

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

MARKUS SCHÄR¹

Zusammenfassung: Das Bundesamt für Statistik (BFS) erstellt periodisch die Arealstatistik über das gesamte Gebiet der Eidgenossenschaft. Da der personelle Aufwand für dessen Erfassung sehr hoch ist, ist der Bund auf der Suche nach einer neuen teil- oder vollautomatisierten Erfassungsmethode. In dieser Arbeit wurde mit Hilfe eines Deep Learning Frameworks, einer aus dem Machine Learning stammenden Methode, die Machbarkeit einer automatisierten Erfassung evaluiert. Dies erfolgte mit Hilfe der Ausarbeitung (Konzept, Implementierung) eines Workflows, sowie erster Evaluierungen anhand von vorliegenden Daten der Arealstatistik des BFS. Untersucht wurden insbesondere der Einfluss der Architektur und der Parametrisierung der für Deep Learning eingesetzten neuronalen Netze sowie die Eignung der Luftbilder als Grundlage. Erste Ergebnisse zeigten auf, dass sich durch eine Basiskonfigurierung der verwendeten neuronalen Netze bereits gute Ergebnisse erzielen lassen, die jedoch hinsichtlich einer zuverlässigen automatisierten Erfassung auf ein Optimierungspotenzial der Methodik hinweisen.

1 Einleitung

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, die sich zur Generierung von Landeskarten nutzen ließen. Der Mensch erkannte mit Leichtigkeit auf den Luftbildern Häuser, Flüsse, Wälder und andere Areale. Im Zeitalter der Digitalisierung hat dieser Ansatz der Arealstatistik neue Bedeutung gewonnen, indem versucht wird, einer Maschine die Klassifizierungsfähigkeit spezifischer Areale zu übertragen.

Die vom Bundesamt für Statistik (BFS) periodisch im Abstand von 12 Jahren erstellte Arealstatistik dient zur Erfassung von Veränderungen der Landnutzung in der Schweiz. Diese wird anhand von flächendeckenden Luftaufnahmen durchgeführt, indem zwei Mitarbeiter des BFS die Landnutzung mit einem Stichprobenraster von 100m Maschenweite in 72 Klassen kategorisieren (ca. 4.1 Mio. Stichprobenpunkte mit einer Abdeckung von je 1ha) und diese dem Kategorienkatalog GEOSTAT zuteilen. Eine Kontrolle der erzielten Kategorisierung erfolgt anhand eines reduzierten Datensatzes, der von einer unabhängigen Person begutachtet und im Falle von Unstimmigkeiten vor Ort abgeglichen wird. Um den Aufwand und die Aktualisierungszeitspanne zu reduzieren, sollen Verfahren untersucht werden, mit denen sich die Arealstatistik des BFS automatisieren lässt. Es handelt sich dabei um eine *Big Data* Problemstellung, für dessen Lösung sich „State of the Art“ Verfahren des maschinellen Lernens anbieten und nachfolgend als Deep Learning Framework umgesetzt werden.

¹ Fachhochschule Nordwestschweiz, Institut für Vermessung und Geoinformation, Gründenstrasse 40, CH- 4132 Muttenz, E-Mail: markus.schaer@g2014.ch

Als Bilddatenbank eignet sich das Schweizerische Orthophotomosaik, welches durch das Bundesamt für Landestopografie zur Verfügung gestellt wird und die Schweiz in zwei Bodenauflösungs-Gebiete aufteilt (Abb. 1).



Abb. 1: Aufteilung der GSD – Gebiete (BUNDESAMT FÜR LANDESTOPOGRAFIE 2017)

Je nach Region beträgt die Bodenauflösung 25cm (Abb. 1 violett) oder 50cm (Abb. 1 grün). Die Standardabweichung (1σ) der Lagegenauigkeit von SWISSIMAGE für beide GSD liegt bei einem Pixel. Beim aktuellen Datensatz wurden vier Spektralkanäle erfasst. Rot, Grün, Blau und Nahes Infrarot.

2 Methodik

Deep Learning ist ein Subthemengebiet von Artificial Intelligence und Machine Learning. Artificial Intelligence, dessen Wurzeln bis 1950 zurückreichen hat durch die wachsende Rechenleistung von Maschinen und der damit verbundenen Möglichkeit des Deep Learnings einen enormen Aufschwung erhalten.

