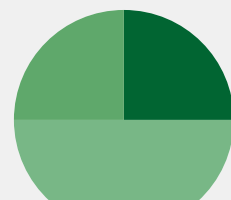


BFS Aktuell



02 Raum und Umwelt

Neuchâtel, September 2023

Einsatz von künstlicher Intelligenz in der Arealstatistik des BFS

Funktionsweise, Fähigkeiten und Leistungsgrenzen des Deep-Learning-Tools ADELE

Seit 2022 trägt erstmals künstliche Intelligenz «im Echteininsatz» zur Erstellung einer öffentlichen Statistik in der Schweiz bei: Das Deep-Learning-Tool ADELE erlaubt es, die arbeitsaufwendigen Luftbild-Interpretationen der Arealstatistik teilweise zu automatisieren. ADELE zeigt exemplarisch, welche neuen Möglichkeiten die künstliche Intelligenz der öffentlichen Statistik bietet und wie ihr Einsatz in der Praxis konkret aussehen kann. ADELE offenbart aber auch gewisse Reibungspunkte zwischen den Verfahren der künstlichen Intelligenz und dem Anspruch, eine präzise Statistik mit auf lange Frist zuverlässigen Zeitreihen zu erstellen.

1 Die BFS-Arealstatistik: Inhalt und Zweck

Die vom Bundesamt für Statistik (BFS) erstellte Arealstatistik gibt Auskunft über die Nutzung und die Bedeckung des Bodens in der Schweiz und darüber, wie sich diese über die Zeit verändern. Damit bildet sie ein unverzichtbares Instrument der räumlichen Langzeitbeobachtung. Dank der Ergebnisse der Arealstatistik lässt sich unter anderem beurteilen, ob und inwieweit die Änderungen bei der Bodennutzung den Zielen der Schweizer Raumentwicklungspolitik entsprechen – und damit auch dem Wunsch nach einem haushälterischen Umgang mit den Landressourcen.

Gegenwärtig liegen die Resultate von insgesamt vier abgeschlossenen und methodisch einheitlichen Erhebungen vor, nämlich 1985, 1997, 2009 und 2018 (siehe auch Kasten rechts). Es handelt sich dabei immer um die Schlussjahre von jeweils mehrjährigen Erhebungsperioden. 2021 wurde mit einer weiteren Erhebung begonnen. Erste Teilergebnisse (Kantone Genf und Waadt) wurden im Sommer 2023 veröffentlicht. Die Resultate für die Gesamtschweiz liegen voraussichtlich 2027 vor.

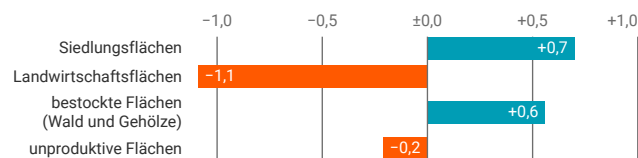
Resultate und detaillierte Methodendokumente zur Arealstatistik

Untenstehende Grafik zeigt ein Beispiel einer Auswertung basierend auf den Daten der Arealstatistik: Zwischen 1985 und 2018 wurde in der Schweiz pro Sekunde eine Fläche von durchschnittlich 0,7 Quadratmetern neu bebaut, das heisst, zu Siedlungsfläche umgenutzt. Zahlreiche weitere Grafiken sowie Detailtabellen, Erklärungstexte und Methodendokumente zur Arealstatistik finden sich auf der entsprechenden BFS-Webseite unter:

www.landuse-stat.admin.ch

Veränderung der Bodennutzung zwischen 1985 und 2018 in m² pro Sekunde

Zu- und Abnahme nach den Hauptbereichen der Arealstatistik



Quelle: BFS – Arealstatistik (AREA)

© BFS 2023

2 Datenquellen und herkömmlicher Erhebungsprozess

Luftbilder als Basis

Basis der Arealstatistik bilden hochaufgelöste digitale Luftbilder, die das ganze Territorium der Schweiz abdecken und die dem BFS vom Bundesamt für Landestopografie zur Verfügung gestellt werden. Die Bilder werden von entsprechend ausgerüsteten Flugzeugen aus aufgenommen, die die Schweiz in regelmässigen Abständen überfliegen. Die Fotos dienen primär der Erstellung und Nachführung der Schweizer Landeskarten; ihre Verwendung für die Arealstatistik stellt eine von zahlreichen Zweitnutzungen dar.

Über 4 Millionen Stichprobenpunkte

Den für die Arealstatistik verwendeten Luftbildern wird in einem ersten Schritt ein Raster mit Stichprobenpunkten im Abstand von je 100 Metern überlagert. Somit kann sich die Statistik auf insgesamt 4,1 Millionen Stichprobenpunkte stützen. Die Punkte sind georeferenziert und fix: Ihre Position bleibt über die verschiedenen Erhebungen hinweg genau gleich. Für jeden einzelnen Stichprobenpunkt wird bei jeder Erhebung anhand der jüngsten Fotos von Neuem die Bodennutzung sowie die Bodenbedeckung bestimmt. Danach werden die Ergebnisse auf das gesamte Territorium hochgerechnet.

