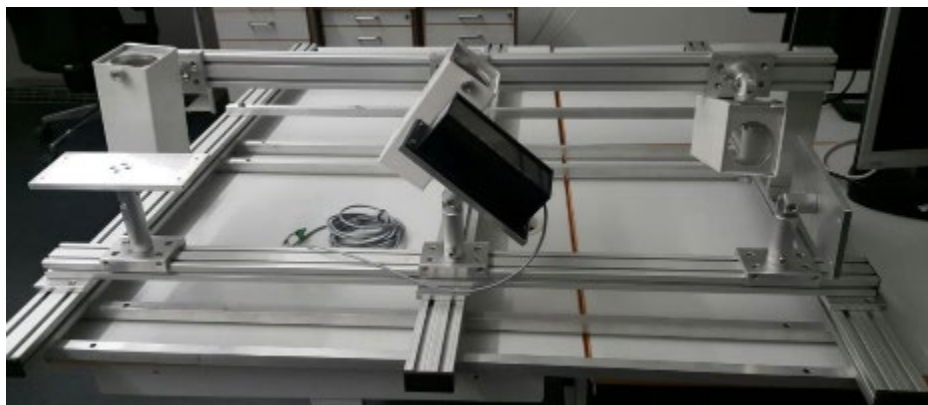


Abbildung 8: API für die synchronisierte Aufnahmen von Bildern

Zur Datenerfassung im Tunnel wurde eine Dachträgerplattform aus Aluminium hergestellt. Sie ist modular aufgebaut ist und ermöglicht den flexiblen Ein- und Ausbau von Kameras und LEDs. Lage und Ausrichtung der Kamera und LEDs bezogen auf die Erfassung eines Rings der abgebildeten Oberflächen können je nach Projektanforderung festgelegt und geändert werden, wie in Abbildung 9,a zu sehen ist.



(a)



(b)

Abbildung 9: Systemsetup mit 3 Kameras und LEDs: (a) Flexibilität beim Einbau der Kameras und LEDs auf der Trägerplattform (b) Befahrung eines Straßentunnels mit einem Fahrzeug der HFT

Abbildung 9,b wurde bei einer Testfahrt aufgenommen, in der die Aufnahmen in einem Straßentunnel in Stuttgart erfolgten. Die Kameras sind so ausgerichtet, dass sie die der Fahrtrichtung zugewandte Seite des Tunnels erfassen.

Bei der Entwicklung und dem Aufbau der Hardware wurde auf die an der HFT vorhandenen Kameras zurückgegriffen. Dadurch musste die Beschaffung von Hardware bei Viscan nicht abgewartet werden, zumal unter den Corona-Bedingungen der Zeitplan für diese Beschaffung sehr unsicher war. Auch wenn das Setup aus vorhandenen Kameras und beschafften LEDs, mit dem die Untersuchungen durchgeführt wurden, mit den Parametern des finalen Messsystems nicht vollumfänglich übereinstimmt (siehe Tabelle 1), konnte mit den in den Experimenten erprobten Systemkomponenten das Projekt erfolgreich bearbeitet und der Nachweis der Funktionstüchtigkeit des Gesamtsystems erbracht werden, wie im Folgenden noch deutlich wird.

Tabelle 1: Konzeption der Kamera- und LED-Parameter des Messsystems sowie das Setup aus vorhandenen Kameras und beschafften LEDs, mit dem die Untersuchungen durchgeführt wurden

Parameter	konzipiert	Im Experiment erprobt
Kamera	CMOS, Bereichsscan	CMOS, Bereichsscan
Auflösung (Anzahl Pixel)	> 12 MP	5 MP
Sensorformat	1.1"	1.1"
Pixelgröße [μm]	3.45	3.45
Brennweite (Objektiv)	12 mm	6 mm

FPS	> 30 Hz	2-4 Hz
Verschluss	Global	Global
Schnittstelle	USB 3.0	USB 3.0
Pixel-Bittiefe	10 – 12 bits	8 bits
Chroma	Color	Color and mono
Belichtungsdauer	5 μ s – 30 s	250 μ s
Anzahl Kameras	10	3
Kamera FOV _w x FOV _h	60° x 46°	70° x 60°
Anzahl LED-Leuchten	10	3
LED Beleuchtungswinkel	30°	30° and 50°
LED Pulsbreite	5 μ s – 30 s	250 μ s
LED Mode	flash	flash
Arbeitse Entfernung	4 - 5 m	3 - 8 m
Erfassungsgeschwindigkeit	60 - 65 km / h	36 km / h
Lage-/Tiefengenaugigkeit	1 mm	1 mm (relativ) < 10 cm (absolut)

AP 3 geometrische Kalibrierung des Multikamerasystems

Ein wichtiger Schritt bei der Aufnahme von Bildern ist die Kamerakalibrierung, um die intrinsischen, extrinsischen und Objektivverzeichnungsparameter der Kameras zu bestimmen. Für ein Mehrkamerasystem sind sowohl Einzelkopf- als auch Mehrkopfkalibrierungen wichtig. Der Workflow für die Kamerakalibrierung wurde im HFT-Labor mithilfe der MATLAB Calibrator-App und eines Schachbrettmusters definiert und kann für das endgültige System angepasst werden. Die Kalibrierung im Labor hat die Anlage eines Testfeldes und die Einmessung von Testfeldkoordinaten und anderen Referenzdaten sehr vereinfacht. Bei Multikopf-Kalibrierungen wird eine Kamera als Primärkamera betrachtet und zwei andere Kameras (Sekundärkameras) werden in Bezug auf die Primärkamera kalibriert. Die Kalibrierungsparameter sind in Tabelle 2 aufgeführt.

Tabelle 2: Ergebnis der kalibrierten intrinsischen Parameter der in den Untersuchungen verwendeten Kameras

Calibration Parameters	Primary Camera	Secondary Cameras
radial distortions (k1, k2, k3)	-0.142, 0.103, -0.018	-0.131, 0.041, 0.065
tangential distortions (p1, p2)	-0.001, 0.0004	-0.001, 0.0002
Skew (b)	-0.993	-1.194
focal length (f) [pix]	1804.1	1792.4
principal points (xc, yc) [mm]	4.233, 3.489	4.297, 3.443
Mean Reprojection Error [pix]	0.33	0.31

Die extrinsischen Parameter der Kameras werden auch als Parameter der äußeren Orientierung (EO) bezeichnet. Sie werden im Rahmen der Kalibrierung in einem lokalen Koordinatensystem bestimmt und zu Näherungswerten für die Bündelausgleichungsverfahren bei der Zuordnung der Bilder in den Bildblöcken aufbereitet. Die hierfür notwendigen Einzelschritte sind in Abbildung 10,b aufgeführt. Es hat sich gezeigt, dass die EO-Parameter auf der Grundlage der Tunnelabmessungen, der Trägerplattform sowie der Installationswinkel der Kamera in Bezug auf lokale Koordinatensysteme (Abbildung 10,a) hinreichend genau ermittelt werden konnten.

