

Klassifizierung von 3D ALS Punktwolken mit Neuronalen Netzen

Lukas WINIWARTER^{1,2}, Gottfried MANDLBURGER^{1,3} und Norbert PFEIFER¹

¹Department für Geodäsie und Geoinformation, TU Wien

²Geographisches Institut, Universität Heidelberg · lukas.winiwarter@uni-heidelberg.de

³Institut für Photogrammetrie, Universität Stuttgart

Zusammenfassung

Das Primärprodukt von Airborne Laser Scanning (ALS), eingesetzt für die Erfassung der Topographie, ist eine dichte, unklassifizierte 3D Punktwolke. Für die Ableitung nutzbarer Produkte ist daher eine Einteilung der Punkte in Klassen, Gruppen oder Segmente erforderlich. Für eine semantische Klassifizierung muss die Information, die inhärent in der Punktwolke in Form der geometrischen Verteilung der Punkte zueinander vorhanden ist, berücksichtigt werden. Dies geschieht üblicherweise durch Berechnung verschiedener Merkmale aus der Nachbarschaft jedes Punktes, wobei die Methoden dieser Berechnungen in konventionellen Ansätzen händisch festgelegt werden müssen. Um diese auf Expertenwissen beruhenden Entscheidungen zu vermeiden, wird ein Neuronales Netz (NN) vorgestellt, bei dem die relevanten Merkmale auf der Basis von Referenzdaten automatisch aus der Nachbarschaft abgeleitet werden.

Die Implementierung und Anwendung des Algorithmus liefert auf einem Subset der ALS Punktwolke des Landes Vorarlberg mit dem Ziel einer Landbedeckungsklassifizierung eine *overall accuracy* von 82,2 %, wobei in Stadtgebieten deutlich höhere Werte (95,8 %) erreicht werden. Dabei ist eine starke räumliche Korrelation insbesondere mit der Bodenbedeckung zu beobachten. Auf dem Benchmark-Datensatz der ISPRS wurde eine *overall accuracy* von 80,2 % erreicht, was einem Ergebnis im Mittelfeld der teilnehmenden Methoden entspricht.

Besonders in Betracht der wachsenden Anzahl der Attribute, die ALS Punktwolken mit sich bringen, verringert der Einsatz von NN die nötige Interaktion eines Anwenders. Eine entsprechend hohe Klassifizierungsgenauigkeit setzt dabei das Vorhandensein repräsentativer Referenzdaten für das Training voraus.

1 Einleitung

Klassifizierung, also die Einteilung in semantisch zusammenpassende Gruppen, ist ein wichtiger Schritt in der Prozessierung von 3D Punktwolken (OTEPKA et al. 2013). In dieser Arbeit betrachten wir dabei Punktwolken, die mithilfe von Airborne Laser Scanning (ALS) oder dichter Bildzuordnung (Dense Image Matching, DIM) gewonnen wurden und die Topographie, also das Gelände und die Objekte darauf, repräsentieren.

Typische automatische Klassifizierungen von ALS-Punktwolken erreichen dabei *overall accuracies* von 80 - 85 % (GERKE 2014, TRAN et al. 2018). Um die Genauigkeit weiter zu

steigern, wird oft manuell nachklassifiziert. Aus der Computer Vision sind jedoch Klassifizierungsgenauigkeiten jenseits der 95 % bekannt. Dort werden oft Neuronale Netze eingesetzt, welche durch wiederholtes Training (siehe Abschnitt 2.1) lernen, Klassen zu unterscheiden (CIREŞAN et al. 2012, LECUN et al. 1989). In der Bilderkennung kommen dabei sogenannte Convolutional Neural Networks (CNNs) zum Einsatz. CNNs beziehen die lokale Nachbarschaft auf Pixel-Ebene mithilfe von Faltungskernen (Convolution Kernels) ein. Diese Beobachtungen motivieren den Einsatz von Neuronalen Netzen in der Klassifizierung von ungeordneten 3D Punktwolken, um auch dort Nachbarschaften miteinzubeziehen.

2 Klassifizierung

Die Klassifizierung von 3D topographischen Punktwolken kann auf unterschiedlichen Ebenen stattfinden. In dieser Arbeit werden die Einzelpunkte unabhängig voneinander betrachtet, was zu einem Rauschen im Klassifizierungsergebnis führen kann, jedoch ein Ergebnis mit hoher Auflösung auf den originalen Beobachtungen liefert. Dadurch, dass weniger Annahmen über die Objekte in der Szene getroffen wurden, ist i.A. eine bessere Klassifizierung zu erwarten. Alternativen dazu sind die Klassifizierung von Segmenten bzw. Objekten, welche jedoch zunächst detektiert werden müssen, oder von Voxeln, also einer dreidimensionalen Raster-Repräsentation der Punktwolke.