Visuelle Interpretation der Luftbilder

Die Bestimmung der Bodennutzung und -bedeckung an den Stichprobenpunkten geschieht herkömmlicherweise «per menschlichem Auge»: Entsprechend geschulte Interpretinnen und Interpreten betrachten die Luftbilder auf einem speziellen 3D-Stereobildschirm und ordnen jeden Stichprobenpunkt einer der vordefinierten Nutzungs- und Bedeckungskategorien zu. Hierzu

besehen sie sich auf den Fotos die unmittelbare Umgebung des zu interpretierenden Punktes: Liegt dieser auf einem Gebäude? Und falls ja: Um welche Art von Gebäude handelt es sich? Um ein Wohngebäude, ein Industriegebäude, ein landwirtschaftliches Gebäude? Insgesamt wird zwischen 46 verschiedenen Typen der Bodennutzung und 27 Typen der Bodenbedeckung (z. B. «Gebäude oder «Gebüsch») unterschieden. Alle Kategorien der Bodennutzung können einer von vier Hauptnutzungen zugeteilt werden, nämlich: Siedlungsflächen, Landwirtschaftsflächen, bestockte Flächen (Wald, Gehölze) und unproduktive Flächen.

Hilfsinformationen für die Bildinterpretation

Um den Interpretinnen und Interpreten die korrekte Bestimmung der Stichprobenpunkte zu erleichtern, werden ihnen auf zwei zusätzlichen Bildschirmen verschiedene Hilfsinformationen zur Verfügung gestellt. So erlaubt beispielsweise das Gebäude- und Wohnungsregister eine genaue Unterscheidung zwischen Ein- und Mehrfamilienhäusern, das topografische Landschaftsmodell der Schweiz gibt Auskunft über die Strassenbreite, und das Inventar der landwirtschaftlichen Nutzflächen hilft, Spezialkulturen wie Obstanlagen oder Reben zu erkennen. Trotz des Einsatzes dieser und vieler weiterer Hilfsdaten gibt es allerdings vereinzelt auch Stichprobenpunkte, bei denen die genaue Nutzung oder Bedeckung am Bildschirm nicht mit Sicherheit bestimmt werden kann. Solche Punkte werden bei Geländebegehungen vor Ort begutachtet.



Interpretation der Luftbilder am 3D-Stereobildschirm. Auf den Zusatzbildschirmen werden der Interpretin Hilfsinformationen zur Verfügung gestellt.



Luftbild mit Stichprobenpunkten und den Codes für die dort festgestellte Bodennutzung. Der Code 106/11 etwa steht für «befestigte Fläche auf öffentlichem Gebäudeareal».

3 Einsatz von künstlicher Intelligenz in der Arealstatistik: Potentiale und Herausforderungen

Trotz der genannten Hilfsinformationen und des Einsatzes gewisser Unterstützungswerkzeuge (z. B. automatische Plausibilitätsprüfung während der Eingabe) bleibt die visuelle Interpretation der Luftbilder angesichts der enormen Menge von 4,1 Millionen Stichprobenpunkten äusserst aufwendig. Der Gedanke liegt daher nahe, diese Arbeiten mittels des Einsatzes geeigneter Bilderkennungsprogramme zu automatisieren. Die Möglichkeiten der automatischen Bilderkennung haben sich in den letzten Jahren dank des Einsatzes künstlicher Intelligenz in Form maschinellen Lernens (genauere Erklärungen dazu folgen noch) stark erweitert.

Ein Einsatz von künstlicher Intelligenz (KI) im Rahmen der Arealstatistik scheint umso verlockender, als für das «Training» der entsprechenden Programme auf enorme Mengen bereits existierenden «Lernstoffs» zurückgegriffen werden kann: Im Zuge vergangener Arealstatistik-Erhebungen sind bereits Millionen von Luftbildern von Menschen interpretiert worden, und jede einzelne Interpretation wurde samt dem dazugehörigen Bild digital aufbewahrt. Die für das Training der automatischen Bilderkennung sonst vielfach notwendige manuelle Aufbereitung der Lerndaten («Klickarbeit») wurde somit bereits geleistet.

Aufgrund des vermuteten Potentials hat das BFS schon 2017 in einer Machbarkeitsstudie abklären lassen, ob und wie die Luftbildinterpretation der Arealstatistik mittels künstlicher Intelligenz zumindest teilweise automatisiert werden könnte. Dies auch mit dem Ziel, die Erhebung künftig in kürzeren Abständen durchführen zu können. Schnell zeigte sich jedoch, dass im Fall der Arealstatistik nicht nur die Potentiale gross sind, sondern auch die Herausforderungen:

- Die Anforderungen an die Präzision sind bei der Arealstatistik ungleich höher als bei vielen anderen, vor allem kommerziellen Nutzungen der automatischen Bilderkennung.
- Die riesige Zahl der zu bestimmenden Stichprobenpunkte erfordert eine entsprechend grosse Rechenkapazität und -zeit. Angenommen, die Bestimmung eines einzelnen Punktes dauert 1 Sekunde: Bei 4,1 Millionen Punkten entspricht dies einer Berechnungsdauer von 47 Tagen!
- Einige Nutzungs- und Bedeckungskategorien sehen auf den Luftbildern sehr ähnlich aus, sodass für die Bestimmung auch auf Hilfsinformationen zurückgegriffen werden muss. Die eingesetzte künstliche Intelligenz muss daher neben Bild- noch andere Informationen verarbeiten können. Dies erfordert eine komplexere Programm-Architektur und steigert den ohnehin schon grossen Rechenaufwand noch zusätzlich.
- Die Häufigkeit der verschiedenen Nutzungs- und Bedeckungskategorien ist sehr ungleich verteilt. Einige Kategorien kommen so selten vor, dass für sie nicht genügend Lernstoff für das Training der künstlichen Intelligenz zur Verfügung steht. Ausserdem besteht die Gefahr, dass die seltenen Kategorien von den Algorithmen des maschinellen Lernens gleichsam übergangen werden. Die Häufigkeit der Kategorien muss daher mit Hilfe statistischer Verfahren ausbalanciert werden.