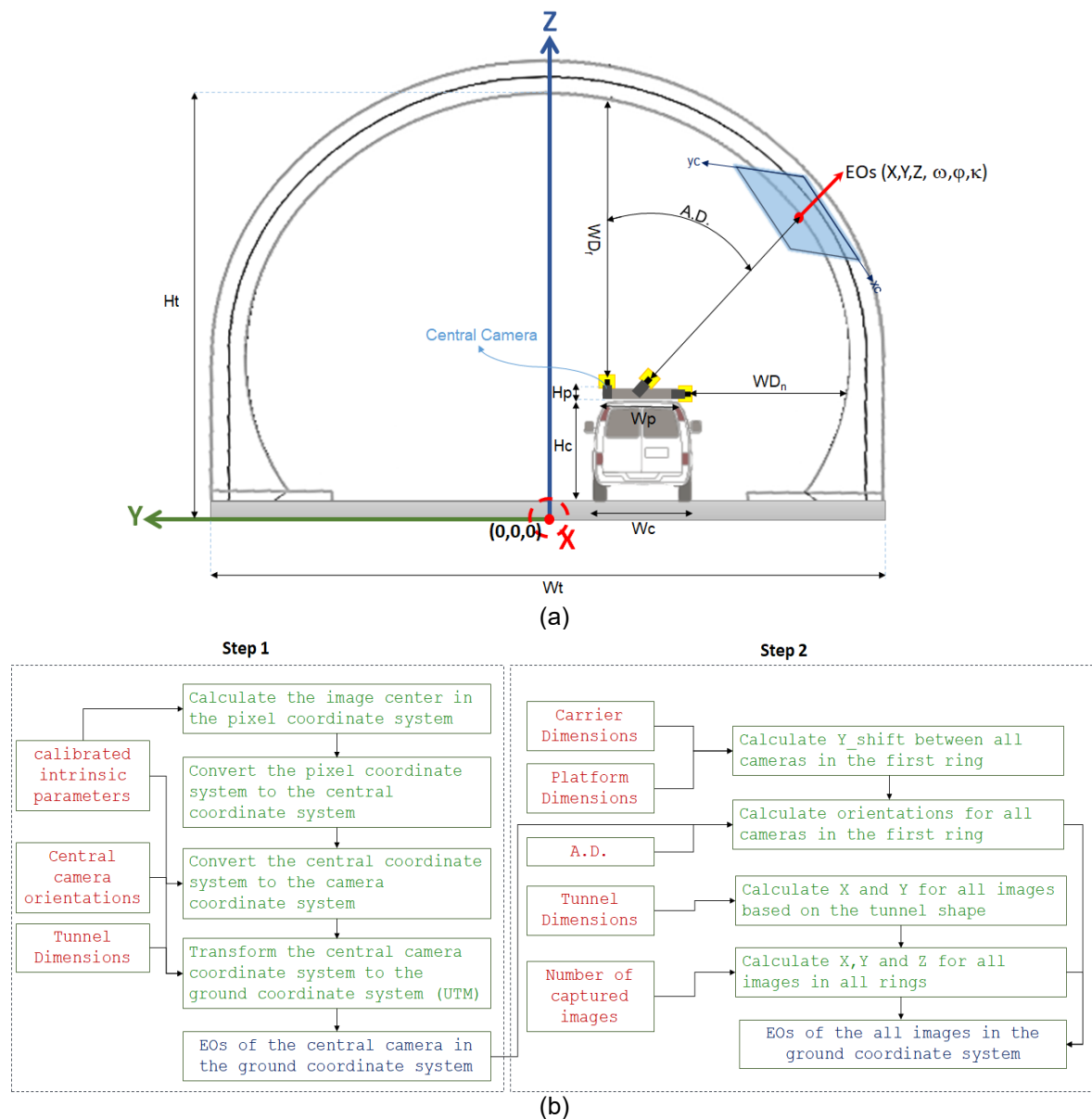


Abbildung 10: Bestimmung der Kamerapositionen und -ausrichtungen in Bezug auf das Referenzkoordinatensystem der Trägerplattform (a) Prinzipskizze, (b) Prozessierschritte

Abbildung 11 zeigt die als Näherungswerte für die Bündelausgleichungsverfahren ermittelten EO-Parameter aller aufgenommenen Bilder am Beispiel eines geraden und eines gekrümmten Tunnels. In Abbildung 12 ist das erste Experiment mit einem Testblock mit 1600 Bildern mit und ohne geschätzte Näherungswerte der EO-Parameter dargestellt. Die Bündelausgleichungsverfahren und damit verbundene Punktwolkenerzeugung aus Verknüpfungspunkten der Bildzuordnung ist mit den genähert ermittelten EO-Parametern effizienter und erfolgreicher.

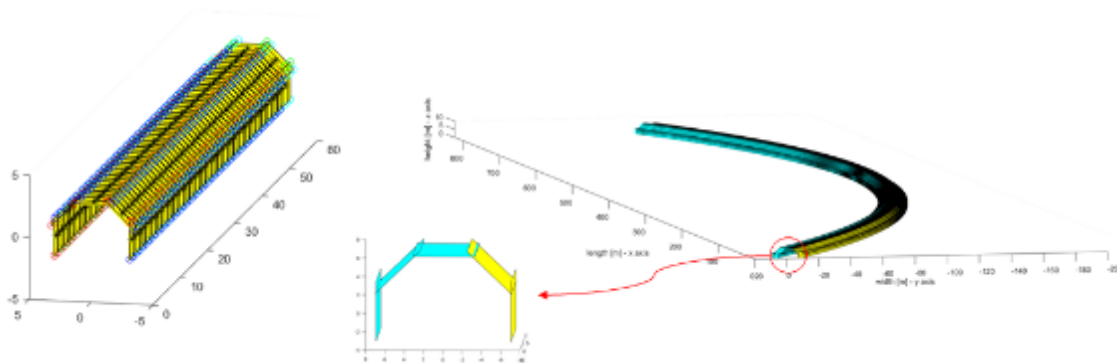


Abbildung 11: Genähert ermittelte EO-Parameter für gerade und gekrümmte Tunnels

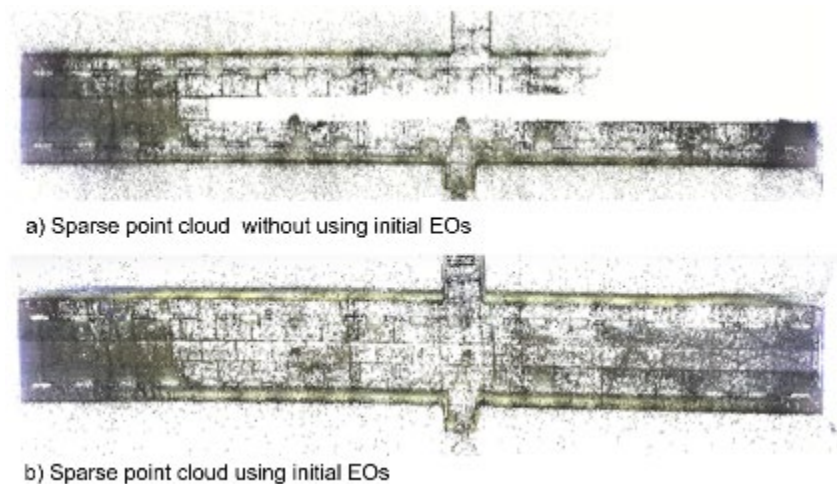


Abbildung 12: Punktwolkenerzeugung mit und ohne a priori genähert ermittelte EO-Parameter

AP 4 Entwicklung Prozessfluss Photogrammetrie

Um die synchronisierte Erfassung von Bildern mit dem entwickelten System sowie den implementierten Prozessflusses mit der Orientierung der simultan erfassten Bilder bis hin zur Erzeugung von Punktwolken zu erproben, wurde eine Unterführung in der Nähe des Feuerbacher Hauptbahnhofes gewählt. Breite und Höhe des Tunnels betragen ca. 6,2 m bzw.

