

Verwendung künstlicher neuronaler Netzwerke zur Bilderkennung für die Analyse von Straßenmarkierungen in Mobile-Mapping-3D-Punktwolken

JOHANNES WOLF¹, RICO RICHTER¹ & JÜRGEN DÖLLNER¹

Zusammenfassung: Diese Arbeit stellt einen Ansatz zur Verwendung eines etablierten künstlichen neuronalen Netzwerks aus der Bilderkennung für die effiziente semantische Klassifizierung von Straßenmarkierungen aus 3D-Mobile-Mapping-Punktwolkenscans in Straßenumgebungen vor. 3D-Punktwolken werden häufig als Datenquelle für Analysen von Geodaten verwendet. Für ihre Verarbeitung wird in vielen Anwendungsfällen eine semantische Klassifikation benötigt. Neben expliziten Regeln zur Unterscheidung durch geometrische Attribute können maschinelle Lerntechniken eingesetzt werden, um die Markierungen unter Verwendung zuvor trainierter künstlicher neuronaler Netzwerke zu identifizieren. Die eingegebenen 3D-Punktwolken werden um ein semantisches Attribut pro Punkt erweitert und es können Shapes für individuelle Straßenmarkierungen abgeleitet werden, die sich zur Weiterverarbeitung anbieten.

1 Einleitung

Im Bereich der Geodaten finden 3D-Punktwolken breite Verwendung. Sie haben sich als wertvolle Datenquelle für Analysen etabliert, da sie ohne großen Aufwand erfasst und verarbeitet werden können und detaillierte Informationen der aufgenommenen Umgebung beinhalten (VOSSELMAN et al. 2004). Technisch werden sie als ungeordnete Sammlung einzelner Messpunkte vorgehalten, die jeweils Koordinaten in drei Dimensionen besitzen sowie mögliche weitere Attribute, wie beispielsweise Intensitätswerte bei der Erfassung durch LiDAR, haben können (RICHTER et al. 2013). Die ungeordneten und unstrukturierten Punkte einer 3D-Punktwolke benötigen in vielen Anwendungsfällen eine semantische Klassifizierung für die weitere Nutzung (NIEMEYER et al. 2012). Semantische Klassifizierung bezeichnet den Prozess, jedem einzelnen Punkt ein zusätzliches Attribut zuzuordnen, das die Art des Objektes beschreibt, zu dem dieser Punkt gehört, beispielsweise „Auto“, „Laternenpfahl“ oder „Verkehrsschild – rechts abbiegen“. Wenn individuelle Objekte und deren semantische Klassen identifiziert wurden, können diese für verschiedene Anwendungsfälle genutzt werden, wie zum Beispiel zur Erstellung oder Aktualisierung von Straßenkatastern (CAROTI et al. 2005), zur Analyse von Lichtraumprofilen (MIKRUT et al. 2016) oder 3D-Modellierung (VOSSELMAN et al. 2003). Typische semantische Klassen dienen zur grundlegenden Unterscheidung zwischen Boden, Vegetation und Gebäuden. Es können darüber hinaus aber je nach konkreter Anwendung auch viele weitere und sehr detaillierte Klassen, wie zum Beispiel Autos, Straßenmarkierungen, Verkehrsschilder oder Bordsteinkanten benötigt werden (PU et al. 2011). Abb. 1 zeigt eine 3D-Punktwolke, für die

¹ Hasso-Plattner-Institut, Digital Engineering Fakultät, Universität Potsdam, Prof.-Dr.-Helmert-Str. 2-3, 14482 Potsdam, E-Mail: [johannes.wolf, rico.richter, juergen.doellner]@hpi.de

semantische Klassen ermittelt wurden und in der die Punkte in jeweils einer Farbe pro Klasse dargestellt werden.