4 Das KI-Tool ADELE: Aufbau und Funktionsweise «im Prinzip»

Unter anderem aufgrund der beschriebenen Schwierigkeiten wird eine vollständige Automatisierung der Luftbild-Interpretation mittels künstlicher Intelligenz kurz- wie mittelfristig kaum möglich sein. Nichtsdestotrotz entschloss sich das BFS, zusammen mit sachkundigen Partnern ein entsprechendes KI-Tool zu entwickeln. Denn schon eine teilweise Automatisierung der Interpretationsarbeiten vermag diese erheblich zu beschleunigen. Da sein Kernmodul mit dem Verfahren des sogenannten «Deep Learning» arbeitet, wurde das KI-Tool «ADELE» genannt, was für «**A**realstatistik **DE**ep **LE**arning» steht.

Im Folgenden wird zunächst beschrieben, wie die automatische Bodennutzungs- und Bodenbedeckungsbestimmung mit ADELE «im Prinzip» funktioniert, also ungeachtet der praktischen Limitierungen des Tools. Anschliessend soll gezeigt werden, wie ADELE in der Praxis mit der menschlichen Bildinterpretation kombiniert wird, sodass die Erstellung der Arealstatistik an Geschwindigkeit gewinnt, ohne dass die Ergebnisse an Qualität und Zuverlässigkeit verlieren.

4.1 Das ADELE-Kernelement: ein Deep-Learning-Modul zur «visuellen» Luftbild-Interpretation

Wie gesehen, stützen sich die menschlichen Interpretinnen und Interpreten für die Bestimmung der Bodennutzung und -bedeckung an den Stichprobenpunkten in erster Linie auf die in den Luftbildern «enthaltenen» Informationen: Sie nehmen eine visuelle Interpretation vor. Gleiches gilt auch für ADELE: Die Kernkompetenz des Tools besteht darin, auf denselben Luftbildern, die bisher von Menschen interpretiert wurden, die Zonen rund um die Stichprobenpunkte auf Muster hin abzusuchen, die eine Zuteilung der Punkte zu jeweils einer der 46 Bodennutzungs- und 27 Bodenbedeckungskategorien der Arealstatistik erlauben. Konkret wird zu jedem Stichprobenpunkt ein Bildausschnitt von 50 auf 50 Metern untersucht, was 200 x 200 und somit 40 000 dreifarbigen Pixeln entspricht. Die entsprechenden Farbwerte bilden die Datenbasis für die Bildinterpretation durch ADELE.

Doch wie funktioniert diese Interpretation? Die Antwort auf diese Frage ist ebenso einfach wie schwierig: ADELE bestimmt die Bodennutzung und -bedeckung anhand eines statistischen Modells, bestehend aus einer riesigen Zahl von vom Tool selbst eingestellten Parametern. ADELE wurde nicht etwa von menschlichen Programmierinnen und Programmierern mit formalen Regeln des Typs «Farbe Weiss = Gletscher» gefüttert. Denn es wäre schlichtweg unmöglich, alle Regeln manuell zu programmieren, die für die korrekte Interpretation der Luftbilder nötig sind. Stattdessen wurde dem Tool durch den Menschen die Fähigkeit verliehen, sich auf Basis älterer Arealstatistik-Luftbilder und der dazugehörigen menschlichen Bildinterpretationen die korrekte Bestimmung der Bodennutzung und -bedeckung selbst «anzu-trainieren». Dazu kommt in der Kernkomponente von ADELE ein künstliches neuronales Netzwerk zum Einsatz, das nach dem schon mehrfach angesprochenen Prinzip des sogenannten

Deep Learning

«Deep Learning», auf Deutsch «tiefes Lernen», ist die Bezeichnung für eine bestimmte Art des maschinellen Lernens und somit für einen Teilbereich des grossen Gebiets der künstlichen Intelligenz. Die Grundprinzipien des Deep Learning sind schon seit Jahrzehnten bekannt. Einer erfolgreichen Anwendung standen aber lange Zeit vor allem beschränkte Rechenkapazitäten sowie ein Mangel an geeigneten Eingangsdaten für die nötigen Lernprozesse im Weg. Seit diese Hindernisse überwunden sind, erlebt das Deep Learning einen grossen Boom und ermöglicht gerade bei der automatischen Bild- und Spracherkennung Ergebnisse, die bis vor Kurzem unerreichbar schienen.

Das Grundprinzip des Deep Learning ist ein Stück weit von der Funktionsweise des menschlichen Gehirns inspiriert, weshalb die entsprechenden Algorithmen auch als «künstliche neuronale Netzwerke» bezeichnet werden. Als Hardware fungieren aber normale Computerprozessoren.

