Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Dokumentation «Prototyp AI»

Bild: swisstopo

Auftraggeber	Bundesamt für Statistik
Projektleiter	Prof. Dr. Denis Jordan, FHNW
Autoren	Denis Jordan, Adrian Meyer, Natalie Lack, Maria Schönholzer (FHNW) Reik Leiterer, Gillian Milani (ExoLabs GmbH)
Klassifizierung	Nicht klassifiziert
Status	Abgeschlossen

Änderungsverzeichnis

Datum	Version	Änderung
26.02.2019	1.0	Draft - Dokumentation «Prototyp AI»
05.04.2019	1.1	Revision 1 - Dokumentation «Prototyp AI»
01.05.2019	2.0	Revision 2 - Dokumentation «Prototyp AI»
10.05.2019	3.0	Final – Dokumentation «Prototyp AI»

Inhalt

1 Ausgangslage 5
1.1 Motivation für den Prototyp AI 6
1.2 Rahmenbedingungen für den Prototypen 7
1.3 Voraussetzungen für den Prototypen AI 10
1.3.1 Aktuelles Vorgehen Arealstatistik 10
1.3.2 Grundlagedaten Arealstatistik 13
1.3.3 Infrastruktur Hardware/Software 14
2 Prototyp AI 15
2.1 Ziele 15
2.2 Konzept 15
2.3 Datenaufbereitung 16
2.3.1 Datenaufbereitung Luftbilder 17
2.3.2 Datenaufbereitung Satellitenbilder 19
2.3.3 Datenaufbereitung Hilfs- und Zusatzdaten 21
2.3.4 Performance der Datenvorbereitungsmodule 23
2.4 CNN 23
2.4.1 Systemvoraussetzungen 23
2.4.2 Funktionsweise 24
2.4.3 Datenvorbereitung und Sampling 25
2.4.4 Mögliche Einstellungen/Handling 25
2.5 Post-Klassifikator: Random Forest 27
2.5.1 Methodenwahl 27
2.5.2 Systemvoraussetzungen 28
2.5.3 Funktionsweise 28
2.5.4 Datenvorbereitung und Sampling 28
2.5.5 Mögliche Einstellungen/Handling 31
2.5.6 Parameter-Optimierung 32
2.5.7 Experimentelles Vorgehen 33
3 Sensitivitätsanalysen 35
4 Ergebnisse 37
4.1 Validierung 37
4.2 CNN 43
4.2.1 Experimentelle Ergebnisse 43
4.2.2 Zusammenfassung Erkenntnisse 47
4.3 Ergebnisse RF 49
4.3.1 Land Cover AREA3 49
4.3.2 Land Cover AREA4 50
4.3.3 Land Cover AREA4 mit Zusatzdaten 52
4.3.4 Land Cover – Vergleich Klassifikationsmodelle 54
4.3.5 Change Detection Land Cover 55
4.3.6 Land Use AREA3 57
4.3.7 Land Use AREA4 59
4.3.8 Land Use AREA4 mit Zusatzdaten 62

4.3.9 Change Detection Land Use 63
4.3.10 Kombinierte LU/LC-Klassifikation 65
4.3.11 Variablengewichte der Satellitenkanäle im Vergleich 67
4.3.12 Variablengewichte der AV-Daten im Vergleich 68
4.3.13 Trainingsdauer 68
5 Empfehlungen und Optimierungspotential 69
5.1 Datengrundlage/Datenverwendung 69
5.2 CNN 70
5.3 RF 71
5.4 Sensitivitätsanalysen und Validierung 72
6 Schlussfolgerungen und Ausblick 73
7 Literatur 75