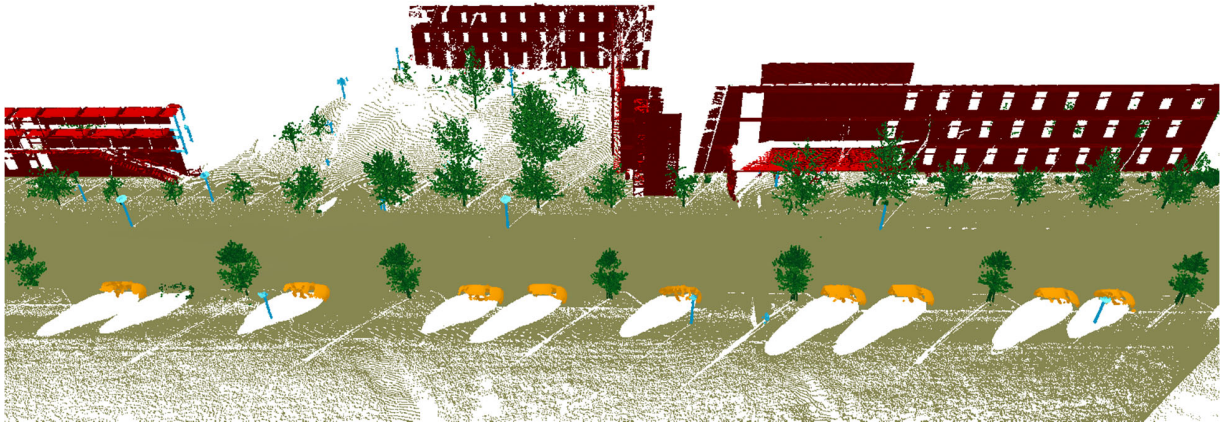


Abb. 1: Nach semantischer Klasse eingefärbte 3D-Punktwolke: Boden (braun), Vegetation (grün), Gebäude (rot), Fahrzeuge (orange), pfahlartige Strukturen (blau)

Diese Arbeit stellt einen Ansatz zur Verwendung eines künstlichen neuronalen Netzwerks aus der Bildanalyse für die semantische Klassifizierung vor. In den letzten Jahren wurden viele Techniken für die automatische Analyse von Bildern und zugehörige Frameworks entwickelt (PULLI et al. 2012). Hier wird gezeigt, dass diese auch zur Klassifizierung von 3D-Punktwolken eingesetzt werden können. Manche Objekte im Straßenraum erstrecken sich nicht in drei Dimensionen oder lassen sich ohne Verlust auch in einem zweidimensionalen Raum darstellen. Straßenmarkierungen befinden sich beispielsweise flach auf dem Boden und ihre Höhe unterscheidet sich, wenn überhaupt, nur marginal von der umgebenden Oberfläche. Sie eignen sich daher besonders für die visuelle Erkennung, weil sie ohne Schwierigkeiten in zweidimensionalen Bildern abgebildet werden können.

Die Erkennung von Straßenmarkierungen in einer Draufsicht aus 3D-Punktwolken bietet gegenüber der Erkennung in Panoramabildern den Vorteil, dass die Markierungen ohne Verzerrung erkannt und in Umringe mit präzisen Positionsangaben als Shape-Dateien überführt werden können.

2 Verwandte Arbeiten

3D-Punktwolken sind ein geeigneter Zwischenschritt zur automatischen Erstellung von 3D-Stadtmodellen (SCHWALBE et al. 2005) und Landschaftsmodellen für viele Anwendungsbereiche in der Stadtplanung von Kommunen, für Firmen und Einzelpersonen (VOSSELMAN et al. 2001). Katasterdaten können mit 3D-Punktwolken kombiniert werden, um interaktive Visualisierungen für die Erkundung und Analyse zu erstellen (ARINGER & ROSCHLAUB 2014). Sehr dichte Punktinformationen können analysiert werden und die Ableitung großer Modelle kann ohne großen manuellen Aufwand erfolgen (RICHTER & DÖLLNER 2013). Neben aus der Luft erfassten Daten werden auch Mobile-Mapping-Techniken zur Erfassung eingesetzt (LI 1997). Mobile-

Mapping-Scans können genutzt werden, um beispielsweise automatisch Straßennetze abzuleiten, Straßenoberflächen zu analysieren (JAAKKOLA et al. 2008) oder Gebäudefassaden zu rekonstruieren. Für viele Anwendungen ist die automatisierte Analyse der aufgenommenen 3D-Punktwolken eine zwingende Voraussetzung. Semantische Klassifizierung kann auf unterschiedliche Weise automatisiert durchgeführt werden: Semantische Pro-Punkt-Informationen können aus der Topologie einer 3D-Punktwolke explizit abgeleitet (CHEN et al. 2017) oder durch die Anwendung von Deep-Learning-Konzepten (BOULCH et al. 2017) ermittelt werden.

Im topologischen Ansatz werden explizite Regeln definiert, wie unterschiedliche semantische Klassen durch geometrische Attribute erkannt werden können (GRILLI et al. 2017). 3D-Punktwolken können in lokale Gruppen mit gleichen Eigenschaften, wie zum Beispiel ähnlich ausgerichtete Oberflächennormalen, segmentiert werden (RABBANI et al. 2006). Jedes dieser Segmente kann anschließend in Bezug auf Größe und Orientierung analysiert werden. Große, vertikale Flächen können so beispielsweise als Gebäudefassaden identifiziert werden, während Punktgruppen mit stark variierenden Oberflächennormalen in der Regel Teil von Vegetation sind (WOLF et al. 2019).