Funktionsweise

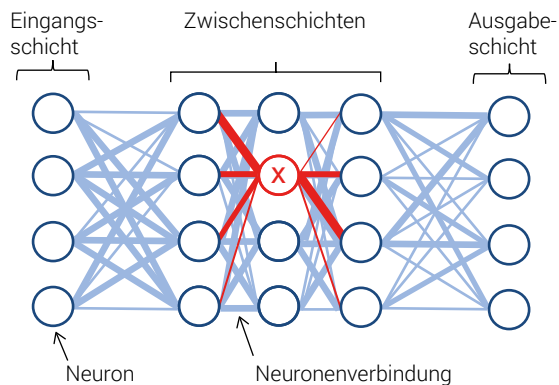
Die künstlichen neuronalen Netzwerke des Deep Learning bestehen aus einer grossen Menge an simulierten Neuronen (künstliche Nervenzellen), die in Schichten organisiert sind und untereinander Ladungssignale austauschen (G2). Die erste Neuronenschicht wird dabei als *Eingangsschicht* bezeichnet. Hier werden die Input-Daten «abgelegt», die vom künstlichen neuronalen Netzwerk untersucht werden sollen. Im Fall der Luftbild-Interpretation der Arealstatistik entspricht die Eingangsschicht den insgesamt 120 000 Farbwerten für die 40 000 Pixel des untersuchten Bildausschnitts (vgl. Kapitel 4.1). Jeder einzelne gemessene Farbwert wird somit als Ausgangssignal eines Neurons der Eingangsschicht eingesetzt.

In den *Zwischenschichten* des Netzwerks findet die eigentliche Datenverarbeitung statt. Die Zahl dieser Schichten kann sehr gross und das Netzwerk somit entsprechend «tief» sein – daher auch die Bezeichnung «Deep Learning». Moderne Deep-Learning-Netzwerke können zudem komplizierte Faltungsstrukturen aufweisen, die die Leistung des Netzwerks zusätzlich vergrössern. Dies ist auch bei dem für die Arealstatistik verwendeten Typ des «Convolutional Neural Network» der Fall.

Ein bestimmtes Neuron «X» in einer der Zwischenschichten erhält immer Signale von Neuronen aus der jeweils vorgelagerten Neuronenschicht. Je nachdem, wie stark die vorgelagerten Neuronen «senden» und wie stark die Verbindungen zum Neuron X sind, ist die Summe der Signale, die bei X eintreffen, grösser oder auch kleiner. Das empfangene Ladungssignal wird in der Folge vom Neuron X verstärkt oder auch abgeschwächt und an die Neuronen der nächsten Schicht weitergeleitet. Dabei wird das Signal aufgesplittet: Je nach Stärke der Verbindung erhält ein bestimmtes Neuron der nächsten Schicht einen grösseren oder auch kleineren Anteil des von X ausgehenden Ladungssignals. Entscheidend dabei ist: Die Stärke jeder einzelnen Verbindung im künstlichen neuronalen

Künstliches neuronales Netzwerk mit drei Zwischenschichten

G2



Quelle: BFS – Arealstatistik (AREA)

© BFS 2023

Netzwerk kann separat variieren, und Gleiches gilt auch für die «Signalverstärkungs- oder Abschwächungsleistung» eines jeden Neurons. Es gibt also zahllose kleine «Hebel», mit denen der Signalfluss durch die verschiedenen Schichten des Netzwerks verändert und somit bestimmt werden kann, welche Teile des von der Eingangsschicht ausgehenden Signals bei einem bestimmten Neuron der letzten Schicht eintreffen.

Diese letzte Schicht ist die sogenannte *Ausgabeschicht*. Ihre Neuronen entsprechen im Fall der Arealstatistik den einzelnen Kategorien der Bodennutzung bzw. -bedeckung. Sie besteht also aus 46 Neuronen bei dem für die Bestimmung der Bodennutzung und aus 27 Neuronen bei dem für die Bestimmung der Bedeckung verwendeten Netzwerk. Das Ziel ist nun, die Verstärkungs- oder Abschwächungsleistung der zahlreichen Neuronen in den Zwischenschichten und die Stärke der Verbindungen zwischen diesen so einzustellen («parametrieren»), dass bei jedem möglichen Eingangs-Luftbild immer dasjenige Neuron in der Ausgabeschicht das stärkste Signal empfängt, das für die korrekte Nutzungs- oder -bedeckungskategorie steht. Wenn dem neuronalen Netzwerk zum Beispiel ein Luftbildausschnitt mit einem Stichprobenpunkt auf einem See als Input gereicht wird, dann sollte in der Ausgabeschicht möglichst nur dasjenige Neuron, das für die Nutzungskategorie der «stehenden Gewässer» steht, ein Signal empfangen und mithin «aktiviert» werden.

Lernprozess

Die genannten Einstellungen in den Zwischenschichten enthalten somit die «Regeln», nach denen das Netzwerk die Eingangsdaten interpretiert. Diese Regeln werden dem neuronalen Netzwerk allerdings nicht etwa von einem menschlichen Programmierer oder einer Programmiererin vorgegeben, sondern vom neuronalen Netzwerk selbst ausgearbeitet. Genau darin besteht seine «künstliche Intelligenz».