Für eine erfolgreiche Klassifizierung müssen Merkmale bestimmt werden, die die Klassen unterscheidbar machen. Dabei gibt es solche, die bereits vom Sensor gemessen werden (etwa Signalstärke, Anzahl der Echos pro Laserpuls, Echoweite, ...), und solche, die aus der lokalen Nachbarschaft der Punkte abgeleitet werden können. Letztere beschreiben z.B. ob die Nachbarschaft um einen Punkt eine Ebene repräsentiert.

WEINMANN et al. (2013) haben gezeigt, dass eine Reihe für Klassifizierungsaufgaben relevanter Merkmale aus dem sogenannten *structure tensor*, der Kovarianzmatrix der Punktkoordinaten abgeleitet werden kann. Um Rotations- und Skaleninvarianz zu erhalten werden die Elemente der Matrix mit dem Nachbarschaftsradius normalisiert, und die Eigenwerte der Matrix im Sinne einer Hauptachsentransformation (Principal Component Analysis, PCA) berechnet (JUTZI und GROSS 2009).

Verhältnisse bzw. einfache algebraische Kombinationen dieser Eigenwerte liefern gut zu interpretierende geometrische Maße wie die *Linearität*, die *Planarität* oder die *Streuung* (WEINMANN et al. 2015). Die dazugehörigen Eigenvektoren zeigen außerdem die Richtung der Hauptkomponenten an, etwa ist der Eigenvektor zum kleinsten Eigenwert gleich dem Normalvektor der ausgleichenden Ebene der Punkte (PFEIFER 2002, 101f.).

2.1 Überwachte Klassifizierung

Überwachte Klassifizierung (*supervised classification*) basiert auf dem Vorhandensein von Trainingsdaten. Für diese Daten ist das erwartete Ergebnis bekannt. Der verwendete Klassifizierungsalgorithmus passt die Parameter so an, dass das Ergebnis der Trainingsdaten möglichst gut den erwarteten Ergebnissen entspricht. Im Gegensatz dazu stehen unüberwachte Verfahren (*unsupervised classification*), bei denen vom Algorithmus selbst Gruppierungen in den Daten gefunden und zugewiesen werden (z.B. k-Means-Clustering).

Diese Trainingsdaten müssen zunächst bestimmt werden. Oft kommen hierbei automatische Klassifikatoren zum Einsatz, deren Ergebnisse manuell nachbearbeitet werden. Wenn in den Trainingsdaten dennoch fehlerhafte oder nicht eindeutige Klassenzuweisungen vorhanden sind, spricht man i.A. von Training unter *label noise* (HICKEY 1996).

2.2 Neuronale Netze als Klassifikatoren

Neuronale Netze bestehen aus einer Menge von Neuronen, die die biologischen Vorgänge in Lebewesen imitieren. Das *Perceptron* ist ein einfaches mathematisches Modell eines solchen Neurons (Rosenblatt 1958). Dabei wird eine gewichtete Summe verschiedener Eingänge gebildet, auf welche eine Aktivierungsfunktion angewandt wird. Dies ist für ein Neuron mit n Eingängen in Gleichung 1 dargestellt.

$$O = a(\sum_{i=1}^n w_i \cdot I_i + b) \quad (1)$$

Dabei werden die Gewichte w_i sowie der *bias* b während des Trainings bestimmt. Die Aktivierungsfunktion a ist im Allgemeinen nicht linear, um die Approximation nicht-linearer Funktionen zu erlauben.

In einem neuronalen Netz werden derartige Neuronen hintereinander geschaltet. Im Falle des *Multilayer Perceptrons (MLP)* bilden die Neuronen Schichten, wobei jedes Neuron der vorhergehenden Schicht mit jedem Neuron der darauffolgenden Schicht verbunden ist. Die Anzahl der Schichten gibt die „Tiefe“ des Netzes an, die Anzahl der Neuronen pro Schicht die Breite (RUMELHART et al. 1985).