Der alternative Ansatz nutzt maschinelles Lernen, um die semantische Klasse von Objekten mit zuvor trainierten künstlichen neuronalen Netzwerken zu ermitteln (ZHOU & TUZEL 2018). Solche Netzwerke nutzen bereits klassifizierte Datensätze für das Training und lernen, die semantische Klasse für individuelle Punkte oder Punktgruppen in neuen, unbekanntem Datensätzen zu bestimmen. In den letzten Jahren wurden mehrere Ansätze entwickelt, die Struktur der 3D-Punktwolken direkt dafür zu nutzen, wie PointNet und ähnliche Netzwerke zeigen (QI et al. 2017). Sie fokussieren sich jedoch zumeist auf Einzelobjekte und benötigen große Trainingsdatensätze. Objekterkennung in Bildern ist seit vielen Jahren ein großes und aktives Forschungsgebiet. Viola et al. stellen einen oft zitierten Algorithmus zur Objekterkennung vor, der beispielsweise zur Gesichtserkennung in Bildern verwendet werden kann (VIOLA et al. 2001). U-Net, das ursprünglich im medizinischen Kontext entwickelt wurde, wird heute in vielen verschiedenen Bereichen zur Bildanalyse eingesetzt (RONNEBERGER et al. 2015). Es ermöglicht die automatische Erkennung von konkreten Bereichen in Bildern, wie beispielsweise Krebszellen aber auch Straßenmarkierungen in Luftbildern (ZHANG et al. 2018).

Anwendung auf 3D-Punktwolken findet U-Net auch bei SnapNet (BOULCH et al. 2018). Dort wird im Gegensatz zur hier vorgestellten Implementierung aus der 3D-Punktwolke ein Mesh generiert, für das zunächst eine Normalenberechnung durchgeführt werden muss. Der hier vorgestellte Ansatz benötigt diese Schritte nicht und fokussiert sich auf die Unterscheidung spezifischer Markierungstypen im Gegensatz zur Erkennung grober semantischer Kategorien.

3 Datensätze

In dieser Arbeit werden drei Datensätze aus unterschiedlichen Städten in Deutschland genutzt. Sie variieren im Hinblick auf Punktdichten sowie die Anzahl der Fahrzeuge, Personen und anderen Objekte, die in der Szene vorhanden sind. Das trainierte Netzwerk kann die zu erkennenden Objekte jedoch in allen Datensätzen mit ähnlicher Genauigkeit erkennen. Verschiedene Bereiche der Datensätze wurden für das Training und die Evaluation genutzt.

Eine typische Straße eines Datensatzes ist in Abb. 2 dargestellt. Die dargestellte 3D-Punktwolke besteht aus 29 000 000 Punkten und deckt etwa 670 Meter Straße mit mehreren Kreuzungen ab. Zur Erstellung eines Trainingsdatensatzes für ein U-Net-basiertes künstliches neuronales Netzwerk, das Straßenmarkierungen erkennen kann, wurden mehrere tausend Markierungen in 600 Bildern manuell klassifiziert. Data Augmentation, in diesem Fall Rotation in 90°-Schritten und Spiegelung, wurde verwendet, um den Datensatz auf das Achtfache zu vergrößern.

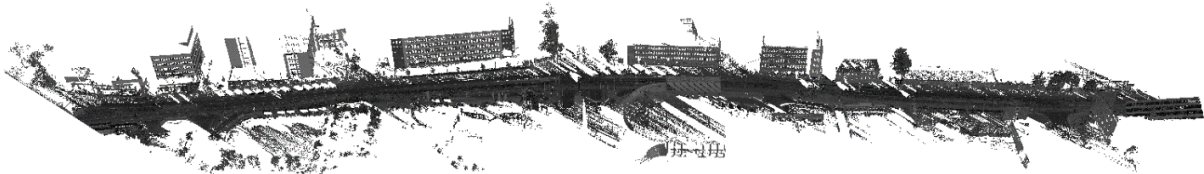


Abb. 2: 3D-Punkt wolke einer Straße, die als Eingabe für die Erkennung von Straßenmarkierungen genutzt wurde

4 Konzept und Implementierung

Der hier vorgestellte Ansatz nutzt die Möglichkeiten von Algorithmen zur bildbasierten Objekterkennung, um automatisiert Straßenmarkierungen in 3D-Punkt wolken zu klassifizieren. Straßenmarkierungen sind einfach abzubilden, da sie in einer Draufsicht auf die 3D-Punkt wolke deutlich sichtbar sind. Die durch den Ansatz erkannten Markierungen sind in Abb. 3 dargestellt.



Abb. 3: 3D-Punkt wolke, in der identifizierte Straßenmarkierungen als orangefarbene Shapes eingeblendet sind

Die Implementierung basiert auf einem Pipeline-Konzept, das große Datensätze automatisiert rendern kann, Straßenmarkierungen in den gerenderten Bildern erkennt und die Ergebnisse in die originale 3D-Punkt wolke zurückprojiziert.

Zunächst werden die Eingabe-3D-Punkt wolken wie in Abschnitt 4.1 beschrieben gefiltert. Ein Renderer erstellt anschließend quadratische Bilder, darauf wird in Abschnitt 4.2 näher eingegangen.

Die gerenderten Bilder werden mit dem zuvor trainierten künstlichen neuronalen Netzwerk klassifiziert und die Ergebnisse werden in die 3D-Punktwolke zurückprojiziert, wie in Abschnitt 4.3 beschrieben. Abschnitt 4.4 beschreibt die Ableitung von Shapes für individuelle Markierungen.