Das Finden dieser Regeln setzt ein langwieriges und äusserst rechenintensives «Training» mit enormen Mengen von durch Menschen vorinterpretierten Daten voraus. Vor Beginn des Trainings wird allen Neuronenverbindungen im Netzwerk rein zufällig eine gewisse Stärke zugeteilt, und ebenso zufällig wird auch die Signalverstärkungs- oder Abschwächungsleistung der einzelnen Neuronen eingestellt. Danach lässt man das Netzwerk die Trainingsdaten nach diesen zufälligen Regeln abarbeiten. Im Fall der Arealstatistik wurden dem Netzwerk Millionen von Luftbildausschnitten aus früheren Erhebungen gereicht. Die von dem rein zufällig parametrisierten Netzwerk vorgenommenen Kategorisierungen sind als Bildinterpretationen natürlich völlig unbrauchbar. Doch das ist vorderhand irrelevant, denn das Ziel des ersten Datendurchlaufs ist ein anderes: Nach dessen Abschluss kann bei jedem Luftbild für jedes einzelne Neuron der Ausgabeschicht festgestellt werden, wie stark der berechnete Wert vom echten, also menschlich bestimmten Wert abweicht. Auf Basis

dieser «Fehlerinformationen» kann das Netzwerk sich dann im Zuge eines weiteren, noch aufwendigeren Rechenprozesses adäquat parametrieren oder anders ausgedrückt: lernen. Hierzu geht das Netzwerk nun genau den umgekehrten Weg und arbeitet sich von der Ausgabeschicht Richtung Eingabeschicht durch («Rückpropagierung» der gefundenen Fehler). Das Regelwerk, das dabei auf den Zwischenschichten «angelegt» wird, ist dermassen umfangreich und kompliziert, dass es für den Menschen nur ansatzweise nachvollziehbar ist.

Um den Lernprozess zu verkürzen und zugleich die Qualität der Bildinterpretationen zu steigern, wird bei der Arealstatistik ein neuronales Netzwerk verwendet, das von dessen Hersteller schon einem Grundtraining mit anderen Bilddaten unterzogen worden war. Zum Zeitpunkt des Trainingsbeginns mit den Arealstatistik-Daten war es daher bereits in der Lage, nützliche Bildmerkmale wie Kanten oder Ecken zu erkennen, und konnte dieses «Wissen» auf die Luftbilder der Arealstatistik übertragen (sogenanntes «Transfer Learning»).

«Deep Learning» funktioniert. Nähere Erklärungen zu dieser speziellen Form der künstlichen Intelligenz finden sich im Kasten auf den Seiten 4 und 5.

Wird dem trainierten Deep-Learning-Modul ein bestimmter Stichprobenpunkt zur Interpretation vorgelegt, so steht am Ausgang des Rechenprozesses allerdings nicht etwa eine abschliessende Antwort im Sinne von: «Beim Punkt X liegt die Nutzungskategorie A vor». Vielmehr liefert das Programm für jede der 46 Bodennutzungs- und 27 Bodenbedeckungskategorien eine separate Wahrscheinlichkeitsangabe (Prädiktionswahrscheinlichkeit). Dieses Ergebnis kann dann als Input für weitere Rechenschritte verwendet werden.

4.2 Einbezug von Hilfsdaten mittels Random-Forest-Modul

Wie eingangs gesehen, können die menschlichen Interpretinnen und Interpreten für die korrekte Interpretation der Luftbilder eine Reihe von Hilfsinformationen hinzuziehen. Gleiches gilt auch für ADELE: Die Ergebnisse der Bildinterpretation durch das Deep-Learning-Modul werden in einem weiteren Rechenschritt mit zusätzlichen Informationen kombiniert. Dadurch kann die Präzision der Kategorisierung weiter gesteigert werden. Gegenwärtig berücksichtigt ADELE die folgenden Hilfsdaten:

- altimetrische Daten: Höhe über Meer, Hangneigung, Vegetationshöhe usw.;
- bestimmte Daten aus dem topografischen Landschaftsmodell von swisstopo;
- per Satellit gemessene Farbwerte (unter anderem Infrarot) am Stichprobenpunkt zu verschiedenen Jahreszeiten;
- weitere Geodaten: amtliche Vermessung, Bauzonen, landwirtschaftliche Nutzungsflächen usw.

Für die Kombination der Deep-Learning-Bildinterpretation mit den verschiedenen Hilfsdaten wird separat ein anderes Verfahren des maschinellen Lernens namens «Random Forest» eingesetzt (für Erklärungen siehe Kasten auf Seite 6). Dabei werden drei separate Modelle trainiert:

- Das erste Random-Forest-Modell bestimmt die Bodennutzung am Stichprobenpunkt.
- Das zweite Modell bestimmt die Bodenbedeckung.
- Das dritte Modell entscheidet lediglich, ob sich die Bodennutzung und -bedeckung am Stichprobenpunkt gegenüber der letzten Erhebung verändert hat (ja/nein). Ihm wird somit eine einfachere Aufgabe gestellt als den ersten beiden Modellen. Entsprechend treffender sind seine Vorhersagen (höhere Prädiktionswahrscheinlichkeit).

Random Forest

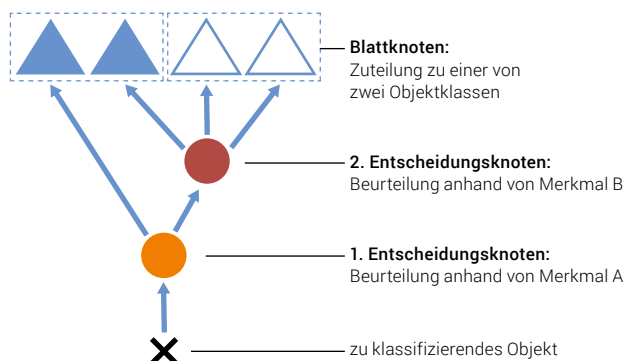
Als «Random Forest» wird ein Verfahren des maschinellen Lernens bezeichnet, das häufig eingesetzt wird, wenn Objekte anhand bestimmter Merkmale «klassifiziert», also einer bestimmten Klasse von Objekten zugeordnet werden sollen. Das Verfahren ist besonders stark darin, sehr unterschiedliche Objekt-Merkmale zu verarbeiten, die zueinander nur in einem losen Zusammenhang stehen. Genau dies ist der Fall bei den sehr heterogenen Hilfsdaten (siehe Kapitel 4.2), die für die Zuteilung der Arealstatistik-Stichprobenpunkte zu den Kategorien der Bodennutzung und Bodenbedeckung herangezogen werden.