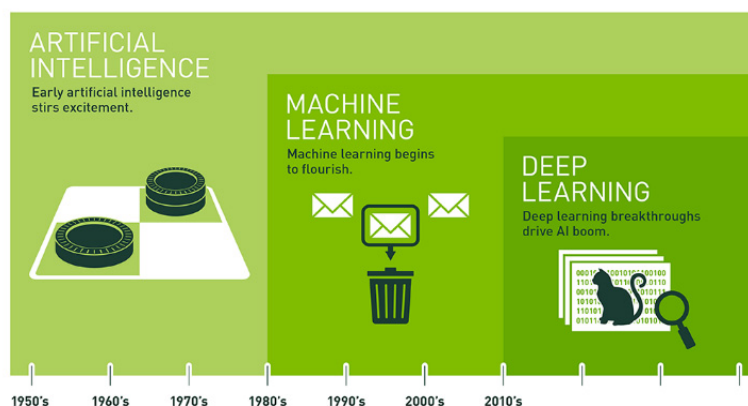


Abb. 2: Hierarchy of AI, Machine Learning and Deep Learning (COPELAND 2016)

Der oft benutzte, umfassende Begriff des maschinellen Lernens, deckt bereits einfachere der Maschine vorgegebenen Aufgaben ab. Ein klassisches Beispiel für eine einfache Anwendung des maschinellen Lernens ist die SPAM – Regel, welche Nachrichten beispielsweise anhand des Betreffs automatisch vorsortiert.

Deep Learning, auf Deutsch tiefergehendes Lernen, ist dahingehend entwickelt worden, menschliche Wahrnehmungsmechanismen maschinell anzuwenden, indem Netzwerkstrukturen des zentralen Nervensystems nachgebildet werden. Die Methodik hat sich zur automatisierten Bearbeitung von Bildinformationen als besonders geeignet erwiesen und ermöglicht mit Hilfe eines «synthetischen Nervensystems» die überwachte oder unüberwachte Kategorisierung von Daten. Grundsätzlich besteht ein für Deep Learning verwendetes neuronales Netz aus drei Hauptelementen. Den funktional getrennten Layern, den darin enthaltenen Neuronen und den übergreifenden Verbindungen.

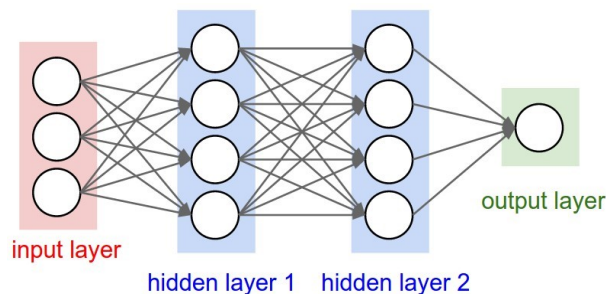


Abb. 3: 3-Layer Netzstruktur (STANFORD UNIVERSITY 2017)

Nach REY & WENDER (2010) werden die Layer in drei Bereiche aufgeteilt. Dem Netz werden wiederholt über den Input Layer (Abb. 3 rot) Informationen in Form von Zahlen übergeben. Diese Informationen werden in den Hidden Layern (Abb. 3 blau) verarbeitet und weitergegeben. Der Output Layer (Abb. 3 grün) führt anhand der erhaltenen verarbeiteten Informationen die Kategorisierung durch.

In den Layern befinden sich die sogenannten Neuronen. Diese führen die eigentliche Modifikation der Eingabewerte durch. Dabei spezialisieren sich die einzelnen Neuronen im Rahmen der vorgegebenen mathematischen Parameter auf bestimmte Merkmale des Gesamtbildes. Diese Merkmale verarbeiten die Neuronen in neue Merkmale welche wiederum in Zahlenform über die Verbindungen an neue Neuronen weitergegeben werden. Die Extraktion solcher Merkmale basiert beispielsweise auf vorgegebenen Filtern, Kantenoperatoren oder Faltungen.

Die Neuronen werden durch die Verbindungen Layer-übergreifend verbunden. Jede Verbindung besitzt eine Gewichtung. Damit ein neuronales Netz und somit der Deep Learning Algorithmus eingesetzt werden kann, muss das Netz «angelernt» werden. Dazu werden anhand eines bereits kategorisierten Datensatzes dem neuronalen Netz der Input sowie der Output bereitgestellt. Dies ist ein Trainingsprozess welcher idealerweise eine große Anzahl an unabhängigen und repräsentativen Daten erfordert und dabei die Gewichte der Verbindungen solange anpasst, bis das Netz zwischen In – und Output ein möglichst gutes Bindeglied bildet. Die Gewichte der Verbindungen repräsentieren das «Wissen» eines Deep Learning Algorithmus. Die Berechnung der Gewichte geschieht in der Regel mithilfe des iterativen Gradientenabstiegsverfahrens (Delta-Regel), der zu einer Minimierung von Fehlklassifizierungen führt.