4.1 Vorverarbeitung der 3D-Punktwolken

Die Implementierung soll skalieren, so dass auch mehrere Milliarden Punkte, wie sie typischerweise bei der Erfassung ganzer Städte entstehen, effizient analysiert werden können. Eine geeignete Datenreduktion ist daher ein wesentlicher Aspekt der Verarbeitung. Um Straßenmarkierungen zu erkennen, muss nur die Straße selbst entlang der befahrenen Bereiche betrachtet werden. Während der Aufnahme der 3D-Punktwolken wird eine Trajektorie aufgezeichnet, die den exakten Pfad des Messfahrzeugs beschreibt. Die 3D-Punktwolke kann mithilfe dieser auf einen etwa 20 Meter breiten Streifen zurechtgeschnitten werden. Wenn keine Trajektorie vorliegt, kann der relevante Bereich durch eine Analyse der lokalen Punktdichte ermittelt werden, weil Bereiche in räumlicher Nähe zum scannenden Fahrzeug eine höhere Auflösung besitzen als weiter entfernte Bereiche. In den verbleibenden Daten werden Ausreißer durch eine Ausreißer-Erkennung entfernt, um Störungen in der 3D-Punktwolke zu entfernen, die das Rendern einer Draufsicht behindern könnten. Alle Punkte mit beispielsweise weniger als fünf benachbarten Punkten in einer Entfernung von einem halben Meter können als Ausreißer markiert und entfernt werden. Dieser Verarbeitungsschritt kann durch eine heuristische Suche beschleunigt werden, die eine räumliche Datenstruktur, wie zum Beispiel ein dreidimensionales Raster, nutzt, in das alle Punkte platziert werden. Punkte innerhalb einer Rasterzelle, die weniger als eine bestimmte Anzahl von Punkten beinhaltet, können als Ausreißer markiert werden. Für den speziellen Anwendungsfall dieser Arbeit ist ein solcher heuristischer Ansatz ausreichend, weil die wesentlichen Informationen alle in dichten Bereichen der 3D-Punktwolke liegen.

Ein Bodenerkennungs-Schritt identifiziert Bodenpunkte basierend auf deren relativer Höhe und Orientierung (MENG et al. 2009). Darüber befindliche Punkte können entfernt werden, sodass Punkte aus Gebäuden oder Vegetation nicht analysiert werden. Der Algorithmus unterteilt die zu untersuchende Fläche in ein reguläres zweidimensionales Raster. Für jede Rasterzelle wird der niedrigste Höhenwert der darin befindlichen Punkte ermittelt. Diese Werte repräsentieren ein vereinfachtes Bodenmodell. Nachdem das Raster so initialisiert wurde, werden Scanlines genutzt, um alle Bodenpunkte der 3D-Punktwolke zu finden. Diese Scanlines bewegen sich entlang der Achsen in positive und negative Richtung sowie diagonal durch das Raster. Der Algorithmus berücksichtigt, welcher Anstieg in welche Scan-Richtung ermittelt wurde und wie sich die Höhe zwischen Punkten und dem niedrigsten Wert der Umgebung verhält. Für jede Scanline werden mögliche Bodenpunkte separat ermittelt. Im Anschluss wird durch eine Mehrheitsentscheidung festgelegt, welche Punkte als Bodenpunkte betrachtet werden.

Die verbleibende 3D-Punktwolke beinhaltet nur noch Bodenpunkte entlang der Trajektorie des Messfahrzeugs ohne Ausreißer. Nach dem Vorverarbeitungsschritt enthielten die 3D-Punktwolken der Testdatensätze durchschnittlich noch etwa 60% der ursprünglichen Punkte.

4.2 Rendering der Bilder

Für den zu betrachtenden Bereich können nun Bilder in einer Draufsicht gerendert werden, die nacheinander durch das künstliche neuronale Netzwerk verarbeitet werden sollen.