2,8 m. Die Gesamtlänge des Tunnels beträgt ca. 57,2 m. Bei diesem Test wurden 1620 RGB-Bilder im .jpg-Format aufgenommen und auf einem Laptop für die Punktwolkenerzeugung gespeichert. Zur Beurteilung der Genauigkeit und zur Bewertung des gesamten Prozesses werden die Referenzdaten mit dem Faro Focus3D X-330 Laserscanner erfasst. Die Ergebnisse der Bündelorientierung und Punktwolkenerzeugung mit Agisoft Metashape [7] sind in Abbildung 13 dargestellt. In diesem Szenario werden während der Bildorientierung etwa 655.000 Verknüpfungspunkte erzeugt; der RMS-Rückprojektionsfehler beträgt 0,21 Pixel. Bei maximal 6 m Entfernung bis zur Tunneldecke entspricht die Pixelgröße ca. 3,5 mm, so dass eine nachbarschaftliche Genauigkeit der Punkte von 0,7 mm erzielt wurde, die in Tabelle 1 als relative Lage- und Tiefengenauigkeit (< 1 mm, relativ) aufgeführt wurde.

Darüber hinaus werden bei der Erstellung der dichten Punktwolke durch Semi-global Matching etwa 3.000.000 Punkte einschließlich der Farbinformationen rekonstruiert. Wie in Abbildung 14 dargestellt, betragen die Punktabstände zwischen der mit Bildern erzeugten Punktwolke und den Referenzdaten (Faro-Punktwolke) etwa 0,048 m bzw. 0,044 m für den Mittelwert und die Standardabweichung der Abstände. Die Einfärbung der Abweichungen macht deutlich, dass es Bereiche gibt, in denen größere Abweichungen auftraten, die große Masse der Punkte aber nur im mm-Bereich voneinander abwich.

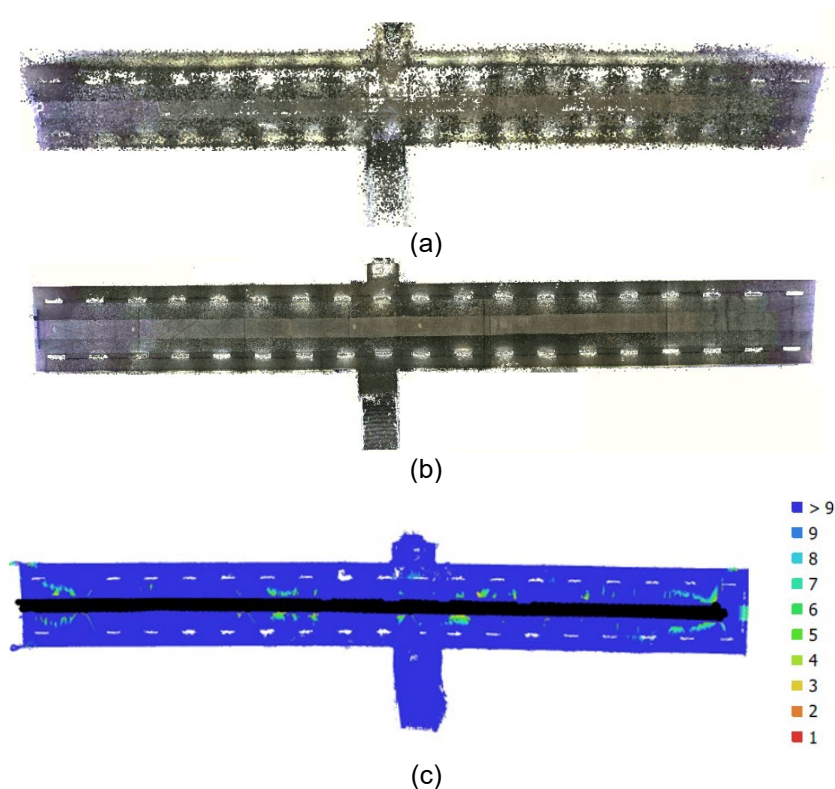


Abbildung 13: Bündelorientierung und Punktwolkenerzeugung mit Agisoft Metashape
(a) dünne Punktwolke, (b) dichte Punktwolke, (c) Überlappungskarte (Verknüpfungspunkte)

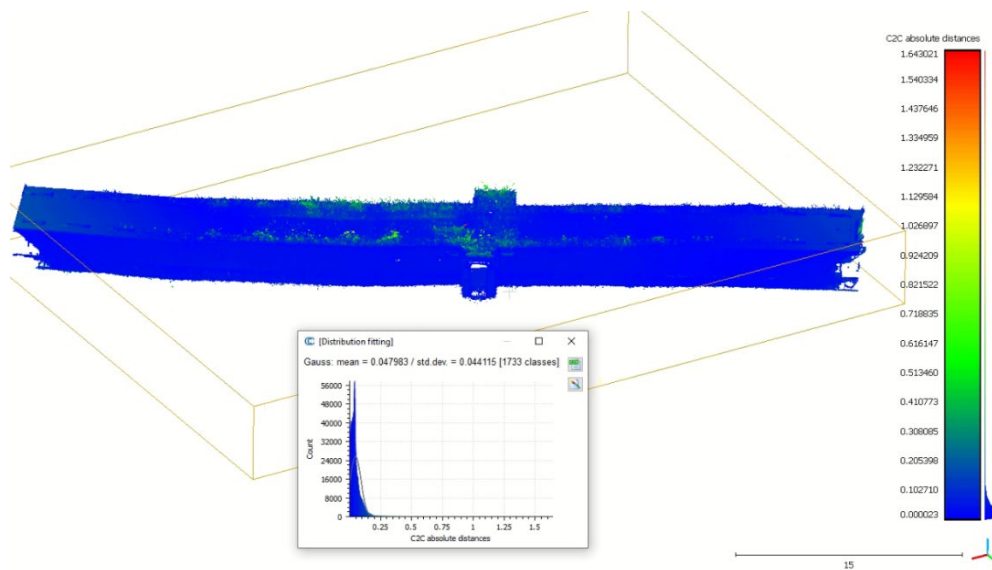


Abbildung 14: Abweichung der photogrammetrischen Punktwolke von der durch Laserscanning erfassten Referenzpunktwolke

Ein weiteres Experiment besteht darin, die Bilder eines langen Tunnels zu verarbeiten, indem der gesamte Bildblock in mehrere sich überschneidende Bildblöcke unterteilt wird, für jeden Block eine Punktwolke erzeugt wird und die Punktwolken dann mithilfe einer punktbasierten Registrierungstechnik zusammengeführt werden. Das Ergebnis dieses Experiments ist in Abbildung 15 dargestellt. Die mittlere Distanz und die Standardabweichung der Punktabstände zwischen der fusionierten Punktwolke und der ursprünglichen Punktwolke betragen 0,07 m bzw. 0,08 m. Die Zusammenführung der Punktwolken der drei Blöcke funktionierte gut, solange die Überlappung zwischen den Blöcken nicht unter 20% der Länge eines der Blöcke fiel.