Glossar

AI	Artificial Intelligence
ALTI3D	Digitales Geländemodell der Schweiz
AP	Arbeitspaket
API	Application Programming Interface
AREA	Arealstatistik
AV	Amtliche Vermessung
B	Blau
BFS	Bundesamt für Statistik
CD	Change Detection
CHM	Canopy Height Model
CUDA	Compute Unified Device Architecture
CNN	Convolutional Neural Network
CPU	Central Processing Unit
CSV	Comma-separated Values
DS	Data Samples
FCIR	NIR – Rot – Grün (Bildkanal-Kombination)
DEM	Digital Elevation Model
DL	Deep Learning
DNN	Deep Neural Network
DOM	Digitales Oberflächenmodell
DSM	Digital Surface Model
DTM	Digital Terrain Model
EVI	Enhanced Vegatation Index
F	F-Mass
FHNW	Fachhochschule Norstwestschweiz
G	Grün
GEE	Google Earth Engine
GLM	Generalized Linear Model
GPU	Graphics Processing Unit
GWR	Eidgenössische Gebäude- und Wohnungsregister
HDF	Hierarchical Data Format
IDLE	Integrated Development and Learning Environment
IGEO	Institut Geomatik der FHNW
KK	Kappa Koeffizient
kNN	K-Nearest-Neighbor
LC	Land Cover
LFI	Schweizerisches Landesforstinventar
LK	Landeskarte
LU	Land Use
LULC	Land Use/ Land Cover
M	Variabel-Knotensplits
ML	Machine Learning
MNDWI	Modification of Normalized Difference Water Index
MS	Meilenstein
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NIR	Nahes Infrarot
OA	Overall Accuracy
ÖREB	Kataster der öffentlich-rechtlichen Eigentumsbeschränkungen
PA	Producer’s Accuracy
R	Rot
RAM	Random-Access Memory
RGB	Rot – Grün – Blau (Bildkanal-Kombination)
RF	Random Forest
RS	NIR – Rot – Grün – Blau (Bildkanal-Kombination)
SGD	Stochastic Gradient Descent Optimizer
STATENT	Statistik der Unternehmensstruktur
STATPOP	Statistik der Bevölkerung und der Haushalte
SVM	Support Vector Machine
SWIR	Short Wavelength Infrared
TIR	Thermal Infrared
TLM	Topografisches Landschaftsmodell
UA	User‘s Accuracy
UI	User Interface
WSL	Eidgenössische Forschungsanstalt für Wald, Schnee und Landschaft

Zusammenfassung

Im vorliegenden Bericht zum Pilotprojekt „IA pour la statistique de la superficie“ ist die Entwicklung und Evaluation eines Prototypen („Prototyp AI“) dokumentiert, welcher die Anwendung von Methoden des maschinellen Lernens im Rahmen des Erhebungsprozesses der Arealstatistik ermöglicht. Der Prototyp AI ist hierbei so konzipiert, dass der vollständige Ablauf, von der Datenvorverarbeitung bis hin zur Datenauswertung, automatisiert durchgeführt werden kann, aber durch die offene Gestaltung des Quellcodes auch eine unkomplizierte Anpassung an spezifische Nutzerbedürfnisse ermöglicht ist.

Der aktuelle Erhebungsprozess der Arealstatistik ist charakterisiert durch die hochkomplexe Datengrundlage (in Hinblick auf Datenumfang und -heterogenität), eine semantisch anspruchsvolle Struktur der Zielklassen (in Hinblick auf Klassentiefe und zum Teil offen gehaltene Klassendefinitionen - vgl. Kapitel 1.3.1) und einem sehr hohen Anteil des manuellen Arbeitsaufwandes. Zudem besteht der Anspruch, die Klassenzuweisung transparant, reproduzierbar und mit einer sehr hohen Zuverlässigkeit auszuführen.

Dementsprechend wurde der Prototyp AI so entwickelt, dass er die zur Verfügung stehenden Geodaten einlesen, harmonisieren und miteinander verknüpfen kann. Dies umfasst auch Daten, die bisher keine systematische Verwendung fanden (z.B. Vegetationshöhenmodelle) oder erst für den nächsten Erhebungsprozess zur Verfügung stehen werden (z.B. die Sentinel2-Satellitendaten). Als Methoden des maschinellen Lernens wurde ein Deep Learning Klassifikator basierend auf einer *CNN* – Struktur (Convolutional Neural Network) und ein zusätzlicher *Random Forest* implementiert. Diese Verfahren gehören zu den zur Zeit leistungsstärksten Ansätzen der Bildauswertung und zeigten in vorläufigen Studien (vgl. Hochuli 2018) bereits vielversprechende Ergebnisse im Kontext der Klassifikation von Landbedeckung und -nutzung. Um die Auswertung der für das maschinelle Lernen typischen Ergebnisse (= Wahrscheinlichkeiten der Klassenzuweisung) zu erleichtern, wurde zudem ein Validierungsmodul entwickelt, welches die üblichen und schnell interpretierbaren Genauigkeitsmetriken in tabellarischer und grafischer Form zur Verfügung stellt. Der Prototyp AI ist in seiner Funktionsweise sowohl umfassend im vorliegenden Bericht, auf der Ebene des Quellcodes im Dokument «Anhang B – Prototyp AI» und im Quellcode selber dokumentiert.