Ein Neuronales Netzwerk lässt sich auch als Transformation der Eingänge (in unserem Fall die Attribute bzw. Merkmale der Punkte) in einen erweiterten Raum darstellen, in dem die Klassen leicht separierbar sind. In der letzten Schicht werden die Ausgangswerte normalisiert, um eine Art Wahrscheinlichkeit für jede Klasse zu erhalten. Das *argumentum maximi (argmax)* dieser Wahrscheinlichkeit gibt die wahrscheinlichste Klasse zurück. Die Transformation basiert auf Basis der durch das Netzwerk angenäherten Funktionen. Dabei wurde bereits 1989 gezeigt, dass ein Netz mit der Tiefe eins ausreicht, um eine stetige Funktion beliebig genau zu approximieren, wenn es breit genug ist (CYBENKO 1989). Die Verwendung von mehreren, weniger breiten Schichten verringert jedoch die Anzahl der Neuronen und erlaubt das Erlernen generischerer Funktionen. Dabei wird etwa zunächst entschieden, ob ein Punkt Teil einer ebenen Fläche ist. Erst danach wird unter Berücksichtigung weiterer Informationen dem Punkt die Klasse Gebäude oder Boden zugewiesen (GOODFELLOW et al. 2016 3f.).

2.3 PointNet und PointNet++

Eine weitere Schwierigkeit bei Neuronalen Netzen ist die Repräsentation der Punktwolke, die inhärent ungeordnet und unregelmäßig ist (OTEPKA et al. 2013). Die Reihenfolge der Neuronen in einem Netz ist wesentlich, daher kann nicht einfach ein Punkt einer Nachbarschaft als Eingabeneuron repräsentiert werden. Um dieses Problem zu umgehen, approximieren QI, LI et al. (2017) eine Funktion f , die auf einer Punktmenge $X = \{x_1, x_2, \dots, x_n\}$ operiert, über die Verknüpfung zweier Funktionen g und h , wobei g eine symmetrische Funktion ist, deren Ergebnis nicht von der Reihenfolge der Argumente abhängig ist (siehe Gleichung 2).

$$f(\{x_1, x_2, \dots, x_n\}) = g(h(x_1), h(x_2), \dots, h(x_n)) = g(h_1, h_2, \dots, h_n) \quad (2)$$

Eine plausible Wahl für g ist eine Aggregationsfunktion wie etwa $g = \max(h_i)$ oder $g = \frac{1}{n} \sum_i(h_i)$. h hingegen wird durch ein *Multilayer Perceptron* repräsentiert, welches auf jedem Punkt gleichermaßen operiert. Um die Ableitung mehrerer Merkmale zuzulassen, sind die Funktionen g und h vektorwertig. Damit lässt sich nun ein Merkmalsvektor aus einer Menge von n Punkten x_i berechnen, der die Verteilung dieser Punkte in Raum und ggf. in weiteren Attributen bzw. Merkmalen beschreiben kann (Qi, Li et al. 2017).

In einer Erweiterung der Methode wird dieser Feature-Vektor nicht auf die gesamte Punktwolke, sondern nur auf kleine Nachbarschaften angewandt. Um diese Informationen dennoch in einen globalen Kontext betten zu können, werden für Nachbarschaftszentren, die aus der originalen Punktwolke stammen, in einem größeren Suchradius weitere Nachbarschaftsmerkmale berechnet. Dabei kommt *farthest point subsampling* zum Einsatz, wobei unterstellt wird, dass sich großräumige Nachbarschaften kleinräumig nicht auswirken.

Um am Ende die Klassifizierung der originalen Punktwolke zu erhalten, werden die so berechneten Merkmale jeweils mithilfe der drei nächstliegenden Nachbarschaftszentren mittels inverser Distanzgewichtung zurückinterpoliert. Ein einfaches *MLP* sorgt danach für eine weitere Transformation.

Damit besitzt jeder Punkt eine Reihe von Informationen: Die vom Sensor stammenden Attribute, lokale Nachbarschaftsmerkmale und interpolierte globale Nachbarschaftsmerkmale auf verschiedenen Niveaus. Diese werden aneinandergehängt und bilden die Eingabe zu einem weiteren *MLP*, welches die finale Klassifizierung vornimmt (Qi, Su et al. 2017).

Hierbei sei auf die Ähnlichkeit zu CNNs hingewiesen, welche ebenfalls auf verschiedenen Größenordnungen (durch Pooling bzw. Resampling der Bilder) Nachbarschaftsinformationen (durch die Faltungskerne) berücksichtigen.