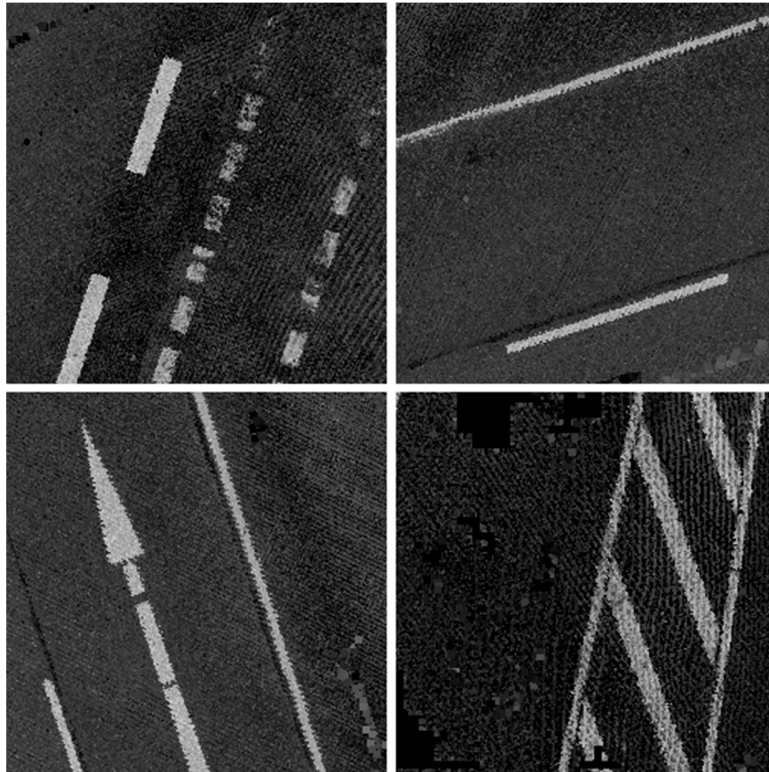


Abb. 1: Beispiele für gerenderte Bilder aus einer 3D-Punktwolke, die Fahrbahnmarkierungen zeigen

Der genutzte Renderer verwendet eine 3D-Punktwolke als Eingabe und generiert eine Menge von Bildern der Auflösung 128×128 Pixel in orthogonaler Projektion, wie in Abb. 4 dargestellt wird. Der dargestellte Bereich hat eine Größe von etwa $4,5 \times 4,5$ Metern. Die Position der Bilder wird so festgelegt, dass der komplette Bereich der zuvor gefilterten 3D-Punktwolke abgedeckt wird und sich die Bilder jeweils ein kleines Stück überlappen. Jedes Bild beinhaltet einen Kanal mit den Intensitätswerten des in jedem einzelnen Pixel jeweils sichtbaren Punktes sowie einen Kanal mit der ID des Punktes, der an dieser Position gerendert wurde. Zweiterer wird für die Rückprojektion der Klassifizierungsergebnisse in die 3D-Punktwolke benötigt.

Für bessere Ergebnisse werden die Punkte der 3D-Punktwolke als Paraboloiden gerendert, von denen in dichten Bereichen jeweils nur die Spitzen sichtbar sind. Rendering-Ergebnisse mit verschiedenen Primitiven werden in Abb. 5 dargestellt. Durch die Verwendung von Paraboloiden werden in weniger dichten Bereichen mehr Pixel ausgefüllt, um Löcher im generierten Bild zu vermeiden, während in dichten Bereich gleichzeitig scharfe Kanten einzelner Markierungen erhalten bleiben (siehe Abb. 5e).

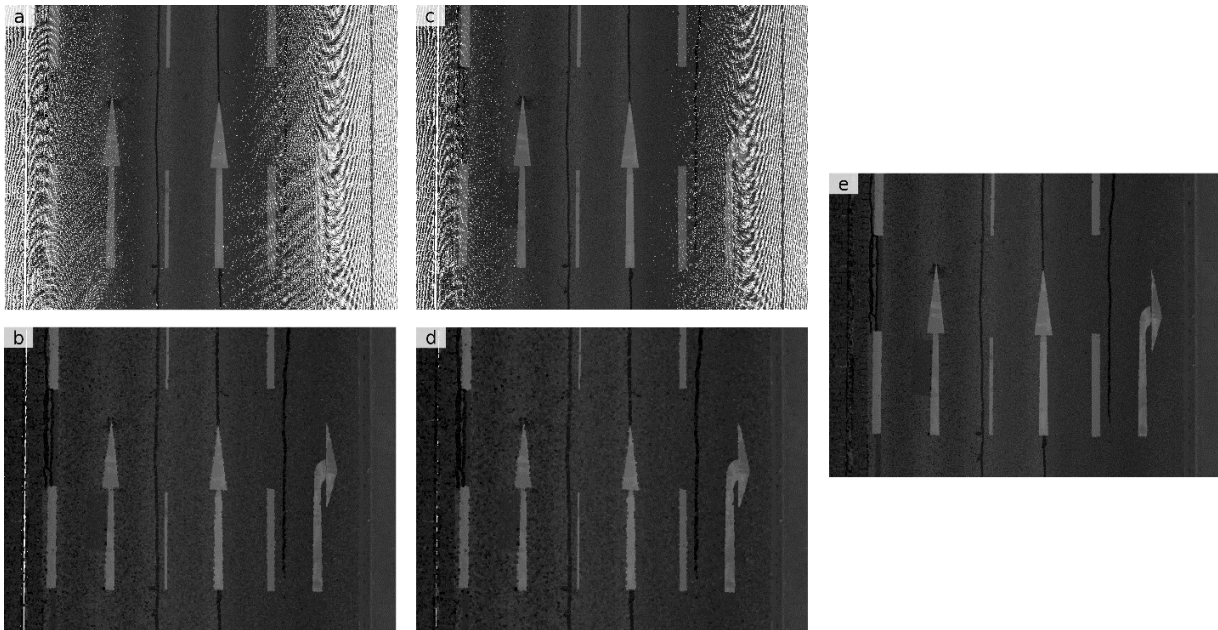


Abb. 2: 3D-Punktwolke gerendert mit unterschiedlichen Primitiven: a) Kleine Punkte (2 Pixel). b) Große Punkte (6 Pixel). c) Kleine Quads (1×1 cm). d) Große Quads (3×3 cm). e) Paraboloid