Die Funktionsweise des Prototyps AI wurde dahingehend evaluiert, dass verschiedene Klassifikationsdurchläufe ausgeführt und deren Ergebnisse hinsichtlich Plausibilität, Stabilität und Zuverlässigkeit interpretiert wurden. Im Vergleich zu Studien mit ähnlichen Klassifikationsansätzen, konnten hohe Gesamt-Genauigkeiten und zum Teil sehr hohe Klassengenauigkeiten erreicht werden. Insbesondere für Klassen der Landbedeckung/-nutzung, die nicht den urbanen Flächen zugeschrieben werden, zeigten die Klassifikatoren eine vielversprechende Performance. Darüber erlaubt der Prototyp AI auch verschiedene Ansätze zur direkten Ausweisung von zwischen zwei Zeitpunkten gleichbleibenden bzw. sich verändernden Klassenzuweisungen. Die vorgestellten Ergebnisse sind hierbei aber eher als Tests der Funktionsweise des Prototyps AI zu verstehen, da davon ausgegangen werden kann, dass sich die jeweiligen Klassifikationsergebnisse in Abhängigkeit einer optimierten Parametrisierung noch deutlich verbessern werden. Eine entsprechend strukturierte Vorgehensweise zur Parametrisierung und Validierung wird im vorliegenden Bericht diskutiert.

Das BFS verfügt mit dem Prototyp AI über einen als einzigartig einzustufenden Workflow, welcher auch ausserhalb der Arealstatistik Anwendung im Bereich der Klassifizierung von Geodaten finden kann. Die dynamische und offene Struktur erlaubt neben dem Hinzufügen zusätzlicher Datenquellen und möglicher parallelisierter Klassifikatoren auch die Anpassung der thematischen Ebene (= flexible Klassenausweisungen). Zudem kann die räumliche Dimension ohne nennenswerten Aufwand beliebig angepasst werden (z.B. Änderung der räumlichen Distanz zwischen den Stichprobenpunkten oder die zu berücksichtigende Fläche um den Stichprobenpunkt).

Wir sehen den Prototyp AI als Start einer Reise – und hoffen auf eine vielfältige und zufriedenstellende Anwendung!

1 Ausgangslage

1.1 Motivation für den Prototyp AI

Die Arealstatistik wird seit den 1980er Jahren vom Bundesamt für Statistik (BFS) periodisch alle 12 Jahre, die aktuelle Erfassung 2013/18 mit einem verkürzten Abstand von 9 Jahren und ab 2019 in einer rollenden Erhebung alle 6 Jahre bereitgestellt. Sie dient zur Erfassung des Zustandes der Landbedeckung (Land Cover = LC) und Landnutzung (Land Use = LU) und bildet ein unverzichtbares Instrument der räumlichen Langzeitbeobachtung (Bundesamt für Statistik 2018b). Grundlage der Arealstatistik sind 4.1 Millionen Stichprobenpunkte (= Hektarpunkte) mit einem regelmässigen Raster von 100m Maschenweite, welche primär auf Basis von Luftbilddaten von einem Team geschulter Interpret*innen nach einem Katalog von 46 LU-Kategorien und 27 LC-Kategorien beurteilt werden (Bundesamt für Statistik 2004).

Mit der rollenden Erhebung sollen zukünftig moderne Verfahren angewendet werden, um den Erfassungsprozess zumindest teilweise zu automatisieren. Ziel des Projektes "Prototyp AI" des BFS ist es, das Potenzial automatisierter Verfahren auf der Basis künstlicher Intelligenz abzuschätzen und zu diesem Zweck einen Software-Prototypen zu konzipieren, zu implementieren und hinsichtlich systematisch durchgespielter Konfigurationen zu evaluieren. Aufgrund der vom Auftragnehmer, der Fachhochschule Nordwestschweiz (FHNW) und der ExoLabs GmbH (ExoLabs), gelieferten Arbeiten soll das BFS den Einsatz von maschinellem Lernen für die nächste Erhebung der Arealstatistik hinsichtlich Nutzbarkeit beurteilen können und einen entsprechenden Einsatz realisieren.