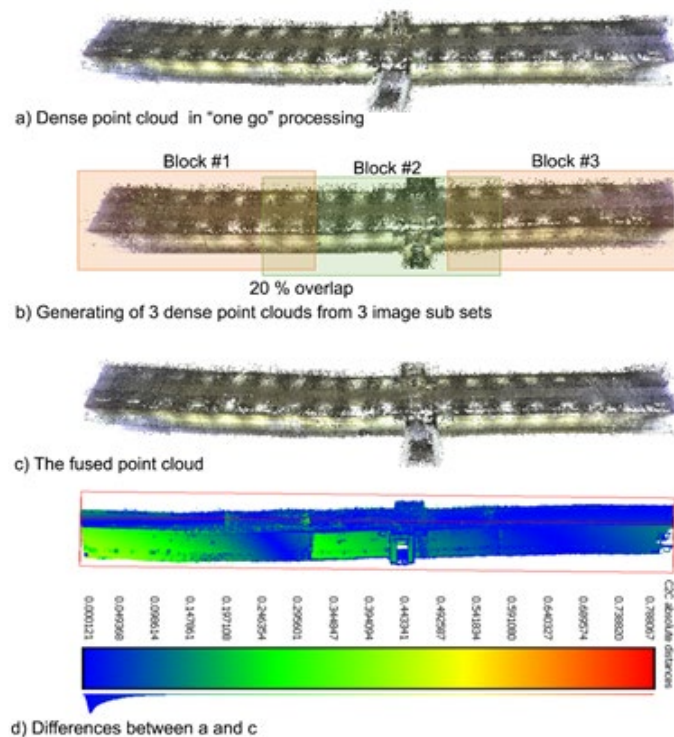


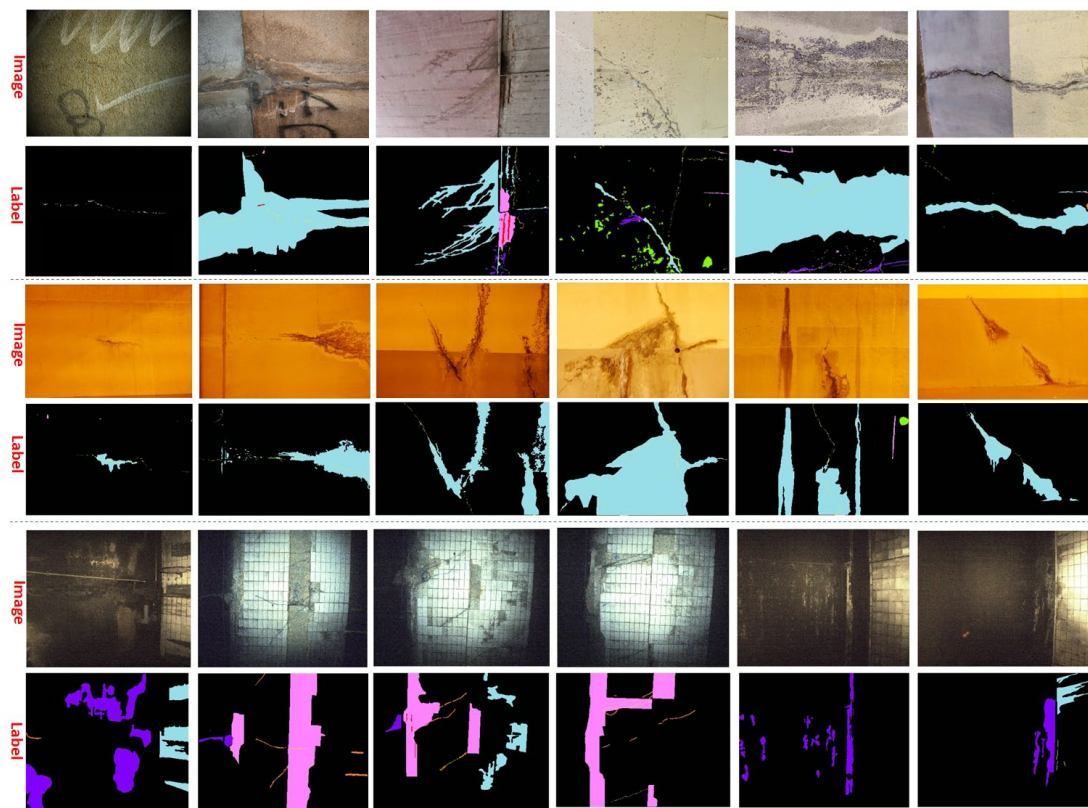
Abbildung 15: Blockweise Bündelorientierung und Punktwolkenerzeugung

AP 6 Entwicklung neuronaler Deep Learning Netze für die Objekterkennung in Tunnelbauwerken

Verschiedene neuronale Netze wie Deeplab und Unet wurden sowohl für die Schadens- als auch für die Objekterkennung entwickelt. Das vortrainierte Netzwerk für beide Netzwerke ist die Bilddatenbank ImageNet mit dem Backbone ResNet 50. Auf dieser Grundlage werden Netze mit unterschiedlichen Einstellungen und Hyperparametern anhand zu erstellender Trainingsdatensätze trainiert. Die Trainingsdaten umfassen Aufnahmen von Tunneloberflächen, die sowohl mit MV-Kameras als auch mit DSLR-Kameras aufgenommen wurden. Der erste Schritt für die Durchführung der semantischen Segmentierung besteht darin, die Bilder zu labeln, so dass neben den Originalaufnahmen auch gelabelte Bild für das Training vorhanden sind.

Für jede der beiden Aufgaben - Schadensdokumentation von Rissen, Abplatzungen, etc. und Erkennung der Ausstattungsfeatures des Tunnels – wurden über einen Zeitraum von 6 Monaten etwa 500 Bilder gelabelt, wie in Abbildung 16 dargestellt. Für die Schadenserkennung wird eine Bibliothek mit 9 Klassen verschiedener Schadensarten (Riss, Abplatzung, Rost, Rostfahne, Delamination, Bewehrungsstahl, Ausblühungen, Moos, Oberflächenveränderung) angelegt. Bezüglich der Erkennung der Ausstattungsfeatures werden 10 Klassen verschiedener Objekte (Verkehrsschilder, Leuchte, Kamera, Axialventilator, Ampel,

Kabel, Alarmlautsprecher, Entlüftung, Elektronikbox, Kabelbox) definiert. Die Bibliothek, in der die Originalbilder und die gelabelten Bilder gespeichert sind, stellen eine wichtige Ressource für Weiterentwicklungen nach Ablauf des Projekts dar. Sollen weitere Schäden oder Objekte den Deep Learning Netzen hinzugefügt werden, sind lediglich neue Trainingsläufe mit entsprechenden Validierungen durchzuführen.