4.3 Klassifizierung

Die gerenderten Bilder werden als Eingabe für das zuvor trainierte künstliche neuronale Netzwerk verwendet. Das Ergebnis ist eine Ausgabemaske für jedes Eingabebild, in dem jedem Pixel die Art des dort erkannten Markierungstyps, bzw. die Information „keine Straßenmarkierung“ zugeordnet ist. Nachdem U-Net ermittelt hat, welche Pixel zu welcher Art Straßenmarkierung gehören, können die zugehörigen Punkte in der 3D-Punktwolke über den Punkt-ID-Kanal ermittelt werden. Diesen Punkten wird in einem neuen Attribut ihre semantische Klasse zugeordnet.

Die Zahl der Punkte in dem von einem Bild abgedeckten Ausschnitt ist zumeist höher als die Anzahl der Pixel des Bildes. Mehrere Punkte der 3D-Punktwolke sind demzufolge verdeckt und ihre IDs können nicht aus den klassifizierten Bildern ermittelt werden. Daher wird in einem Nachbearbeitungsschritt allen Punkten in der direkten Nachbarschaft von identifizierten Punkten ebenfalls das entsprechende semantische Attribut zugewiesen.

4.4 Shape-Erstellung

Für jede Gruppe zusammenhängender Punkte der gleichen semantischen Klasse kann ein Shape generiert werden (ESRI 1998). Je nach Markierungstyp werden eine konvexe Hülle oder das am besten passende Rechteck um die Punkte herum aufgespannt, wodurch Shapes für die Straßenmarkierungen erzeugt werden, wie sie in Abb. 3 dargestellt sind. Die resultierenden Dateien können anschließend in beliebigen GIS-Anwendungen für die weitere Verarbeitung genutzt werden.

5 Evaluierung

Der Evaluierungsdatensatz umfasst mehrere 3D-Mobile-Mapping-Punktwolken, die mit der gleichen Erfassungshardware in den gleichen Städten aufgenommen wurden, aus denen auch die Trainingsdaten stammen. Das künstliche neuronale Netzwerk wurde fünf Stunden auf einer Nvidia GeForce 1080 Ti trainiert. Nach manueller Prüfung wurden bei einem Durchsatz von 7,5 Millionen Punkten pro Minute 91% der Straßenmarkierungen korrekt identifiziert. Das entspricht einer Verarbeitung von etwa 300 Metern erfasster Straße pro Minute. Abb. 6 zeigt einen Ausschnitt einer automatisch generierten Shape-Datei mit erkannten Straßenmarkierungen.

Hauptsächlich wird die Geschwindigkeit der Analyse durch Lese- und Schreiboperationen beschränkt, wenn die einzelnen Bilder gerendert und auf die Festplatte geschrieben werden. Eine In-Memory-Lösung könnte den Ansatz um ein Vielfaches beschleunigen.

Fehler in der Erkennung lassen sich zum größten Teil in nicht erkannte Straßenmarkierungen in Bereichen allgemein hoher Intensitätswerte in nassen Bereichen der Fahrbahn und in falsch zugeordnete semantische Klassen, wie zum Beispiel den falschen Linientyp, zusammenfassen.

