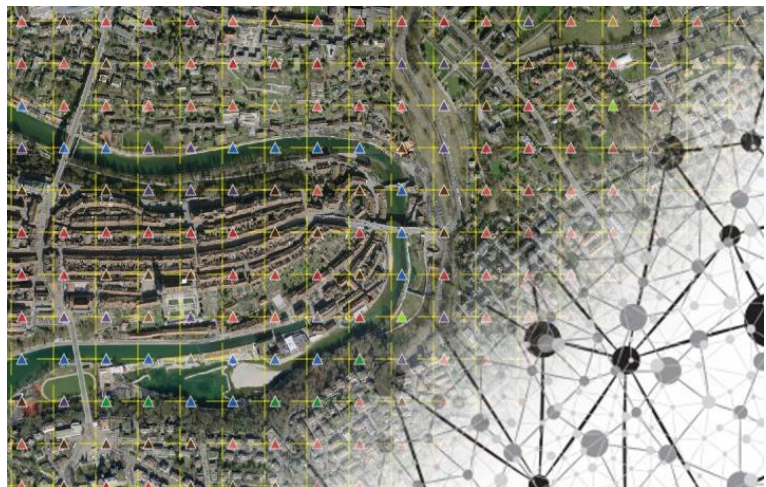


Bachelor-Thesis 2017

Einsatz von Deep Learning zur Aktualisierung der Arealstatistik der Schweiz - erste Untersuchungen



Autoren: Markus Schär

Examinator: Prof. Dr. Stephan Nebiker
Prof. Dr. Denis Jordan

Expertin: Anita Bertiller

Einsatz von Deep Learning zur Aktualisierung der Arealstatistik der Schweiz - erste Untersuchungen

Am Institut Vermessung und Geoinformatik der Fachhochschule Nordwestschweiz wurde analysiert, ob die Erfassung der Arealstatistik der Schweiz mithilfe eines Deep Learning Frameworks, einer Anwendung aus dem Machine Learning, effizient automatisierbar ist. Die erhaltenen Erkenntnisse sind aufschlussreich und bilden die Grundlage für weitere Anwendungen in der Fernerkundung.

Schlagworte: Deep Learning, Fernerkundung, Machine Learning, Arealstatistik, Luftbilder, Artificial Intelligence

1. Einführung

Ein Bild sagt mehr als tausend Worte. In den 1920er Jahren wurden in der Schweiz die ersten Luftbilder erstellt. Dadurch erhielt man einen neuen Blickwinkel auf die Landschaft Schweiz und somit auch neue Informationen. In den Anfängen wurden anhand der Luftbilder Karten ergänzt, wie zum Beispiel die Siegfried - Karte. Der Mensch erkannte mit Leichtigkeit auf den Luftbildern Häuser, Flüsse, Wälder und so weiter. Im Zeitalter der Digitalisierung wird versucht, diese Interpretationsfähigkeit der Maschine beizubringen. Dadurch können mehr Daten in kleineren Zeitabständen erfasst, verarbeitet und analysiert werden.

2. Ausgangslage

Die Arealstatistik der Schweiz ermöglicht eine Klassifizierung der Bodennutzung aufgrund von ca. 4.2 Mio. Rasterpunkten mit je einer Grösse von 1ha. Basis für die Erfassung sind die Luftbilder des Bundesamtes für Landestopografie. Dabei wird jeder Punkt durch einen Mitarbeiter des Bundesamtes für Statistik (BFS) angeschaut entsprechend klassiert. In einem zweiten Schritt wird ein reduzierter Datensatz von einem weiteren Mitarbeiter kontrolliert. Dieser Prozess beansprucht sehr viel Zeit und personelle Ressourcen. Da die Arealstatistik in Zukunft in kürzeren Perioden veröffentlicht werden sollte und zugleich weniger Personal zur Verfügung stehen wird, ist das BFS auf der Suche nach einer neuen Erfassungsmethode.