(a)

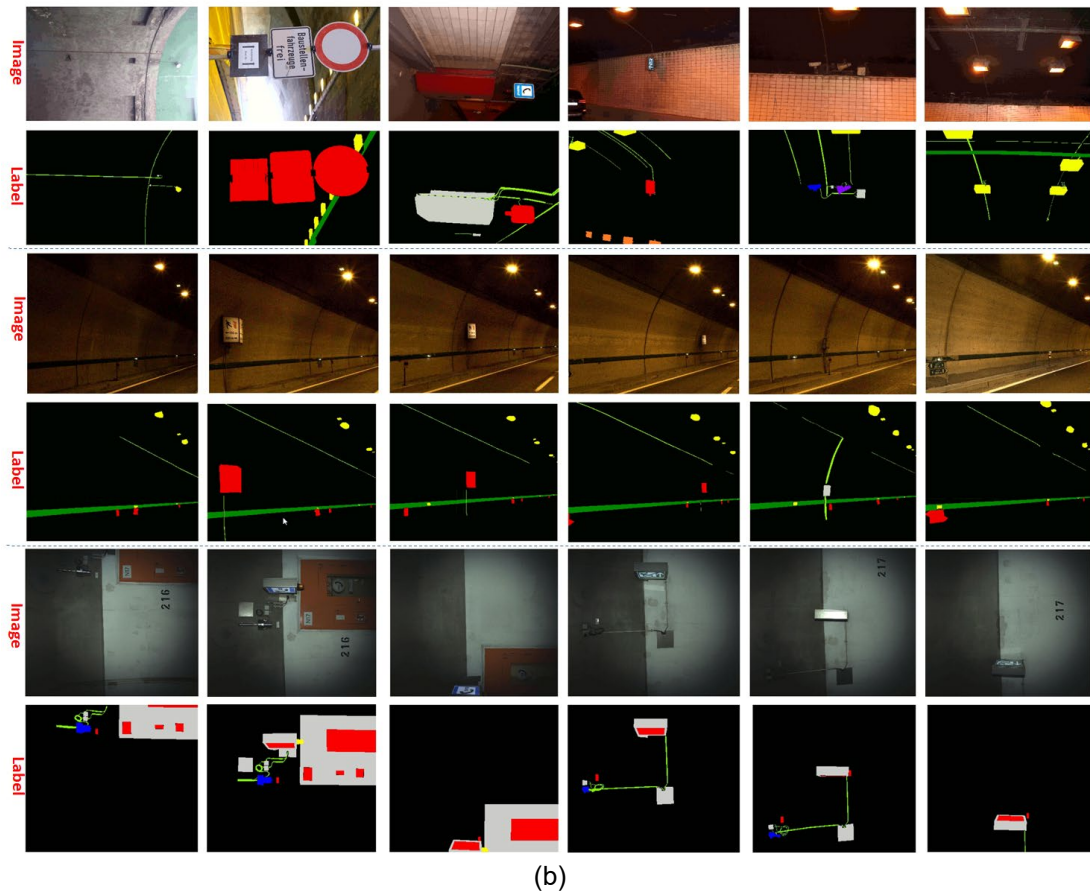


Abbildung 16: Beispielbilder und entsprechende Labels für (a) Schäden und (b) Objekte in den verschiedenen Trainingsdatensätzen

Um einen möglichst umfassenden Einblick in die Leistungsfähigkeit der Deeplab- und Unet-Netzwerke für die Schadenserkennung zu bekommen, werden zwei verschiedene Trainingsdatensätzen trainiert und anschließend anhand von Testdaten bewertet. Die Schadenserkennung stellte sich als anspruchsvoll dar, so dass weitere Analysen zur Verbesserung unternommen wurden. Es hatte sich herausgestellt, dass die Schadensdefinitionen in einigen Bildern nicht eindeutig war und die Genauigkeit des Labelings verbessert werden musste. Als weiteren Versuch zur Verbesserung der Erkennungsrate werden die beiden CNNs Unet++ [8] und SA-Unet [9] untersucht. Die Ergebnisse der Schadenserkennung sind in Tabelle 3 auf der Grundlage der Intersection of Union (IoU) für verschiedene Szenarien zusammengefasst.

Tabelle 3: Genauigkeitsbewertungen verschiedener CNNs zur Schadenserkennung

Datas et	Number of raw data	Number of training data	Number of augmen ted training data	Number of test Images	IoU_ave Deeplab	IoU_ave Unet	IoU_ave Unet++	IoU_ave SAUnet
1	490	450	16354	40	32.6	34.5	28.6	31.6
2	220	200	23751	20	21.2	24.8	24.5	25.5

Die durchschnittlichen Erkennungsraten auf Pixelniveau liegen für alle Arten von Schäden bei etwa 34,5 % bzw. 25,5 % für Bilder, die mit DSLR- bzw. MV-Kameras aufgenommen wurden. Die prädizierten Label sind in Abbildung 17 visualisiert.

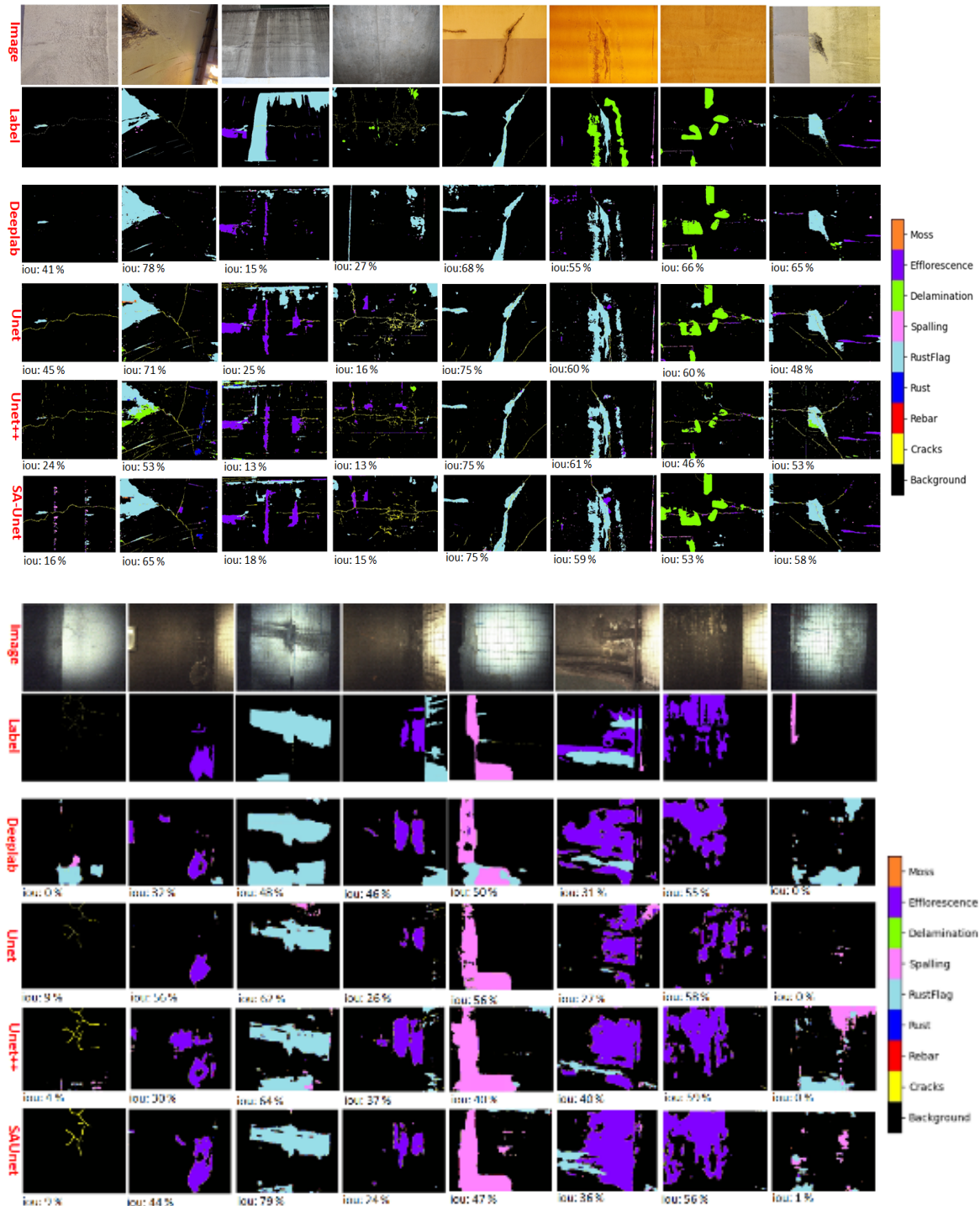


Abbildung 17: Ergebnisse von Deep Learning zur Schadenserkennung in verschiedenen Testdatensätzen