Bisherige Untersuchungen haben darauf hingewiesen, dass sich durch Deep Learning (DL) eine vielversprechende Umsetzungsmethode des maschinellen Lernens zur (teil-)automatisierten Erfassung der Arealstatistik anbietet (Machbarkeitsstudie BFS-Nummer do-t-02.02.01-Study, Bachelor Thesen FHNW 2017-2018). DL hat, insbesondere in der Klassifizierung von Bilddaten, bereits eine herausragende Bedeutung in einer Vielzahl von Anwendungsfeldern einnehmen können und zeigt konzeptionelle Vorteile zu klassischen, auf dem Feature Engineering basierenden Verfahren des maschinellen Lernens. Hinsichtlich der vorliegenden Problemstellung der Arealstatistik sind jedoch die in vielen Anwendungsbereichen zum Einsatz kommenden vordefinierten und vortrainierten neuronalen Netze nicht direkt anwendbar, da diese zunächst für RGB-Bildmaterial konzipiert sind, und nicht für eine Prozessierung von Multisensor-Daten, in unterschiedlicher räumlicher und zeitlicher Auflösung, und zusätzlicher Daten, wie sie im Kontext der Arealstatistik erforderlich sind. Dazu bedarf es weiterführender Ansätze des maschinellen Lernens, die beispielsweise Deep Learning Netzwerkarchitekturen zulassen, mit welchen sich Verfälschungen aufgrund dünn besetzter Klassen adressieren lassen, oder mit denen sich Messdaten unterschiedlicher Qualität dem Klassifikator präsentieren und von diesem prozessieren lassen.

Die Wahl einer methodisch adäquaten und im Rahmen des Projektes "Prototyp AI" umsetzbaren Methode stellt einen wesentlichen Teil der Arbeit hinsichtlich einer erfolgreichen Umsetzung des zu realisierenden Prototyps dar. Dementsprechend werden im vorliegenden Bericht Konzepte und Implementierung des Prototyps aufgezeigt und die erzielten Ergebnisse dargestellt und diskutiert. Im Rahmen der inhaltlich und zeitlich sehr ambitionierten Projektarbeiten haben die FHNW und ExoLabs in Abstimmung mit dem BFS strategische ad-hoc Entscheidungen zur Umsetzung des Prototyps treffen müssen. Diese erfolgten jedoch in Gegenüberstellung mit alternativen Lösungen. Letztere bilden eine Grundlage für weiterführende und optionale projektbezogene Arbeiten, welche in einem gesonderten Abschnitt vorgestellt werden.

Ziele

Innerhalb der vom BFS festgelegten Vorgaben hinsichtlich funktioneller und technischer Spezifikationen des Prototyps sollen Varianten in Bezug auf (1) Netzwerkarchitekturen wie CNN und Random Forest, (2) die Behandlung schwach besetzter Klassen im Lernprozess, (3) die Möglichkeit adaptiver Netzwerkkonfigurationen, mit der sich der Lernalgorithmus anhand neuer Daten an die durch den Operateur festgelegte Ground Truth neu konfiguriert, (4) Methoden zur Präsentation von Zusatzdaten und (5) Möglichkeiten zur Implementierung einer LULC-Veränderungsanalyse diskutiert und definiert werden. Die statistischen Auswertungen sollen sich einerseits

an den klassischen Metriken, wie sie etwa in der Machbarkeitsstudie vorgestellt wurden, orientieren und darüber hinaus Verfahren anbieten, die einen aussagekräftigen Vergleich der Methoden ermöglichen. Ebenso soll der Funktionsumfang, insbesondere hinsichtlich Benutzer- und Datenschnittstellen, definiert werden.

1.2 Rahmenbedingungen für den Prototypen

Die Rahmenbedingungen für das vorliegende Projekt wurden in der Offerte abgesteckt und entsprechend der Entwicklungen und Bedürfnisse im Verlauf der Projektarbeit dynamisch angepasst. Das Projekt wurde hauptsächlich an den Standorten Institut Geomatik (IGEO) der FHNW (Muttenz) und bei ExoLabs (Zürich) durchgeführt. Die systemspezifischen Implementierungen wurden direkt im BFS und in Absprache mit den jeweils zuständigen Personen durchgeführt.