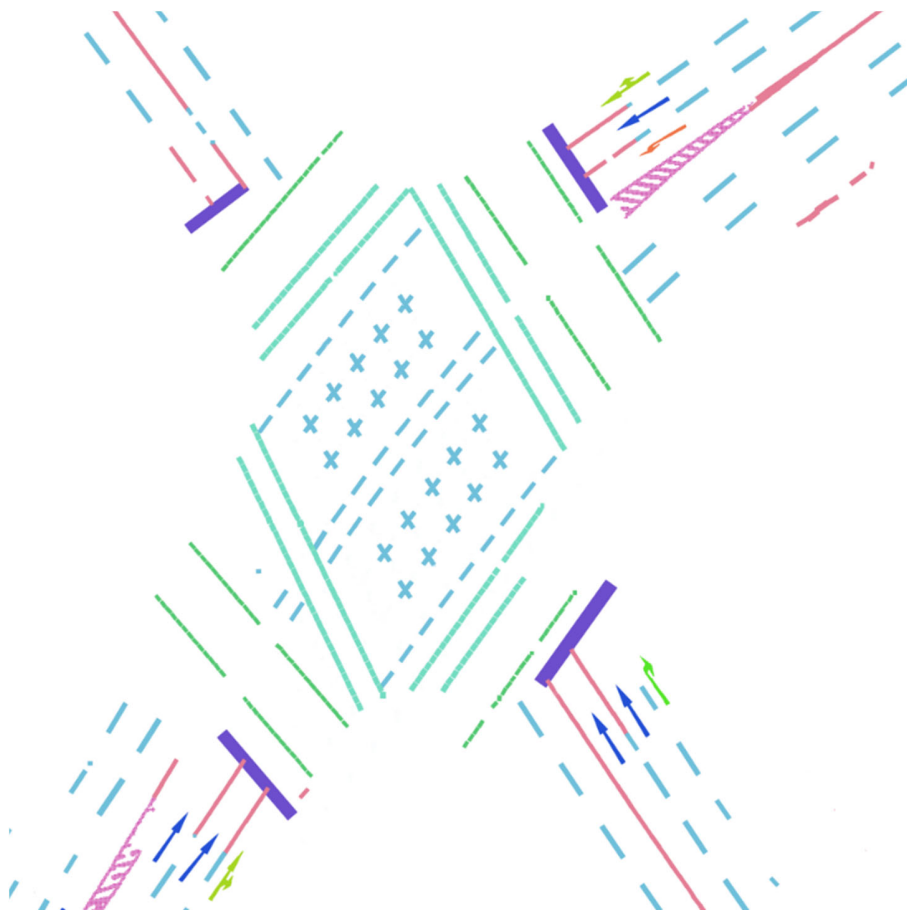


Abb. 3: Ausschnitt einer Shape-Datei mit Shapes pro erkannter Straßenmarkierung. Unterschiedliche Farben geben die jeweilige semantische Klasse an, die für die Markierung erkannt wurde.

6 Fazit und Ausblick

Die in dieser Arbeit vorgestellte Implementierung zeigt, dass ein künstliches neuronales Netzwerk für Bildanalyse zur Klassifizierung von Straßenmarkierungen in 3D-Mobile-Mapping-Punktwolken eingesetzt werden kann. Durch die Verwendung einer geeigneten Rendering-Technik werden detaillierte Bilder des erfassten Bodens generiert, die als Eingabe für das künstliche neuronale Netzwerk verwendet werden können. Es ist möglich, die identifizierten Markierungen in die 3D-Punktwolke zurück zu projizieren und Shape-Dateien abzuleiten, die in GIS-Anwendungen verwendet werden können.

Die Implementierung kann neben Straßenmarkierungen auch auf Kanaldeckel und ähnliche Objekte auf der Straße verwendet werden, dafür wird lediglich ein entsprechender Datensatz für das Training des künstlichen neuronalen Netzwerkes benötigt.

Es ist davon auszugehen, dass ein ähnlicher Ansatz auch für dreidimensionale Objekte wie Autos oder Verkehrsschilder eingesetzt werden kann, wenn ein paar Anpassungen vorgenommen werden. Dies soll in zukünftigen Arbeiten evaluiert werden. Mehrere Bilder können in horizontaler Projektion aus verschiedenen Winkeln um zuvor ermittelte Segmente herum gerendert werden, (zum Beispiel vier Bilder von jeder Seite). Ein zusätzlicher Tiefenkanal im gerenderten Bild enthält einen Teil der dreidimensionalen Information und ermöglicht eine einfache Unterscheidung zwischen Vorder- und Hintergrund. Ein solcher Ansatz könnte dann viele weitere Objekte im Straßenraum klassifizieren, die bisher nicht abgedeckt sind.

7 Danksagung

Wir danken dem Amt für Geoinformation, Vermessung und Kataster der Stadt Essen für die Bereitstellung von Daten, die im Rahmen der hier vorgestellten Arbeit verwendet wurden.

Außerdem bedanken wir uns bei Tobias Pietz und dem Bachelorprojekt 2018/2019 für ihren Beitrag zu diesem Thema.

8 Literaturverzeichnis

- ARINGER, K. & ROSCHLAUB, R., 2014: Bavarian 3D Building Model and Update Concept Based on Lidar, Image Matching and Cadastre Information. *Innovations in 3D Geo-Information Sciences*, Springer, 143-157.
- BOULCH, A., GUERRY, J., LE SAUX, B., & AUDEBERT, N., 2018: SnapNet: 3D Point Cloud Semantic Labeling with 2D Deep Segmentation Networks. *Computers & Graphics* **71**, 189-198.
- BOULCH, A., LE SAUX, B. & AUDEBERT, N., 2017: Unstructured Point Cloud Semantic Labeling Using Deep Segmentation Networks. *Proceedings of 3DOR* **2**.
- CAROTI, G., PIEMONTE, A. & PUCCI, B., 2005: Terrestrial Laser Scanning as Road's Cadastre Revision and Integration Support. *ISPRS Workshop Italy-Canada 2005, 3D Digital Imaging and Modeling: Applications of Heritage, Industry* **1**, 1-3.