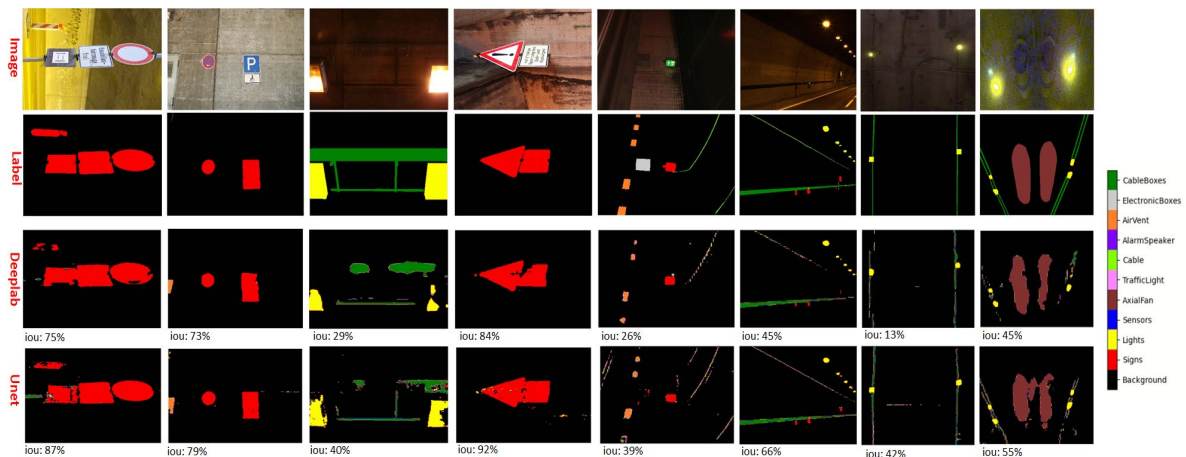
Für die Aufgabe der Objekterkennung werden Deeplab- und Unet-Netzwerke mit 8 verschiedenen Trainingsdatensätzen trainiert und anschließend mit Testdaten evaluiert. Die

Bilder wurden überwiegend in zwei Stuttgarter Tunnel aufgenommen, dem Wagenburgtunnel und dem Heslacher Tunnel. Die Techniken der Datenerweiterung (Augmentation) wie Spiegeln, Drehen und Skalieren werden verwendet, um den Datensatz zu vergrößern und die Erkennungsrate zu verbessern. In Tabelle 4 sind die Ergebnisse der Objekterkennung auf der Grundlage der Intersection of Union (IoU) für verschiedene Szenarien zusammengestellt.

Tabelle 4: Genauigkeitsbewertungen verschiedener CNNs zur Objekterfassung

Dataset	Number of raw data	Number of training data	Number of augmented training data	Number of test Images	IoU_ave Deeplab	IoU_ave Unet
1	98	88	10120	10	63.1	67.1
2	100	90	10350	10	69.9	64.9
3	198	178	10146	20	64.0	66.5
4	298	268	10452	30	60.8	66.4
5	379	340	11900	39	73.9	77.0
6	379	340	11900	39	74.4	75.7
7	577	518	36700	59	69.7	73.3
8	522	473	33583	49	54.6	78.9

Die besten Ergebnisse bei der Objekterkennung erzielte Unet mit einer Erkennungsrate von 78,9 %, womit ein sehr hohes Niveau erreicht wurde. Die prädizierten Label sind in Abbildung 18 dargestellt.



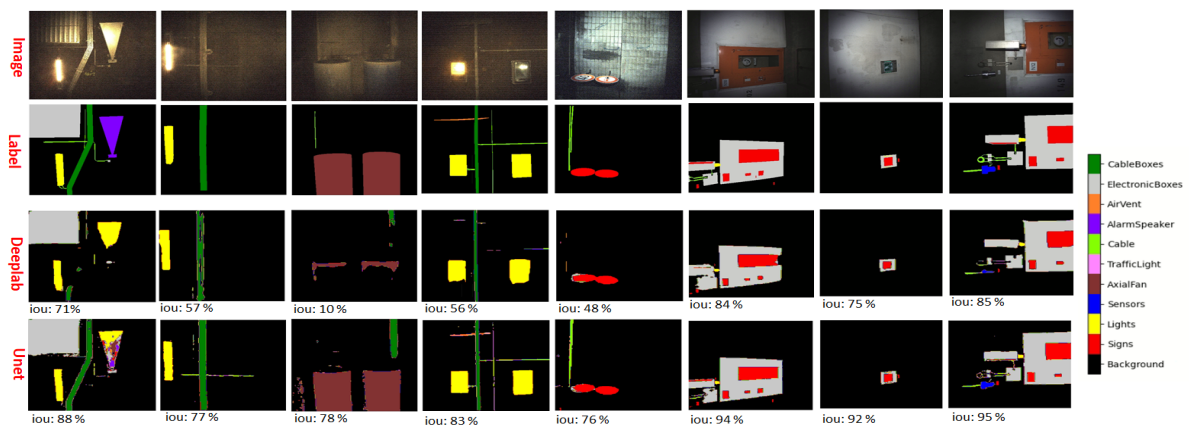


Abbildung 18: Ergebnisse von Deep Learning zur Objekterkennung in verschiedenen Testdatensätze

AP 7 Überführung von Bilddaten in 3D-Punktwolken

Mit dem entwickelten System werden die Bilder des gesamten Tunnelbereichs in Farbe erfasst. In Tabelle 1 sind die Parameter des zur Entwicklung und Erprobung gewählten Setups mit drei 5 MP Kameras dargestellt. Die nominale Bodenpixelauflösung (GSD) beträgt dabei 1,7 bis 2,9 mm bei einem Arbeitsabstand von 3 bis 5 m.

Der wichtigste Parameter, der sich direkt auf die Bildqualität auswirkt, ist der Unterschied zwischen dem Sichtfeld (Öffnungswinkel FOV) der Kamera und dem der LED. Der Öffnungswinkel der Kamera beträgt etwa 70° horizontal und 60° vertikal. Für die experimentelle Erprobung wurden zwei LEDs mit Linsen von 50° FOV und eine dritte LED mit 30° FOV eingesetzt. Aufgrund der kleineren LED FOVs sind die Ecken der Bilder, wie in Abbildung 19 zu sehen ist, abgedunkelt.