Funktionsweise

«Random Forest» bedeutet auf Deutsch «Zufallswald». «Wald» daher, weil das Verfahren ein bestimmtes Objekt – im Falle der Arealstatistik einen Stichprobenpunkt – klassifiziert, indem es die Entscheidungen einer Vielzahl einzelner «Entscheidungsbäume» zu Rate zieht, die zusammen gleichsam einen «Wald» bilden. Dabei wird geschaut, welcher Klasse die meisten Entscheidungsbäume des Waldes das fragliche Objekt zuordnen.

Entscheidungsbäume sind Verkettungen von Wenn-dann-Entscheidungen, in die alle bekannten Merkmale eines Objekts einfließen können: Liegt ein bestimmter Arealstatistik-Stichprobenpunkt gemäss den Hilfsdaten aus der amtlichen Vermessung auf einem Gebäude? Wenn ja, dann dürfte es sich um den Bodenbedeckungstyp «Gebäude» handeln. Wenn nein, dann wird ein weiteres Merkmal des Stichprobenpunkts zu Rate gezogen usw. Werden diese Entscheidungsketten grafisch dargestellt (G3), ergeben sich aufgrund der vielen Verästelungen baumartige Gebilde – daher die Bezeichnung «Entscheidungsbaum». Im Fall von ADELE fliessen insgesamt 280 verschiedene Merkmale in die Entscheidungsbäume ein. Zum einen handelt es sich dabei um die 73 von der Deep-Learning-Luftbildinterpretation (siehe Kapitel 4.1) generierten Wahrscheinlichkeitsangaben zu den 46 Bodennutzungs- und 27 Bodenbedeckungskategorien, zum anderen um insgesamt 207 zusätzliche Merkmale aus diversen Hilfsdaten.

Beispiel eines einfachen Entscheidungsbaums G3



Quelle: BFS – Arealstatistik (AREA)

© BFS 2023

Lernprozess

Die Entscheidungsbäume müssen mittels grosser Datenmengen aus früheren Erhebungen trainiert werden: Auf Basis der zu den Stichprobenpunkten verfügbaren Merkmalsinformationen zum einen und den menschlichen Klassenzuteilungen zum anderen wird die Ast-Struktur der Entscheidungsbäume so lange optimiert, bis die «Entscheidungen» der Bäume denen der menschlichen Interpreten möglichst nahekommen.

Beim Training der zahlreichen Entscheidungsbäume wird darauf geachtet, dass sich deren Aststrukturen relativ stark voneinander unterscheiden: Der «Zufallswald» soll nicht einer Monokultur aus Entscheidungsbäumen des immer gleichen Typs gleichen, sondern einer Mischkultur aus sehr unterschiedlichen Baumarten. Dadurch sind die endgültigen Klassifikationen, für die ja die Entscheidungen aller Bäume des Waldes berücksichtigt werden, breiter abgestützt. Um einen «vielfältigen» Wald zu erzeugen, wird die Reihenfolge und somit die Bedeutung der verschiedenen Merkmale in den einzelnen Entscheidungsbäumen mittels Zufallsauswahl variiert – daher die Bezeichnung «Zufallswald».

5 Verwendung von ADELE in der Praxis

5.1 Leistungsgrenzen und entsprechende Massnahmen

Bisher wurde erläutert, wie das KI-Tool ADELE «im Prinzip» funktioniert. In der Praxis ist es jedoch nicht möglich, die Bestimmung der Bodennutzung und -bedeckung vollständig ADELE zu überlassen. Dies vor allem aus zwei Gründen:

Erstens erreicht das Tool für einen beträchtlichen Teil der Kategorien nicht das Präzisionsniveau, das für eine kohärente Fortschreibung der Arealstatistik-Zeitreihen notwendig ist. Obwohl ADELE laufend verbessert wird, dürfte sich daran auch mittel- bis langfristig nichts ändern: Dafür sehen sich einige

Kategorien optisch schlicht zu ähnlich, während andere – wie zum Beispiel die Parkanlagen – eine vom Tool kaum zu beherrschende innere Vielfalt an Formen und Farben aufweisen. Folglich muss mittels Tests herausgefunden werden, welche Nutzungs- und Bedeckungskategorien von ADELE grundsätzlich mit der erforderlichen Präzision erkannt werden können. Nur diese Teile des ADELE-Outputs dürfen dann in die publizierten Endergebnisse der Arealstatistik einfließen. Ausserdem können nur jene Stichprobenpunkte berücksichtigt werden, bei denen die von ADELE selbst ermittelte Wahrscheinlichkeit einer korrekten Bestimmung eine bestimmte Schwelle überschreitet. Alle übrigen Stichprobenpunkte müssen weiterhin von Menschen interpretiert werden.