3 Methoden

Um die Methoden von PointNet und PointNet++ auf ALS-Daten anzuwenden, müssen diese vorprozessiert werden. Da diese Ansätze in erster Linie auf Punktwolken basieren, die aus 3D-CAD-Modellen abgeleitet werden, ist die Punktmenge mit 4096 oder 2048 Punkten pro Objekt vergleichsweise gering. Eine ebenfalls von Qi, Li et al. (2017) vorgestellte Variante zur „Semantic Scene Segmentation“ bietet den besten Anhaltspunkt, wobei hier Laserscans von Umgebungen im Gebäudeinneren betrachtet wurden. Diese unterscheiden sich von ALS-Daten durch ihre starke räumliche Begrenzung. Die hier vorgestellte Variation der „Semantic Scene Segmentation“ wird *alsNet* genannt.

3.1 Batching

ALS-Daten orientieren sich in erster Linie an der Erdoberfläche. In einem projizierten Koordinatensystem wird diese als 2-Mannigfaltigkeit gesehen. Dieser Idee folgend, extrahieren wir auf Basis eines 2D-Gitters jeweils 200,000 Punkte aus dem Datensatz, indem wir die 200,000 nächsten Nachbarn der vorher über die mittlere Punktdichte bestimmten Gitterpunkte wählen. Dies führt zu annähernd kreisförmigen *batches*, die einander derart über-

lappen, dass jeder Punkt in mindestens vier *batches* zu liegen kommt. Die Begrenzung auf 200,000 Punkte kommt dabei aus der Limitierung moderner Grafikkarten, die in ihrem Speicher diese Punktmenge inklusive der abgeleiteten Merkmale vorhalten können müssen. Auf einer NVIDIA GeForce 1080Ti sind etwa 11GB Grafikspeicher verfügbar.

3.2 Prozessierung

Diese *batches* werden nun auf drei verschiedenen Nachbarschaftsniveaus nach der Methode von PointNet++ (Qi, Li et al. 2017) prozessiert. Die Hyperparameter des neuronalen Netzes werden durch zufällige Kombination aus präselektierten Optionen optimiert, wobei die *overall accuracy* nach etwa 17.500 *batches* Training als Maß für die Optimalität genommen wird. Das Ergebnis der *batch*-weisen Klassifizierung ist eine normalisierte Wahrscheinlichkeit jeder Klasse pro Punkt.

Um nun ein eindeutiges Ergebnis zu erhalten, werden die Wahrscheinlichkeiten der Klassen über die vier oder mehr *batches*, in denen der Punkt präsent ist, gemittelt, sowie die Standardabweichung berechnet. Auf Basis dieser mittleren Wahrscheinlichkeiten wird nun die wahrscheinlichste Klasse bestimmt. Die Standardabweichung kann als Maß verstanden werden, wie eindeutig das Ergebnis der Klassifizierung ist.

4 Daten

alsNet wurde auf zwei verschiedenen Datensätzen getestet. Dabei sind die räumlichen Ausdehnungen und die Anzahl der Punkte sehr unterschiedlich. Dies hat auch einen Einfluss auf die Ergebnisse.

4.1 ALS Vorarlberg

Der ALS-Datensatz des Landes Vorarlberg ist als Open Data frei verfügbar. Es handelt sich um eine Befliegung aus dem Jahr 2011, wobei eine Fläche von etwa 2.700 km² mit einer Punktdichte zwischen 10 und 20 Punkten pro Quadratmeter erfasst wurde. Aus diesem Datensatz wurden 6 Kacheln á 2,5 x 2,5 km² als Training und 4 ebenso große Kacheln als Validierungsdatensätze extrahiert. Dabei wurde auf eine ausgewogene Kombination aus Stadtgebiet, Land- und Forstwirtschaft sowie hochalpinem Gelände wertgelegt.

Die Referenzklassifizierung wurde durch die Produkte *TerraScan*, *TerraSolid* und *TerraModeler* der Firma *Terrasolid Ltd.* sowie einer manuellen Überarbeitung unter Berücksichtigung von existierenden Vektordatensätzen insbesondere zu Seilbahnbauten durchgeführt. Die genutzten Klassen sind am Schema der ASPRS LAS-Definition orientiert (ASPRS 2011). Die häufigsten Klassen hierbei sind Boden, niedrige, mittlere und hohe Vegetation, sowie Gebäude. Wasser und alle weiteren Klassen treten nur selten auf.