3. Deep Learning

Eine mögliche neue Methode ergibt sich aus Deep Learning. Deep Learning ist ein Subbereich des Machine Learnings und gehört somit zu der Thematik Künstliche Intelligenz. Der Begriff «deep», in der deutschen Übersetzung «tief», bezieht sich auf die tiefe Netzarchitektur des genutzten neuronalen Netzes. Diese besitzen wie das biochemische Vorbild, das menschliche Gehirn, Funktionsschichten, Neuronen und Verbindungen. Diesem Netz wird das Luftbild übergeben.

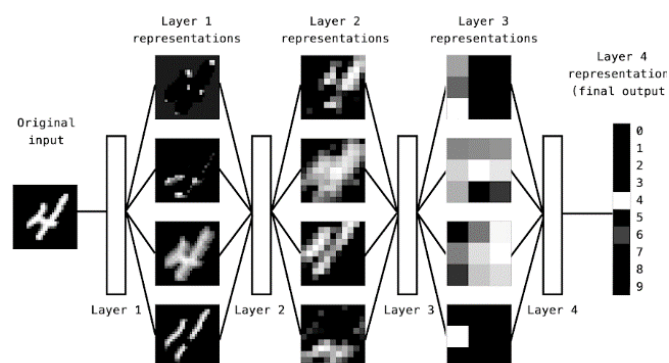


Abb. 1 Vereinfachtes Modell für Klassifikation von Bildern mit handgeschriebenen Ziffern. (Chollet, 2017a)

Jedes Neuron extrahiert aus dem Bild, mithilfe von Faltungsfilttern, Kantenfiltern und so weiter, Merkmale und gibt diese der nächsten Schicht mit Neuronen weiter, bis genügend Merkmale extrahiert werden konnten um eine Klassifikation durchzuführen. Die Analyse zeigte, dass Kenntnisse über die Methodik für eine optimale Nutzung unerlässlich sind.

Sämtliche Prozesse wurden mithilfe der Programmiersprache Python implementiert. Das genutzte Deep Learning Framework «TensorFlow» von Google Inc. wurde mithilfe der High Level API «Keras» angesteuert.

4. Erkenntnisse

Bei der Auswertung zeigte sich, dass der Datensatz gewisse Probleme für den Deep Learning Algorithmus beinhaltet. Zum einen sind einzelne Kategorien stark unterrepräsentiert, wodurch die Maschine zu wenig Erfahrungswerte bilden kann, und zum anderen beinhalten Kategorien Merkmale anderer Kategorien.



Abb. 1 Öffentliches Gebäude (links), Abb. 2 Industrie- und Gewerbeareal (rechts) Stufe 72 Kategorien

Unterschiede in den beiden Abbildungen 1 & 2 sind nicht klar erkennbar. Dennoch werden die beiden Situationen anhand der Vorgaben der Arealstatistik unterschiedlich kategorisiert. Werden die Kategorien und dadurch ähnliche Merkmale zusammengefasst, ist eine bessere Klassifikation möglich.



Abb. 3-6 v.l.n.r.: Siedlungsflächen, Landwirtschaftsflächen, Bestockte Flächen, Unproduktive Flächen auf Stufe 4 Kategorien

Überrascht hat, dass der Deep Learning Algorithmus kein deutlich sichtbares Problem hat mit der punktbasierten Erfassungsmethodik. Da das Training auf punktbasierten Klassifikationen basiert, wird diese durch die Maschine simuliert. Die relative Position der Merkmale innerhalb des Bildausschnitts ist somit ebenfalls ein Klassifikationskriterium.

5. Fazit

Die Erfassungsmethodik Deep Learning ist nach ersten Analysen nur teilweise geeignet, um die Arealstatistik leistungsstark und zuverlässig erfassen zu können. Der Datensatz hat bestimmte Eigenheiten, welche den Deep Learning Algorithmus verunsichern. In unserem

Versuch wurden nur die Luftbilder als Basis genutzt. Mit zusätzlichen Informationen könnten den eruierten Problemen entgegengewirkt werden und Deep Learning kann als zuverlässige und effiziente Klassierungsmethode für die Fernerkundung genutzt werden.

Autor: Markus Schär markus.schaer@g2014.ch

Examinator: Prof. Dr. Stephan Nebiker stephan.nebiker@fhnw.ch

Prof. Dr. Denis Jordan denis.jordan@fhnw.ch

Experte/in: Anita Beriller abertiller@sigmaplan.ch