Zweitens wird eine vollständige Automatisierung der Nutzungs- und Bedeckungsbestimmung ganz ohne menschliche «Begleitung» aus prinzipiellen Überlegungen auch gar nicht angestrebt. ADELE muss auch in Zukunft immer wieder mit frischen Trainingsdaten in Form menschlich interpretierter Kategorisierungen versorgt werden, um:

- dem System die Verwendung gänzlich neuer oder zumindest neu strukturierter Hilfsdaten beizubringen;
- zu verhindern, dass sich gewisse Interpretationsfehler im System über die Jahre akkumulieren;
- die ADELE-Entscheidungsregeln laufend an die «physischen Veränderungen» der Kategorien selbst anzupassen. Die «Industrie- und Gewerbeareale» beispielsweise sahen vor einigen Jahrzehnten auf den Luftbildern noch deutlich anders aus als heute und dürften sich auch in Zukunft «visuell weiterentwickeln». Entsprechend werden die von ADELE auf Basis heutiger Trainingsdaten entwickelten Entscheidungsregeln irgendwann zur Identifizierung dieser Nutzungskategorie nicht mehr geeignet sein.

Um dem zweiten Problemkreis zu begegnen und somit die Qualität der Erhebung auf lange Sicht zu gewährleisten, muss eine gewisse Anzahl von Stichprobenpunkten ausgeschieden werden, bei denen die Bestimmung der Bodennutzung und -bedeckung weiterhin immer von Menschen vorgenommen wird. Konkret wurden für die Arealstatistik insgesamt knapp 1 Million solcher Punkte definiert.

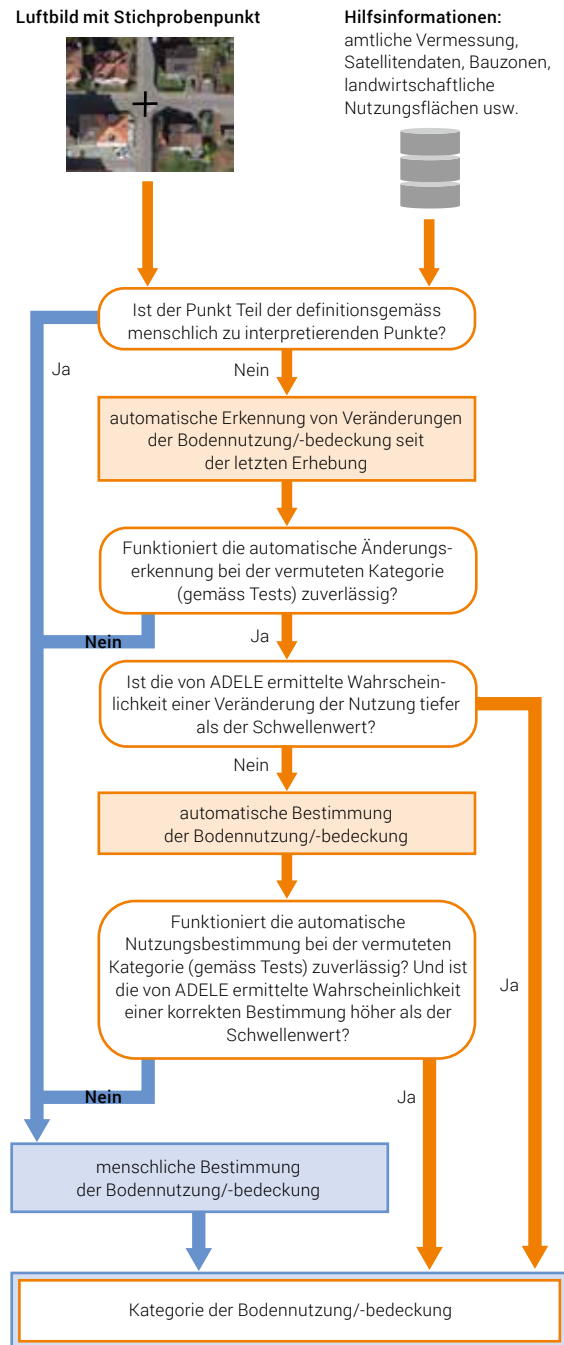
5.2 Ablauf der halbautomatisierten Nutzungsbestimmung

Wie soeben gesehen, muss der Einsatz von ADELE wohlüberlegt eingegrenzt und mit der Arbeit der menschlichen Interpretinnen und Interpreten kombiniert werden, um eine gleichbleibende Qualität der statistischen Ergebnisse zu garantieren. Hierzu wurde ein genauer Prozess definiert (G4). Dieser schreibt vor, dass zunächst geprüft wird, ob es sich bei dem zu bestimmenden Stichprobenpunkt um einen der rund 1 Million Punkte handelt, die aus Gründen der langfristigen Qualitätssicherung notwendigerweise menschlich interpretiert werden müssen (vgl. Kapitel 5.1). Ist dies nicht der Fall, wird von ADELE automatisch geprüft, ob sich die Bodennutzung und -bedeckung an dem Punkt seit der letzten Erhebung verändert hat, wobei die Antwort zunächst nur Ja oder Nein lauten kann. Für die Änderungserkennung wird ein ausschliesslich für diese Aufgabe trainierter Random-Forest-Algorithmus (vgl. Kapitel 4.2) eingesetzt, der sich auf die Ergebnisse des Deep-Learning-Moduls sowie auf verschiedene Hilfsdaten (in diesem Fall vor allem Satellitenbilder) stützt. Kann eine Veränderung mit grosser Sicherheit ausgeschlossen werden, werden die in der letzten Erhebung bestimmten Kategorien definitiv eingesetzt und die automatische Bestimmung ist abgeschlossen. Besteht dagegen eine gewisse Veränderungswahrscheinlichkeit, schliesst sich eine automatische Bestimmung der Nutzung an. Deren Ergebnis wird allerdings nur validiert, falls die vermutete Kategorie für eine automatische Bestimmung überhaupt freigegeben wurde, nachdem Tests gezeigt hatten,