Innerhalb der FHNW war Denis Jordan primär für die Organisation und Beratung im Zusammenhang mit den Methoden des maschinellen Lernens, Natalie Lack mehrheitlich für die Unterstützung im Zusammenhang mit der Anwendung von Geodaten verantwortlich, Adrian Meyer und Maria Schönholzer führten die Arbeiten bezüglich Softwaredesign und Implementierung des Prototypen aus. Innerhalb von ExoLabs übernahm Reik Leiterer die Koordination mit der Projektleitung der FHNW und intern mit den Mitarbeitern von ExoLabs. Gillian Milani von ExoLabs war primär für die Datenvorverarbeitung, Produktevaluation sowie für die generelle Implementierung und Operationalisierung der Software verantwortlich. Denis Jordan war verantwortlich für die gesamte Projektleitung, Reik Leiterer übernahm die Stellvertretung. Beide stehen hinsichtlich der Koordination der Arbeiten der FHNW und von ExoLabs in regelmässigem Kontakt.

Die Kommunikation wurde hauptsächlich durch das mit Hermes bereit gestellte Projektmanagement organisiert. Der Austausch erfolgt über Internet Meetings oder vor Ort. Der Zeitplan bzw. die Projektstrukturierung wurden hinsichtlich der Ausschreibung (Dokument do-b-02-offre_2018-01) abgestimmt (vgl. Abbildung 1.1).

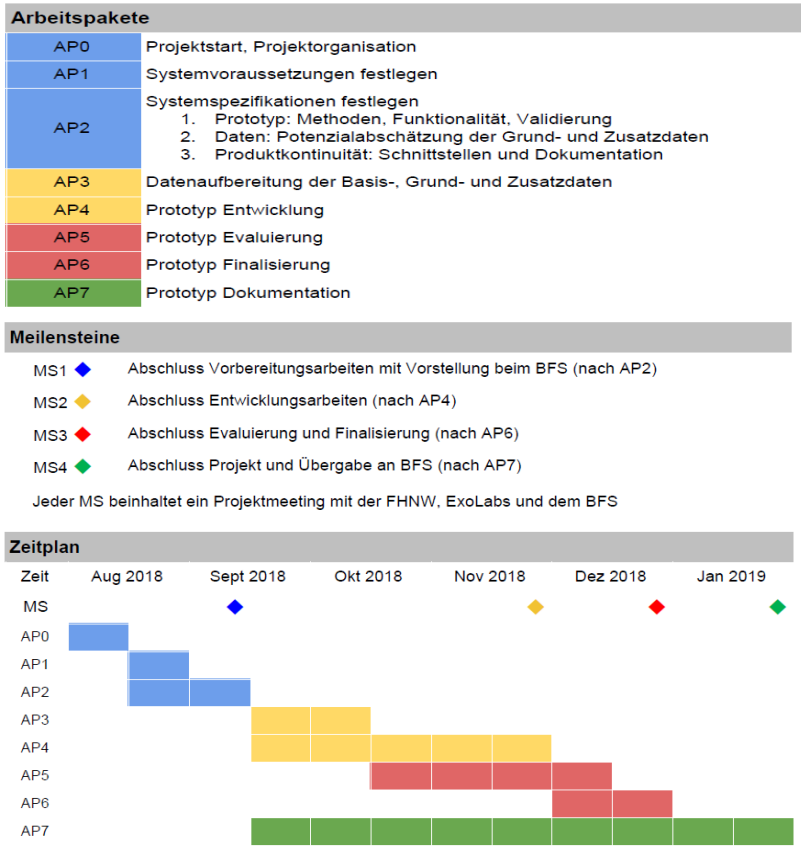


Abbildung 1.1. Projektstruktur und Zeitplan für das Projekt „Prototyp AI“.