4.2 ISPRS Benchmark Vaihingen

Die Referenzklassifizierung wurde auf Basis einer existierenden 2D-Klassifizierung von den Autoren von NIEMEYER et al. (2014) manuell erstellt. Dabei unterscheidet sich die Klassifizierung von jener des Vorarlberg-Datensatzes. Gebäude werden zusätzlich in Fas-

sade und Dachflächen getrennt, Vegetation in niedrige Vegetation, Busch/Hecke, Baum und Strauch. Dabei ist die Unterscheidung dieser Klassen bereits manuell nicht immer eindeutig.

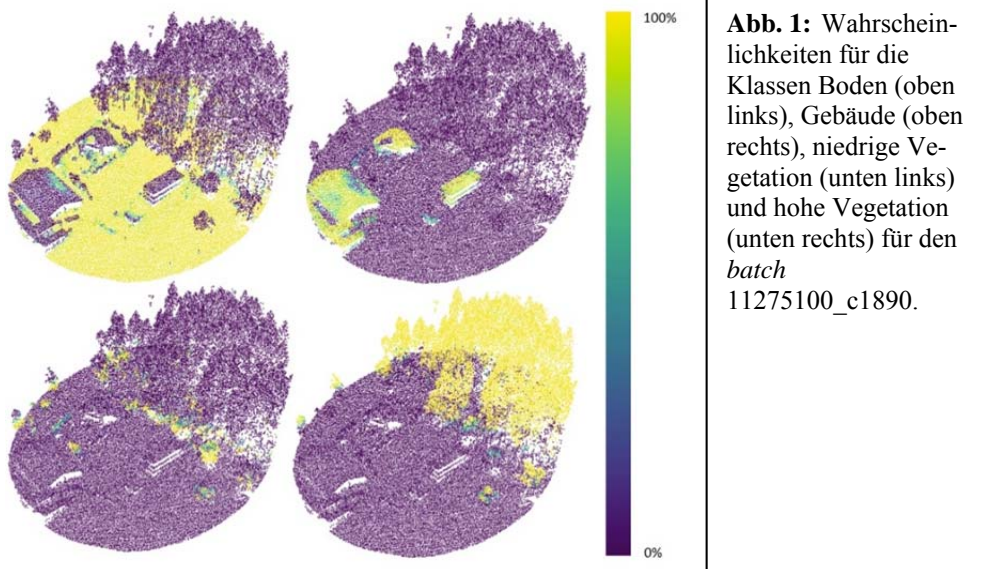
Der Datensatz besteht aus drei getrennten Teilen, einem etwa 360 x 360 m² großen Trainingsgebiet mit ca. 750.000 Punkten und zwei Evaluierungsgebieten, die jeweils etwa 200 x 150 m² Fläche und in Summe ca. 400.000 Punkte beinhalten.

Auf diesem Datensatz lief bis Sommer 2018 ein Benchmark der ISPRS, an dem 25 Methoden bzw. Varianten teilnahmen. Seitdem gibt es die Referenzdaten auch für die Evaluierungsgebiete zum Download (GERKE 2014).

5 Ergebnisse

Für den Vorarlberg-Datensatz wurden aufgrund der großen Datenmenge *batches* zum Training selektiert, die eine breite Repräsentation der vorhandenen Klassen beinhalten. So wurden etwa *batches*, die nur Bodenpunkte enthalten, nicht zum Training verwendet. Nach den bereits erwähnten etwa 17.500 *batches* zeigte sich eine Konvergenz der Klassifizierungsgenauigkeit, die durch weiteres Training nicht weiter erhöht wurde. Die Evaluierung der vier Validierungskacheln zeigte deutlich unterschiedliche Ergebnisse: im urbanen Gebiet Bregenz/Lochau wurde eine *overall accuracy* von 95,8 % erreicht, in den von Forstwirtschaft umgebenen Ortschaften Buchboden und Dalaas 82,6 % bzw. 86,7 %. Im hochalpinen Vergalda-Tal war die Klassifizierungsgenauigkeit hingegen nur noch bei 63,6 %.

Zusätzlich zur Klassenzugehörigkeit konnte pro Punkt die Klassenwahrscheinlichkeit berechnet werden. In Abbildung 1 ist dies beispielhaft für vier Klassen an einem *batch* gezeigt.



Betrachtet man die *overall accuracy* für jeden Batch, so zeigt sich eine Korrelation mit der Landbedeckung. Dies ist in Abbildung 2 für die Kachel Buchboden auf Basis einer topographischen Karte des Landes Vorarlberg dargestellt. Gut erkennbar sind felsige Bereiche im Süden und Süd-Osten, die schlechter klassifiziert wurden.