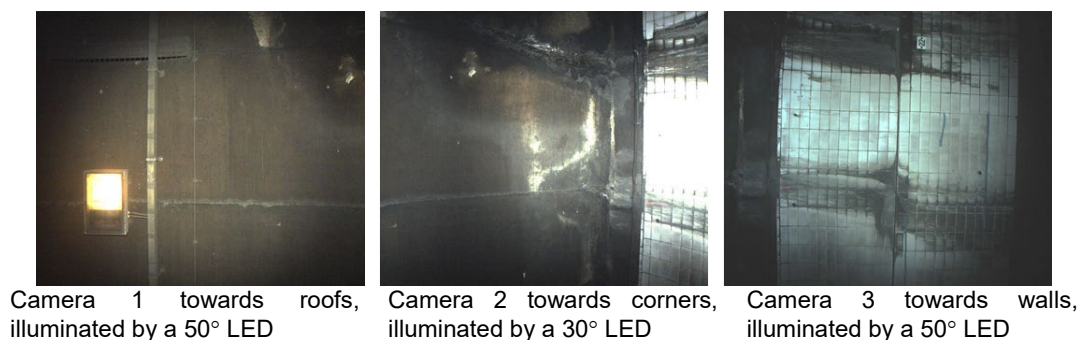


Abbildung 19: Qualität der aufgenommenen Bilder aus dem Wagenburgtunnel

Die USB-3 Schnittstelle eignet sich mit einem Datenstrom von bis zu 5 Gbit/s (in den Versionen 3.1 und 3.2 bis zu 10 bzw. 20 Gbit/s) für die schnelle Speicherung hochauflösender Bilder. Die im gesamten Setup erreichbare Datenrate hängt jedoch nicht nur von der Kamera-Schnittstelle

ab, sondern auch von der USB-Host-Controller-Karte, der Bildspeicherdatenrate der Festplatte und anderen Faktoren ab. Im Idealfall wird jede Kamera mit einem USB 3-Bus angeschlossen, um eine Vollbildrate zu erreichen. In diesem Projekt haben wir einen normalen Laptop mit einem USB 3.0-Anschluss für drei Kameras verwendet, wodurch die tatsächliche Bildrate für Farbbildern auf 2 Hz reduziert wurde. Werden mehr als drei Kameras im Ring einbezogen, verringert sich die Rate entsprechend. Zudem ist die Fahrgeschwindigkeit eingeschränkt, wenn eine entsprechende Überlappung der Bilder in Fahrrichtung vorgegeben ist. Konsequenzen einer niedrigen Bildrate, aber auch einer zu geringen Textur in bestimmten Tunnelabschnitten sind in Abbildung 20 zu sehen. Die Überlappung der Bilder in Ringrichtung von etwa 20 % (Abbildung 20,a) und die texturarmen Bildbereiche mit entsprechend wenig Möglichkeiten zur Bildzuordnung in den Überlappungsbereichen führten dazu, dass die Bündelorientierung und die Erzeugung von Punktwolken nicht erfolgreich waren, wie in Abbildung 20,b zu sehen ist. Die Bilder sind dennoch wichtige Ressourcen für die Tunnelinspektion, da sie für die Schadenserkennung von Rissen, Abplatzungen u.a. verwendet werden.

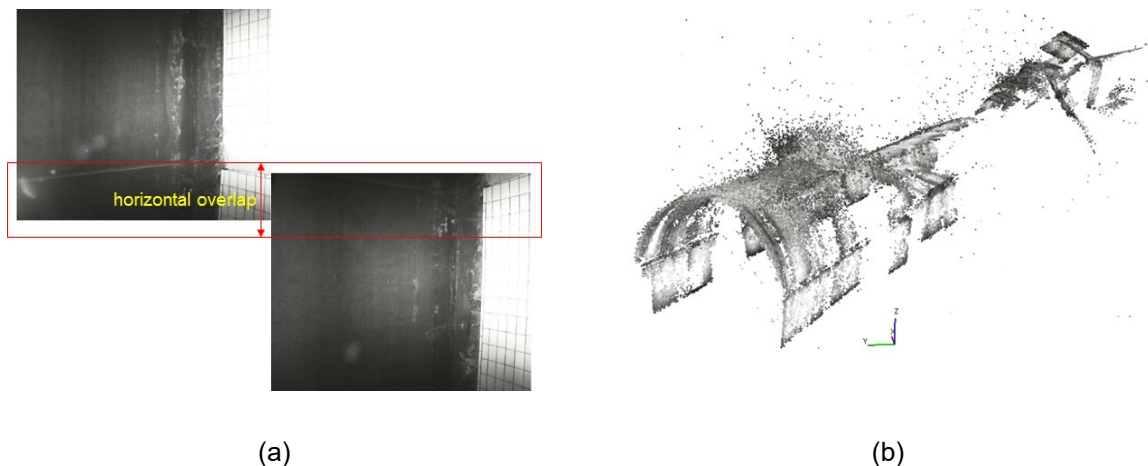


Abbildung 20: (a) Überlappung im Ring, (b) unvollständig rekonstruierte Punktwolke

AP 8 Verbesserung des neuronalen Netzes

Zur Verbesserung der Erkennungsrate bei der Schadens- und Objekterkennung wurden verschiedene Szenarien und Netzwerke untersucht. L2-Regularisierungs- und Klassengewichtungstechniken sowie die Augmentation haben dazu beigetragen, die Probleme der Überanpassung zu verringern. Hinzufügen des „räumlichen Aufmerksamkeitsblocks“ (spatial attention block) zu Unet hat die Detektionsergebnisse verbessert, wie schon in Tabelle 3 und Tabelle 4 und in den Erläuterungen zu Arbeitspaket AP 6 aufgezeigt wurde. Zu Beginn des Projekts wurden für die Schadenserkennung nur die 3 Klassen Riss, Abplatzung und

Rostfahne berücksichtigt. Da die erfassten Bilddaten mehr Schadenstypen enthielten, konnte das Training der Netze auf 9 Klassen erweitert werden. Aufgrund der bereits erwähnten Probleme mit der Schadensdefinitionen, wurde das Labeling verbessert und die Netze mit diesen Trainingsdaten von Grund auf neu trainiert, um sie besser an das Problem anzupassen.

AP 10 Überarbeitung der Methoden anhand von Mess- und Auswerteergebnissen

Die Ergebnisse der Schadens- und Objekterkennung sind in Abbildung 17 und Abbildung 18 für den Wagenburgtunnel und den Hesselacher Tunnel dargestellt. In diesen Ergebnissen wird die Wirksamkeit der Deep-Learning-Methoden zur automatischen Objekterkennung für die Tunnelbilder verdeutlicht. Für das Training eines Netzwerks wird eine Workstation mit einem Grafikprozessor mit mehr als 8 GB Speicher benötigt. Die Verwendung des trainierten Netzes für die semantische Segmentierung weiterer Tunnelaufnahmen kann auf einer normalen Workstation ohne GPU erfolgen.

Die für die Projektentwicklung aus Aluminium hergestellte Plattform (Abbildung 9) ist leichtgewichtig und besitzt dennoch die erforderliche Stabilität und Robustheit auch gegenüber Fahr vibrationen. Die Bildaufnahme wird über eine API auf einem Laptop gesteuert, so dass für die Datenerfassung nur ein Fahrer und ein Operateur benötigt werden. Während der Projektlaufzeit wurden die Komponenten so modifiziert, dass die Plattform tragbar und an verschiedene Fahrzeugtypen anpassbar ist (Abbildung 21). Es wird keine spezielle Ausrüstung benötigt, um die Plattform von Hand auf ein Fahrzeug zu montieren und den Rechner und die Verkabelung im Fahrzeug unterzubringen. Die Plattform mit Ihren Komponenten zur Bildaufnahmen und Datenspeicherung hat die Anforderung einer prototypischen Realisierung des Gesamtsystems erfüllt. Darauf aufbauend hat der Projektpartner Viscan für die zukünftige Verwertung der Projektergebnisse bereits ein weiterführendes Design für den Aufbau der Hardware entwickelt.