Geringfügige Abweichungen wurden entsprechend der Wünsche von Seiten des BFS vorgenommen, die aber zu keinen Verzögerungen im Projektverlauf führten. Die Arbeitsaufteilung zwischen der FHNW und ExoLabs richteten sich nach den Angaben in den Arbeitspaketen (AP), wobei auftretende Abweichungen im Arbeitsaufwand über die in der Offerte verankerten Option kompensiert werden konnten. Projekttreffen in Bezug auf die Meilensteine (MS) wurden entsprechend der Offerte geplant und durchgeführt. Im Rahmen der Arbeiten für MS4 wurden zudem zusätzliche wöchentliche Telefonkonferenzen abgehalten. Nachfolgend sind die Inhalte und der Umfang der jeweiligen Arbeitspakete beschrieben:

AP0 Projektstart, Projektorganisation (Arbeitsaufwand 10 h)
Alle am Projekt mitarbeitenden Personen der FHNW/ExoLabs und das BFS wurden auf den gleichen Stand bezüglich der Ausgangsinformationen gebracht. Sowohl die Zusammenarbeit als auch die Rollen der einzelnen Mitglieder wurden definiert.

AP1 Systemvoraussetzungen festlegen (Arbeitsaufwand 40 h)
Auf der Basis eines innerhalb des Projektmanagements dokumentierten Informationsabgleichs wurden die Systemvoraussetzungen festgelegt. Dies betraf primär die Auswahl der Entwicklungsumgebungen, die im Rahmen der Prototypenentwicklung verwendet werden sollten.

AP2 Systemspezifikationen festlegen (Arbeitsaufwand 180 h)
Innerhalb der vom BFS festgelegten Vorgaben hinsichtlich funktioneller und technischer Spezifikationen des Prototyps wurden die Ziele und Systemspezifikationen des Prototyps diskutiert und definiert. Ebenso wurde der Funktionsumfang, insbesondere hinsichtlich Benutzer- und Datenschnittstellen, festgelegt. Für die gesamte Prototypenentwicklung entsprechend AP3-6 wurde ein Entwicklungsplan vorgelegt, der Umfang und funktionelle Anordnung der einzelnen Softwaremodule aufzeigte. Eine Übersicht über die Spezifikationen und die Verfügbarkeiten der Hilfsdaten wurde erstellt und hinsichtlich Nutzbarkeit innerhalb von AP3-5 diskutiert. Dies betraf insbesondere die Abschätzung des zu erwartenden Mehrwertes dieser Daten in Hinblick auf Verbesserung der Klassifikationsergebnisse und der Nutzbarkeit im operativen Ablauf des Prototyps. Letztendlich erfolgte die Einigung über den Rahmen für die Nutzerfreundlichkeit, einfache Wartung und das Weiterentwicklungspotenzial des Prototyps.

AP3 Datenaufbereitung der Basis-, Grund- und Zusatzdaten (Arbeitsaufwand 80 h)
Ein operationeller Ablauf zur Aufbereitung und Optimierung der Basis-, Grund- und Zusatzdaten wurde etabliert und in einem Softwaremodul umgesetzt. Die so generierten *ready-to-use* - Informationsebenen wurden in den Ablauf der jeweiligen Klassifikationsansätze integriert.

AP4 Prototyp Entwicklung (Arbeitsaufwand 500 h)
Der Prototyp wurde als quelloffenes Computerprogramm übergeben, welches (1) den funktionellen und technischen Spezifikationen gemäss BFS Ausschreibung genügte, (2) die in AP2 festgelegten Methoden des maschinellen Lernens und der Datenverarbeitung abbildete und (3) sowohl die Konfiguration und Steuerung der Eingangsdaten und Klassifikatoren als auch die Evaluation der Ergebnisse über eine Nutzerschnittstelle ermöglichte.

AP5 Prototyp Evaluierung (Arbeitsaufwand 80 h)
Es wurde eine quantitative Beurteilung der beiden Haupt-Prozessschritte, der Veränderungsdetektion und der LU/LC-Klassifikation, sowie der spezifischen Produkte der Klassifikation (Zuverlässigkeit der Erfassung der jeweiligen Kategorien in Hinblick auf deren thematischen Genauigkeit) mittels geeigneter Qualitätsindikatoren vorgenommen. Die Veränderungsdetektion und die Klassifikatoren wurden iterativ optimiert.

AP6 Prototyp Finalisierung (Arbeitsaufwand 40 h)
Die Umsetzung der an den Wünschen des BFS orientierten Optimierungen mit anschliessender Finalisierung des Prototyps wurde abgeschlossen.

AP7 Prototyp Dokumentation (Arbeitsaufwand 100 h)
Eine vollständige und für aktuelle sowie weiterführende Arbeiten zur Arealstatistik dem BFS zugängliche Projektdokumentation wurde erstellt und die Übergabe der gesamten Produktleistung an das BFS vorgenommen.