Um die Separierbarkeit der Klassen genauer untersuchen zu können, wurde eine Konfusionsmatrix erstellt. Diese ist für alle vier Validierungskacheln in Abbildung 3 dargestellt, und zeigt, dass besonders bei Punkten, die in der Referenz als niedrige Vegetation klassifiziert wurden, eine fehlerhafte Zuteilung zur Klasse Boden geschieht. Dies ist insbesondere auch bei der Kachel im Hochgebirge zu beobachten. Die Einbeziehung zusätzlicher Attribute wie der Echobreite sollte helfen, derartige Probleme zu minimieren. Ein geometrischer Grund für diese Fehlklassifizierung könnte die Annahme, sein, dass sich großräumige Nachbarschaften kleinräumig nicht ändern, also ein kontinuierlicher Verlauf der relevanten Merkmale gegeben ist. Bei Boden, der von niedriger Vegetation bewachsen ist, ist dieser kontinuierliche Übergang nicht mehr gegeben, da nur die unterste Schicht tatsächlich Boden ist.

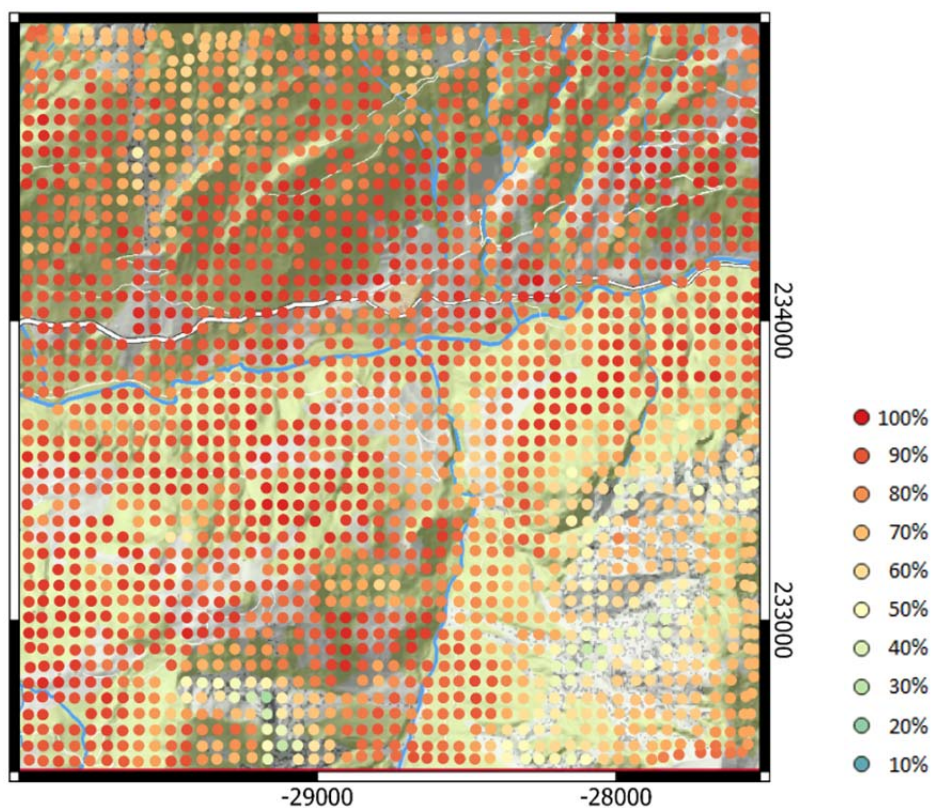


Abb. 2: *Overall accuracies* pro batch für die Kachel Dalaas. Die Kreise geben den jeweiligen Schwerpunkt der *batches* an. Basemap: Land Vorarlberg, Koordinaten: MGI/Gauß-Krüger West.