3 Implementation

Die Implementierung wurde in drei Schritte aufgeteilt, der Datenaufbereitung, der Lern- und Auswertungsphase des Deep Learning Algorithmus und der Auswertung (Abb. 4). Sämtliche Schritte wurden mithilfe der Programmiersprache Python (Python Software Foundation, V. 3.5.3., Beaverton OR) implementiert. Mit Hilfe der High – Level API „Keras“ und dem von Google entwickeltem Deep Learning Framework „TensorFlow“ wurde das neuronale Netz implementiert und für die Aufgabenstellung trainiert.

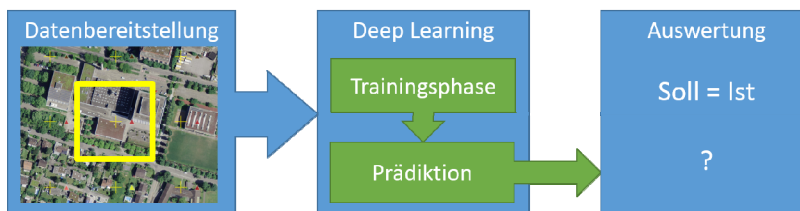


Abb. 4: Generalisierter Workflow

In einem ersten Schritt wurden sämtliche Stichprobenpunkte der Arealstatistik über den Raum mit den Luftbildern verknüpft. Zu jedem Stichprobenpunkte wird zentrisch ein quadratischer Bildausschnitt aus den Luftbildern ausgestanzt. Der Bildausschnitt wird als Input dem Deep Learning Algorithmus übergeben und die bekannte Sollkategorie aus einer bestehenden Arealstatistik als Output definiert.

In einem zweiten Schritt wird das neuronale Netz für die Aufgabenstellung trainiert. Die Daten werden für TensorFlow aufbereitet und mit der Schnittstelle Keras übergeben. Nach der Initialisierung des neuronalen Netzes, werden die Daten nacheinander an das Netz übergeben. Zuletzt werden die berechneten Gewichte der Knoten fixiert und ein unabhängiger Datensatz dem Netz zur Prädiktion übergeben. Dabei wird der Output dem Netz selbst überlassen. Das Ergebnis wird anschließend mithilfe der aus der Fernerkundung bekannten Fehlermatrix mit der der Sollkategorie verglichen.

4 Ergebnis

Der Katalog der Arealstatistik besitzt unterschiedliche Stufen mit bis zu 72 unterschiedlichen Kategorien. Durch die detaillierte Unterteilung der schweizerischen Landschaft entstehen stark unter- und überrepräsentierte Kategorien.

Tab. 1: Über- / unterrepräsentierte Kategorien, Stufe 72

Kategorie	Anteil unausg.	Anteil ausg.	Kategorie	Anteil unausg.	Anteil ausg.
Normalwald	25 %	6 %	Reihen und Terrassenhäuser	0.03 %	0.09 %
Ackerland	17 %	5 %	Lawinen- und Steinschlagverbauungen	0.02 %	0.11 %
Günstige Alpweiden	10 %	6 %	Flugplätze	0.02 %	0.08 %
Naturwiesen	7 %	5 %	Abwasserreinigungsanlage	0.02 %	0.04 %
Stehendes Gewässer	5 %	6 %	Landschaftseingriffe	0.01 %	0.04 %
Gesamtgenauigkeit	36.5 %	37.0 %	mittlere Klassengenauigkeit	24.5 %	30.7 %

In Tabelle 1 sind die fünf Kategorien mit höchster und kleinster Repräsentation aufgeführt. Drei der 72 Kategorien machen zusammen mehr als die Hälfte aller Datenpunkte aus. Bei der Auswertung stellte sich heraus, dass unterrepräsentierte Kategorien eine deutlich schlechtere Klassifikationsgenauigkeit aufweisen.

Um den Einfluss der Verteilung der Kategorien zu untersuchen wurde ein Datensatz erstellt, bei dem die Verteilung ausgeglichener ist. Dadurch blieben die Klassifikationsergebnisse jedoch beinahe gleich, sowohl für die einzelnen Kategorien als auch in der gesamten Auswertung (Gesamtgenauigkeit). Verbessert hat sich aber die mittlere Klassengenauigkeit welche auf eine bessere Genauigkeitsverteilung unter den Kategorien hinweist. Insbesondere unterrepräsentierte Kategorien welche sich durch individuellere Merkmale von anderen Kategorien abheben, werden deutlich besser kategorisiert.

Teils sehr ähnlichen Merkmale in unterschiedlichen Kategorien führen zu folgender Problematik.



Abb. 5a & 5b: Öffentliches Gebäude (links), Industrie – Gewerbe (rechts)

Eine Differenzierung zwischen den Abbildungen 5a und 5b auf alleiniger Basis des Luftbildes durchzuführen, ist nicht eindeutig möglich. Es fehlen eindeutige Merkmale. Dennoch werden die Bilder in der Arealstatistik des BFS unterschiedlichen Kategorien zugeordnet. Solche Situationen scheinen den Deep Learning Algorithmus zu überfordern.