Der Auftragnehmer leistete alle Arbeiten gemäss der vorgelegten Offerte. Dies beinhaltete die von Seiten FHNW und ExoLabs beschriebenen Leistungen in AP0-7. Zusammenfassend handelte es sich um folgende Produkte:

- Prototyp Arealstatistik mit Datenschnittstelle und Benutzerinterface (Softwaremodul)
- Datenmanagement und -aufbereitung (Softwaremodul)
- Evaluierung/Validierung (Softwaremodul)
- Dokumentation aller Produkte (Projektbeschreibung und Dokumentation des Quellcodes)

Kostenzusammenstellung
Die Kostenzusammenstellung orientierte sich am Zeitplan und beruhte auf den in der Offerte aufgeführten Stundenansätzen (Abbildung 1.2). Koordination, Projektmanagement und ähnliches wurden in der folgenden Kostenaufstellung in den jeweiligen AP0-7 integriert. Darüber hinaus wurde teilweise die mit dem BFS vereinbarten Option in Anspruch genommen, um den bei der technischen Umsetzung höher ausgefallenen Arbeitsaufwand auszugleichen.

Kostenplan zur Offerte (zwingender Teil)

	Anzahl Std. pro Tarif (CHF/h)				Anteile		
Arbeitsschritt	**FHNW A** *CHF 133.80*	**FHNW B** *CHF 76.20*	**ExoLabs** *CHF 157.00*	**Std. Total**	**FHNW** %	**ExoLabs** %	**CHF Total**
AP0: Projektstart und -organisation	2.5	2.5	5	10	50	50	CHF 1'310
AP1: Systemvoraussetzungen	10	10	20	40	50	50	CHF 5'240
AP2: Systemspezifikationen	20	76	84	180	53	47	CHF 21'655
AP3: Datenaufbereitung	6	18	56	80	30	70	CHF 10'966
AP4: Prototypentwicklung	50	300	150	500	70	30	CHF 53'100
AP5: Prototypevaluation	6	18	56	80	30	70	CHF 10'966
AP6: Prototypfinalisierung	5	30	5	40	88	13	CHF 3'740
AP7: Dokumentation	10	50	40	100	60	40	CHF 11'428
Total	109.5	504.5	416	1030	54	46	CHF 118'406

Personalkosten (excl. MWSt.)	**CHF 118'406**
MWSt. 7.7 %	CHF 9'117
Personalkosten (inkl. MWSt.)	**CHF 127'523**

Kostenplan zur Offerte (optionaler Teil)

	Anzahl Std. pro Tarif (CHF/h)				Anteile		
Arbeitsschritt	**FHNW A** *CHF 133.80*	**FHNW B** *CHF 76.20*	**ExoLabs** *CHF 157.00*	**Std. Total**	**FHNW** %	**ExoLabs** %	**CHF Total**
Beratung (50 Manntage)	50	200	150	400	63	38	CHF 45'480
Total	50	200	150	400	63	38	CHF 45'480

Personalkosten (excl. MWSt.)	**CHF 45'480**
MWSt. 7.7 %	CHF 3'502
Personalkosten Option (inkl. MWSt.)	**CHF 48'982**

Abbildung 1.2. Kostenzusammenstellung für das Projekt „Prototyp AI“.

1.3 Voraussetzungen für den Prototypen AI

Für eine erfolgreiche Implementierung des Prototyps sind gewisse Voraussetzungen notwendig in Bezug auf Expertenwissen, Datengrundlage und der zur Verfügung stehenden Infrastruktur (Hardware/Software) (Tabelle 1.1). Diese werden nachfolgend im Detail beschrieben.

Tabelle 1.1: Anforderungen für eine erfolgreiche Implementierung des Prototyps AI.