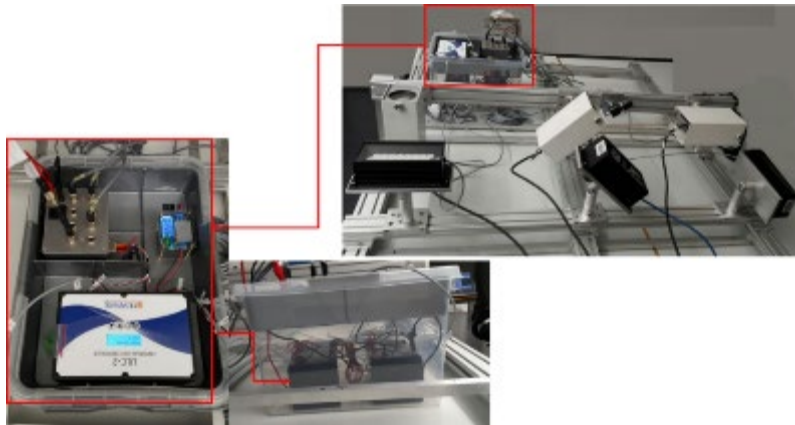


Abbildung 21: Flexibilität der Hardware bezüglich Aufbau und möglicher Modifikation

AP 11 Erstellung eines prototypischen Verfahrens und Vergleichsmessungen

Die Entwicklung und Erprobung des entwickelten Systems erfolgte mit Bilddaten, die bei der Befahrung des Wagenburgtunnels und des Heslacher Tunnels erfasst wurden. Dabei stand die Schadens- und Objekterkennung mittels Deep Learning im Vordergrund. Wie bei allen überwachten Lernverfahren wird ein Trainingsdatensatz bestehend aus den Originalbildern und den gelabelten Bildern erstellt und dieser Datensatz in drei verschiedene Datensätze unterteilt: Training-, Validierung- und Testdaten. Darüber hinausgehende Referenzmessungen werden nicht benötigt. Die Ergebnisse der Auswertung wurden schon in im Berichtspunkt AP 6: Entwicklung neuronaler Deep Learning Netze für die Objekterkennung in Tunnelbauwerken vorgestellt.

Das HFT-Laserscanning-Messsystem kam im Projekt Corona bedingt nicht zum Einsatz. Stattdessen wurde ein terrestrischer Laserscanner der HFT (Faro Focus3D X-330) zur Erfassung der Unterführung am Hauptbahnhof in Feuerbach eingesetzt. Aufgrund der Länge der Unterführung waren für die Erfassung der Punktwolke mehrere Standpunkte notwendig. Über die Ergebnisse der Auswertung wurde in Berichtspunkt AP 4: Entwicklung Prozessfluss Photogrammetrie berichtet. Visualisiert wurden die Ergebnisse in Abbildung 13 und Abbildung 14.

Quellenangabe

[1] <https://github.com/yhlleo/DeepCrack>

[2] <https://www.flir.de/support-center/iis/machine-vision/application-note/configuring-synchronized-capture-with-multiple-cameras/>

[3] www.metaphase-tech.com

[4] <https://arxiv.org/abs/1505.04597v1>

[5] <https://arxiv.org/abs/1606.00915>

[6] www.iimag.de

[7] <https://www.agisoft.com/>

[8] <https://arxiv.org/abs/1807.10165>

[9] <https://arxiv.org/abs/2004.03696>

Summarische Darstellung der erzielten Projektergebnisse

Das Projekt konnte erfolgreich abgeschlossen werden. Die Zielerreichung der im Antrag definierten Ziele wird abschließend nochmals kurz zusammengefasst.

Ziel 1. Mobile Datenerfassung durch auf einem Fahrzeug montiertes Multikamerasystem mit 8x2 4K-(Stereo-)Kameras

Bei der Entwicklung und dem Aufbau der Hardware wurde auf die an der HFT vorhandenen Kameras zurückgegriffen. Auch wenn das Setup aus vorhandenen Kameras und beschafften LEDs, mit den Parametern des finalen Messsystems nicht vollumfänglich übereinstimmt, konnte mit den in den Experimenten erprobten Systemkomponenten das Projekt erfolgreich bearbeitet und der Nachweis der Funktionstüchtigkeit des Gesamtsystems erbracht werden.

Ziel 2. Ausleuchtung der Tunneloberfläche durch Stroboskope

Die Beleuchtung und Synchronisation mit den MetaBright LEDs der Fa. Metaphase war erfolgreich.

Ziel 3. Erkennung von Undichtigkeiten durch Thermografiekamera

Wurde von der HFT nicht bearbeitet.

Ziel 4. Aufnahme von 3 m bis 5 m breiten Bildringen der Tunneloberfläche

Das Konzept Kameras und LED so zu montieren, dass Bildringe der Tunneloberfläche aufgenommen werden, hat sich als bewährt.

Ziel 5. Geschwindigkeit des Messfahrzeuges bis zu 65 km/h

Die konzipierte Erfassungsgeschwindigkeit von bis zu 65 km/h ist in den Systemkomponenten (Belichtungszeit der Kameras, Pulsbreite des LED-Blitzes, Datenrate von Schnittstellen und Datenspeicher, etc.) vollumfänglich berücksichtigt.

Die geplanten umfangreichen Tests konnten aber insbesondere aufgrund der Corona-Einschränkungen nicht im angestrebten Umfang durchgeführt werden. Bei den Experimentierfahrten der HFT wurde 36 km/h nicht überschritten.

Ziel 6. Genauigkeit der Tiefenauflösung ca. 1 mm

Die Lage- und Tiefengenauigkeit (< 1 mm, relativ) wurde wie erwartet erreicht.

Ziel 7. Einsatz von Photogrammetrie/Videogrammetrie zur automatisierten Bearbeitung der Bildverbände und Generierung vermessbarer 3D-Punktwolken.

Wurde in den Experimenten erfolgreich erprobt.

Ziel 8. Deformationsmessungen

Deformationen der Oberfläche, die sich in Rissen, Abplatzungen, etc. darstellen, sind als Teil der Schadensdokumentation über Deep-Learning-Methoden intensiv und mit Erfolg bearbeitet worden.

Ziel 9. Deep-Learning-Methoden zur automatisierten Objekterkennung

Genauigkeitsbewertungen zur Erfassung der Ausstattungsfeatures, insbesondere Schilder, Markierungen, Fahrstreifensignale, Beleuchtungseinrichtungen, Hydranten, Lautsprecher, etc. wurden mit mehreren CNNs durchgeführt, die entsprechend vorab trainiert werden mussten.

Ziel 10. Objekterkennungsrate bei 90 %

Die besten Ergebnisse bei der Objekterkennung erzielte Unet mit einer Erkennungsrate von 78,9 %. Damit konnte die angestrebte Rate von 90 % noch nicht erzielt werden. Da bei Deep Learning Ansätze mit wachsendem Trainingsdatensatz die Erkennungsrate steigt, wird jeder zukünftig bearbeitete Tunnel die bislang auf zwei Tunnel beschränkte Trainingsdaten erweitern, so dass das angestrebte Niveau erreicht werden kann.

Ziel 11. BIM-Objekterfassung

Wurde vom Projektpartner Viscan alleine bearbeitet.

5. Verwertung der Ergebnisse

Fällt in die Zuständigkeit von Viscan.