Halbautomatisierte Bestimmung der Bodennutzung und -bedeckung in der Arealstatistik: Ablauf des Prozesses

G4

Die orange markierten Schritte übernimmt das KI-Tool ADELE



Quelle: BFS – Arealstatistik (AREA)

© BFS 2023

dass ADELE in ihrem Fall grundsätzlich zuverlässig funktioniert. Ausserdem muss die von ADELE ermittelte Wahrscheinlichkeit einer korrekten Bestimmung beim fraglichen Stichprobenpunkt einen bestimmten Schwellenwert überschreiten. Sind diese Bedingungen nicht erfüllt, wird die Bodennutzung und -bedeckung weiterhin menschlich bestimmt.

6 Ergebnis und Ausblick

Der Einsatz des Deep-Learning-Tools ADELE in den Arealstatistik ermöglicht gegenwärtig – Stand 2023 – die automatische Bestimmung von ungefähr 27% der Stichprobenpunkte. Dies entspricht rund 1,1 Millionen Punkten, die nicht mehr von Menschen interpretiert werden müssen, und somit einer ganz erheblichen Zeitersparnis. Zu beachten ist dabei, dass bisher ausschliesslich die automatische Veränderungserkennung im Einsatz steht und ADELE noch keine direkten Kategorisierungen vornimmt.

In Zukunft soll der Anteil der automatisch bestimmten Punkte dank verschiedener Massnahmen vergrössert werden:

- Verwendung zusätzlicher Daten: Angaben aus dem Gebäuderegister, Infrarot-Luftbilder, längere Zeitreihen bei den Satellitenbildern, Angaben zur Bodennutzung und -bedeckung bei benachbarten Stichprobenpunkten usw.
- Verwendung eines besser vortrainierten Neuronen-Netzwerks beim Deep-Learning-Modul.
- Optimierung der sogenannten «Hyperparameter» bei den verwendeten KI-Algorithmen: beispielsweise bessere Einstellung der Anzahl Neuronen(-schichten) beim Deep-Learning-Modul und der Anzahl Entscheidungsbäume bei den Random-Forest-Modellen.

Zu einer Erhöhung des Automatisierungsgrads wird im Laufe der Zeit auch die Tatsache beitragen, dass das Volumen der Trainingsdaten mit jeder neuen Erhebung weiter zunimmt. Dennoch dürfte es schwierig sein, dereinst deutlich mehr als die Hälfte der Punkte automatisch zu bestimmen. Allerdings ist dieser «Grenzwert» auch einer Ressourcenabwägung geschuldet: Mit zunehmendem Automatisierungsgrad wird es nämlich immer schwieriger und somit aufwendiger, diesen noch weiter anzuheben. Ab einem gewissen Punkt ist es dann effizienter, die für eine weitere Automatisierung nötigen Ressourcen direkt auf die menschliche Interpretation der Stichprobenpunkte zu verwenden. Dabei ist auch zu bedenken, dass der menschliche Umgang mit den Grunddaten – die massenhafte Sichtung der Luftbilder am Bildschirm, die Diskussionen im Team, die eingangs beschriebenen Geländebegehungen – nicht nur der Kategorisierung an sich dienen. Aus der menschlichen Interpretationsarbeit heraus entwickelt sich auch ein tieferes Verständnis für die Daten und Kategorien der Arealstatistik und damit eine Expertise, wie sie für eine korrekte Interpretation der statistischen Ergebnisse unverzichtbar ist.

Herausgeber:	Bundesamt für Statistik (BFS)
Auskunft:	Sektion Geoinformation, BFS, arealstatistik@bfs.admin.ch
Inhalt:	Gillian Milani, BFS; Romain Douard, BFS; Ernest Hennig, BFS; Lena Willi-Tobler, BFS; Michèle Gillard, BFS; Anton Beyeler, BFS
Redaktion:	Ferenc Biedermann, BFS
Reihe:	Statistik der Schweiz
Themenbereich:	02 Raum und Umwelt
Originaltext:	Deutsch
Übersetzung:	Sprachdienste BFS
Layout:	Publishing und Diffusion PUB, BFS
Grafiken:	Publishing und Diffusion PUB, BFS
Bilder:	Luftbild: © swisstopo; Foto: © Arealstatistik
Online:	www.statistik.ch
Print:	www.statistik.ch Bundesamt für Statistik, CH-2010 Neuchâtel, order@bfs.admin.ch , Tel. +41 58 463 60 60 Druck in der Schweiz
Copyright:	BFS, Neuchâtel 2023 Wiedergabe unter Angabe der Quelle für nichtkommerzielle Nutzung gestattet
BFS-Nummer:	2261-2300

Die Informationen in dieser Publikation tragen zur Messung der Ziele für nachhaltige Entwicklung (SDG) bei.



Indikatorensystem MONET 2030

www.statistik.ch → Statistiken finden → Nachhaltige Entwicklung → Das MONET 2030-Indikatorensystem