		Estimated						
		Ground 1168235830 (71%)	Low Veg. 58848094 (4%)	Med. Veg. 84927494 (5%)	High Veg. 292859439 (18%)	Building 42226870 (3%)	Water 980070 (0%)	Other 722203 (0%)
Ground truth	Ground 95527480 (37%)	98.4% 920106863	0.8% 7044815	0.1% 469099	0.0% 421766	0.8% 7081315	0.0% 371472	0.0% 32150
	Low Veg. 299366826 (18%)	81.5% 243863071	15.3% 45770070	1.0% 3111230	0.1% 358192	1.9% 5733562	0.2% 509430	0.0% 21271
	Med. Veg. 91017617 (6%)	0.7% 611994	5.5% 5000095	84.9% 77302024	6.7% 6118768	2.2% 1974998	0.0% 48	0.0% 9690
	High Veg. 289747160 (18%)	0.0% 30978	0.0% 12525	1.3% 3798730	98.6% 285551818	0.1% 339739	0.0% 0	0.0% 13370
	Building 28253245 (2%)	4.8% 1367589	1.6% 451071	0.6% 164562	0.0% 8804	92.8% 26209651	0.0% 34	0.2% 51534
	Water 873980 (0%)	91.0% 795156	0.0% 371	0.0% 52	0.0% 218	0.0% 432	8.8% 77231	0.1% 520
	Other 4013692 (0%)	36.4% 1460179	14.2% 569147	2.0% 81797	10.0% 399873	22.1% 887173	0.5% 21855	14.8% 593668

Abb. 3: Konfusionsmatrix für alle vier Validierungskacheln.

Bei der Evaluierung des ISPRS Benchmark Datensatzes wurde mit den auf Vorarlberg optimierten Hyperparametern eine *overall accuracy* von 80,6 % erreicht. Im Vergleich mit den anderen Teilnehmern des Benchmarks ist das ein Ergebnis im Mittelfeld, wobei insbesondere zwei CNN-basierte Methoden mit 85,2 % und 84,9 % nochmals deutlich bessere Klassifizierungen liefern (GERKE 2014).

Für diesen Datensatz wurde die *batch*-Dichte nochmals deutlich erhöht, sodass einzelne Punkte in bis zu 50 *batches* klassifiziert wurden. Dies ist auch durch die unregelmäßigen Umrisse der Daten bedingt. Auf Basis der resultierenden Diskrepanzen in den einzelnen Klassenwahrscheinlichkeiten wurde für jeden Punkt die mittlere Standardabweichung der Wahrscheinlichkeiten über die Klassen berechnet. Es zeigt sich, dass diese besonders dort hoch ist, wo eine Fehlklassifizierung der Daten vorliegt. Dies ist in Abbildung 4 anschaulich gezeigt.

6 Fazit und Ausblick

Die Anwendbarkeit eines End-to-End Deep Learning Ansatzes auf ALS Punktwolken wurde anhand zweier Datensätze gezeigt. Die dabei erreichten Ergebnisse sind durchaus mit anderen, nicht auf Neuronalen Netzen basierten Methoden vergleichbar. Damit wurde gezeigt, dass PointNet dazu fähig ist, die üblicherweise manuell selektierten Punktmerkmale,

die lokale Nachbarschaften beschreiben, automatisch abzuleiten. Dies reduziert nicht nur die nötigen Eingaben des Nutzers, sondern erlaubt außerdem eine automatische Adaption der Merkmale an die gegebene Klassifizierungsaufgabe. Dadurch wird ein allgemeiner Zugang ermöglicht, mit dem etwa auch Baumstämme oder Totholz in Wäldern detektiert werden könnten.

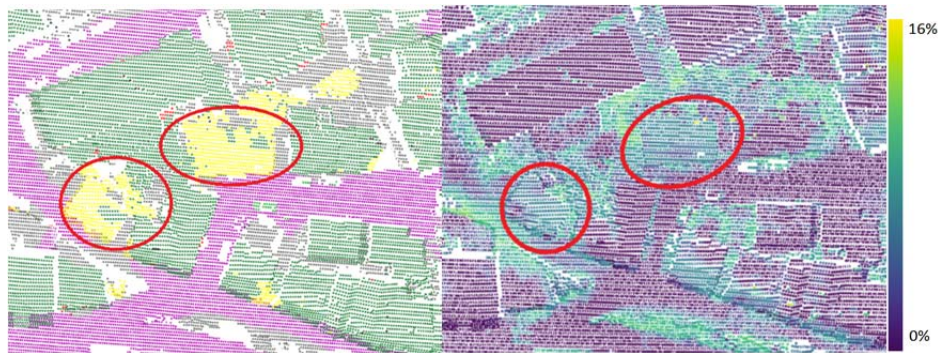


Abb. 4: Ausschnitt des ISPRS Datensatzes. Links: Klassifizierungsergebnis, Rechts: Mittlere Standardabweichung der Klassenwahrscheinlichkeit. Die markierten Gebäude weisen eine erhöhte mittlere Standardabweichung auf, und wurden fehlerhaft als Baum klassifiziert.