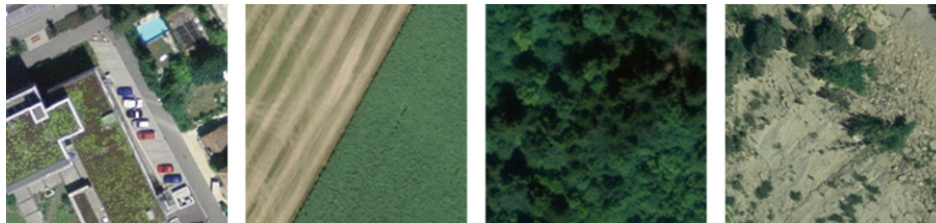


Abb. 6a–6d: v.l.n.r.: Siedlungsflächen, Landwirtschaftsflächen, Bestockte Flächen, Unproduktive Flächen auf Stufe 4 Kategorien

Sind Kategorien wie in Abbildung 6 mit ähnlichen Merkmalen zusammengefasst, resultiert eine klarere Kategorisierung.

Dennoch ließen sich mit Hilfe einer Parameteroptimierung am Deep Learning Algorithmus folgende Ergebnisse erzielen (Tab. 2).

Tab. 2: Resultierte Gesamtgenauigkeiten

Anzahl Kategorien:	4	17	27	72
Gesamtgenauigkeit:	83%	63%	59%	56%

5 Diskussion

Die große Hürde eine automatisierte Erfassungsmethode für die Arealstatik zu finden, welche mit den vorabgehenden menschlich interpretierten Datensätzen harmonisierbar ist, kann mit Deep Learning gemeistert werden. Die Datenverarbeitung ist mit einem Deep Learning Framework und mit Hilfe von High – Level API's ohne stark vertiefte Kenntnisse implementierbar. Die ersten erhaltenen Ergebnisse überzeugen und bestätigen den erarbeiteten Ansatz.

Bei der Wahl des neuronalen Netzes und den damit verbundenen benötigten Parametern sind gewisse Kenntnisse über die Methodik unerlässlich. Es hat sich gezeigt, dass Parameter sowie die Lernregel einen starken Einfluss auf das Resultat haben.

Ein ausgeglichener Datensatz ohne starke über- oder unterrepräsentierte Kategorien scheint eine wichtige Voraussetzung zu sein, um ein Netz effizient trainieren zu können. Dabei ist ein Zusammenspiel zwischen der Anzahl Datensätze und den Merkmalen der jeweiligen Kategorie feststellbar. Kategorien welche sich anhand ihrer Merkmale vom restlichen Datensatz abgrenzen, benötigen weniger Trainingsdaten und können trotz einer Unterrepräsentierung klassiert werden. Um ein besseres Ergebnis zu erhalten, müsste der Datensatz noch weiter optimiert werden.

Das BFS benutzt für die Klassierung der Arealstatistikpunkte teilweise zusätzliche Informationen aus anderen Quellen. Diese zusätzlichen Informationen fehlen dem Deep Learning Framework um zuverlässig eine Klassifikation durchführen zu können. Die Anzahl an Informationen welche dem Framework übergeben werden können ist grundsätzlich unbegrenzt. Damit ähnliche Kategorien zukünftig getrennt werden könnten, müsste geprüft werden, welche Informationen neue Merkmale hinzufügen würden und wie diese an das Deep Learning Framework übergeben werden können.

6 Literaturverzeichnis

BUNDESAMT FÜR LANDESTOPOGRAFIE, 2017: SWISSIMAGE - Das digitale Farborthophotomosaik der Schweiz. https://www.swisstopo.admin.ch/content/swisstopo-internet/de/home/products/images/ortho/swissimage/_jcr_content/contentPar/tabs/items/dokumente/tabPar/downloadlist/downloadItems/588_1464190449870.download/infosi201003deu.pdf, [7.4.2017].

COPELAND, M., 2016: The Difference Between AI, Machine Learning, and Deep Learning? [NVIDIA Blog. The Official NVIDIA Blog. <https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/>, [21.7.2017].

REY, G.D. & WENDER, K.F., 2010: Neuronale Netze: Eine Einführung in die Grundlagen, Anwendungen und Datenauswertung. Bern: Hogrefe, vorm. Verlag Hans Huber, 15.

STANFORD UNIVERSITY, 2017: CS231n Convolutional Neural Networks for Visual Recognition. Stanford University. <http://cs231n.github.io/neural-networks-1/>, [26.7.2017].

Bei sämtlichen aufgeführten Bildausschnitten aus dem Luftbilddatensatz SWISSIMAGE liegt das Urheberrecht beim Bundesamt für Landestopografie (swisstopo, Vertrag JA100071).