- CHEN, D., WANG, R. & PEETHAMBARAN, J., 2017: Topologically Aware Building Rooftop Reconstruction from Airborne Laser Scanning Point Clouds. *IEEE TGRS* **55**(12), 7032-7052.
- ESRI, 1998: ESRI Shapefile Technical Description. <https://www.esri.com/library/whitepapers/pdfs/shapefile.pdf>, zuletzt abgerufen: 2019-11-25.
- GRILLI, E., MENNA, F. & REMONDINO, F., 2017: A Review of Point Clouds Segmentation and Classification Algorithms. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* **42**, 339-344.
- JAAKKOLA, A., HYYPPÄ, J., HYYPPÄ, H. & KUKKO, A., 2008: Retrieval Algorithms for Road Surface Modelling Using Laser-Based Mobile Mapping. *Sensors* **8**, 5238-5249.
- LI, R., 1997: Mobile Mapping: An Emerging Technology for Spatial Data Acquisition. *Photogrammetric Engineering and Remote Sensing* **63**(9), 1085-1092.
- MENG, X., WANG, L., SILVÁN-CÁRDENAS, J.L. & CURRIT, N., 2009: A Multi-Directional Ground Filtering Algorithm for Airborne LIDAR. *ISPRS Journal of Photogrammetry and Remote Sensing* **64**(1), 117-124.
- MIKRUT, S., KOHUT, P., PYKA, K., TOKARCZYK, R., BARSZCZ, T. & UHL, T., 2016: Mobile Laser Scanning Systems for Measuring the Clearance Gauge of Railways: State of Play, Testing and Outlook. *Sensors*, **16**(5), 683-701.
- NIEMEYER, J., ROTTENSTEINER, F. & SOERGEL, U., 2012: Conditional Random Fields for LiDAR Point Cloud Classification in Complex Urban Areas. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **1**(3), 263-268.
- PU, S., RUTZINGER, M., VOSSELMAN, G. & ELBERINK, S.O., 2011: Recognizing Basic Structures from Mobile Laser Scanning Data for Road Inventory Studies. *ISPRS Journal of Photogrammetry and Remote Sensing* **66**(6), 28-39.
- PULLI, K., BAKSHEEV, A., KORNYAKOV, K. & ERUHIMOV, V., 2012: Real-Time Computer Vision with OpenCV. *Communications of the ACM* **55**(6), 61-69.
- QI, C. R., SU, H., MO, K. & GUIBAS, L.J., 2017: Pointnet: Deep Learning on Point Sets for 3D Classification and Segmentation. *IEEE Conference on Computer Vision and Pattern Recognition*, 652-660.
- RABBANI, T., VAN DEN HEUVEL, F. & VOSSELMANN, G., 2006: Segmentation of Point Clouds Using Smoothness Constraint. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* **36**(5), 248-253.
- RICHTER, R., BEHRENS, M. & DÖLLNER, J., 2013: Object Class Segmentation of Massive 3D Point Clouds of Urban Areas Using Point Cloud Topology. *International Journal of Remote Sensing* **34**(23), 8408-8424.
- RICHTER, R. & DÖLLNER, J., 2013: Concepts and Techniques for Integration, Analysis and Visualization of Massive 3D Point Clouds. *Computers, Environment and Urban Systems* **45**, 114-124.

- RONNEBERGER, O., FISCHER, P. & BROX, T., 2015: U-Net: Convolutional Networks for Biomedical Image Segmentation. International Conference on Medical Image Computing and Computerassisted Intervention, Springer, 234-241.
- SCHWALBE, E., MAAS, H.-G. & SEIDEL, F., 2005: 3D Building Model Generation from Airborne Laser Scanner Data Using 2D GIS Data and Orthogonal Point Cloud Projections. The International Archives of Photogrammetry and Remote Sensing **3**, 12-14.
- VIOLA, P., JONES, M., 2001: Rapid Object Detection Using a Boosted Cascade of Simple Features. CVPR (1), 511-518.
- VOSELMAN, G., 2003: 3D Reconstruction of Roads and Trees for City Modelling. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **34**, 3-8.
- VOSELMAN, G. & DIJKMAN, E., 2001: 3D Building Model Reconstruction from Point Clouds and Ground Plans. International Archives of Photogrammetry and Remote Sensing **34**(3/W4), 37-43.
- VOSELMAN, G., GORTE, B. G., SITHOLE, G. & RABBANI, T., 2004: Recognising Structure in Laser Scanner Point Clouds. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **46**(8), 33-38.
- WOLF, J., RICHTER, R. & DÖLLNER, J., 2019: Techniques for Automated Classification and Segregation of Mobile Mapping 3D Point Clouds. 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, 201-208.
- ZHANG, Z., LIU, Q. & WANG, Y., 2018: Road Extraction by Deep Residual U-Net. IEEE Geoscience and Remote Sensing Letters **15**(5), 749-753.
- ZHOU, Y. & TUZEL, O., 2018: Voxelnet: End-to-end Learning for Point Cloud Based 3D Object Detection. IEEE Conference on Computer Vision and Pattern Recognition, 4490-4499.