Außerdem wird mit *alsNet* nicht nur die Verteilung der Punkte im geometrischen Raum, sondern auch im Raum der Attribute und Merkmale berücksichtigt. Mit einer stets wachsenden Anzahl an Attributen, die direkt aus den Messungen oder durch *Data Fusion* (z.B. mit einem Orthophoto) stammen, wird die Berücksichtigung der lokalen Verteilung dieser Attribute immer wichtiger. Gleichzeitig steigt bei herkömmlichen Ansätzen der Aufwand der manuellen Merkmalsselektion.

Deep Learning ist demnach auch zur unmittelbaren Anwendung auf Punktwolken aus ALS geeignet. Der Nachteil der notwendigen großen Menge an Trainingsdaten könnte mithilfe von simulierten Daten, für die die Referenz bereits bekannt ist, verringert werden. Weiters muss nicht für jeden Zweck ein vollständig neues Modell trainiert werden: Etwa haben GIRSHICK et al. (2016) gezeigt, dass es ausreicht, die letzte Schicht eines CNNs mit wenigen Referenzdaten neu zu trainieren, um das Netzwerk an eine andere Zielklassifizierung anzupassen.

Literatur

- ASPRS (2011). *LAS Specification Version 1.4 – R13*. https://www.asprs.org/wp-content/uploads/2010/12/LAS_1_4_r13.pdf (2018-10-24)
- CIREŞAN, D., MEIER, U., & SCHMIDHUBER, J. (2012). Multi-column deep neural networks for image classification. *arXiv preprint arXiv:1202.2745*.
- CYBENKO, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of control, signals and systems*, 2(4), 303-314.

- GIRSHICK, R., DONAHUE, J., DARRELL, T., & MALIK, J. (2016). Region-based convolutional networks for accurate object detection and segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 38(1), 142-158.
- GERKE, M. (2014). *3D Semantic Labeling Contest*. <http://www2.isprs.org/commissions/comm3/wg4/3d-semantic-labeling.html> (2018-10-24)
- GOODFELLOW, I., BENGIO, Y., COURVILLE, A., & BENGIO, Y. (2016). *Deep learning*. Cambridge: MIT press.
- HICKEY, R. J. (1996). Noise modelling and evaluating learning from examples. *Artificial Intelligence*, 82(1-2), 157-179.
- JUTZI, B., & GROSS, H. (2009). Nearest neighbour classification on laser point clouds to gain object structures from buildings. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38(Part 1), 4-7.
- LECUN, Y., BOSER, B., DENKER, J. S., HENDERSON, D., HOWARD, R. E., HUBBARD, W., & JACKEL, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4), 541-551.
- NIEMEYER, J., ROTTENSTEINER, F., & SOERGEL, U. (2014). Contextual classification of lidar data and building object detection in urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 87, 152-165.
- OTEPKA, J., GHUFFAR, S., WALDHAUSER, C., HOCHREITER, R., & PFEIFER, N. (2013). Georeferenced point clouds: A survey of features and point cloud management. *ISPRS International Journal of Geo-Information*, 2(4), 1038-1065.
- PFEIFER, N. (2002). 3D terrain models on the basis of a triangulation. *Geowissenschaftliche Mitteilungen*. Wien: Inst. für Photogrammetrie u. Fernerkundung d. Techn. Univ. Wien.
- QI, C. R., SU, H., MO, K., & GUIBAS, L. J. (2017). PointNet: Deep learning on point sets for 3D classification and segmentation. *Proc. Computer Vision and Pattern Recognition (CVPR), IEEE*, 1(2), 4.
- QI, C. R., YI, L., SU, H., & GUIBAS, L. J. (2017). PointNet++: Deep hierarchical feature learning on point sets in a metric space. In *Advances in Neural Information Processing Systems* (pp. 5099-5108).
- ROSENBLATT, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6), 386.
- RUMELHART, D. E., HINTON, G. E., & WILLIAMS, R. J. (1985). *Learning internal representations by error propagation* (No. ICS-8506). California Univ San Diego La Jolla Inst for Cognitive Science.
- TRAN, THG, OTEPKA, J., WANG, D., PFEIFER, N. (2018): Classification of image matching point clouds over an urban area. *International Journal of Remote Sensing* 39(12):4145-4169.
- WEINMANN, M., JUTZI, B., & MALLET, C. (2013). Feature relevance assessment for the semantic interpretation of 3D point cloud data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 5, W2.
- WEINMANN, M., JUTZI, B., HINZ, S., & MALLET, C. (2015). Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers. *ISPRS Journal of Photogrammetry and Remote Sensing*, 105, 286-304.