Anforderung		
Referenz	Beschreibung	Gewicht*
1.3.1	Expertenwissen	M
1.3.2	Primäre Datenbasis	M
1.3.2	Sekundäre Datenbasis AREA 1-3 Sekundäre Datenbasis TLM3D	M K 3
1.3.2	Hilfsdatensätze	K 1-3
1.3.3	Hardware/Software	M

**Gewicht: M = Muss, K = Kann mit Priorisierung 1-3, wobei 1 = hoch*

1.3.1 Aktuelles Vorgehen Arealstatistik

Grundsätzlich wird die Arealstatistik auf der Grundlage von Stereoluftbildern erhoben. Basierend auf der visuellen Interpretation wird jedem Hektarpunkt jeweils eine Nutzungsform (LU) und ein Bedeckungstyp (LC) zugeteilt. Die Interpretation erfolgt hierbei an einer Arbeitsstation, die eine dreidimensionale Betrachtung ermöglicht. Wenn LULC nicht eindeutig über die Luftbildinterpretation ableitbar ist, werden zusätzliche Datensätze zur Unterstützung herbeigezogen, die über ein Kartenfenster oder über einen Datenbankbezug als Informationen zum Punkt zur Verfügung stehen. Diese Zusatzdaten ermöglichen zum Teil auch, dass eine visuelle Interpretation nicht mehr notwendig ist. Jede LULC-Eingabe wird durch das Erfassungsprogramm auf Plausibilität geprüft (vgl. Plausibilitäts-Konzept AREA4). So ist beispielsweise die Kombination von LU und LC plausibilisiert, womit nicht mögliche Kombinationen ausgeschlossen werden bzw. bei seltenen Kombinationen zu einer Bestätigungsaufforderung führt. Der bisherige Arbeitsablauf am BFS sieht wie folgt aus:

- In AREA4 werden bei der Vorinterpretation alle Seepunkte von Seen abzüglich eines Buffers von 100m ausgeschnitten und müssen nicht interpretiert werden.
- Bei der Interpretation hat das Luftbild Priorität, aber als weitere zeitsynchrone Hilfsdaten kommen folgende Informationen hinzu:
    - Amtliche Vermessung
    - GWR: Bauzonen, Gebäude, Industrie, Attribut Anzahl Wohnungen
    - Differenzkarte Landeskarte (sollte mit TLM ersetzt werden in Zukunft)
    - Google Maps und Street View (Jahrgang wird angezeigt)
    - Swissimage Falschfarben-Infrarotbild
    - Hilfsfenster mit etwa 20 verfügbaren, gesamtschweizerischen Datensätzen (z.B. diverse Bundesinventare wie Moorlandschaften, Flachmoore, Hochmoore, etc.)
    - Als untergeordnete Plausibilitätsprüfung werden zusätzlich zu den nicht zeitsynchronen Hilfsdaten Zusatzinformationen wie Vegetationshöhe, Steigung, Exposition, Höhe über Meer und Sturm-Waldschäden (z.B. Vivian, Lothar) berücksichtigt.
- Eine Person (= Interpret/-in) bearbeitet ein Kartenblatt in ca. 1 Monat (1000 Stichprobenpunkte /Tag).
- Danach überprüft eine 2. Person (Kontrolle) einen Grossteil der ersten Interpretationen und korrigiert bei Bedarf. Die Unterschiede zwischen den Interpretierenden sind im Kapitel 1.3.3 dokumentiert.

- Stichproben, die nicht klar zugeordnet werden konnten, werden durch eine gemeinsame Feldbegehung der Interpreten/-innen erfasst. Pro Kartenblatt tritt dieser Fall bei ca. 100 Stichproben auf.

Für bestimmte LULC-Klassen ist ein gesondertes Vorgehen bezüglich der Interpretation festgelegt. So werden zum Beispiel bei temporären Nutzungen (z.B. Zirkuszelt, Festivalgelände) die langfristige Nutzung interpretiert und Ackerflächen werden in jedem Stadium als Acker ausgewiesen, wobei LC unabhängig vom aktuellen Zustand der Bewirtschaftung immer mit “LC 21 Grasvegetation” ausgewiesen wird. In Bereichen wo LULC nicht eindeutig zu erkennen sind, wie zum Beispiel im Schatten, bei Schneebedeckung oder einer Abdeckung durch Wolken, und auch über die Hilfsdaten keine sichere Veränderung festzustellen ist, wird grundsätzlich keine Änderung der Klassifikation vorgenommen. Darüber hinaus werden im Rahmen der AREA4 alle Hektarpunkte im Bereich stehender Gewässer abzüglich eines Buffers von 100 m bereits bei der Vorinterpretation ausgeschlossen und müssen nicht interpretiert werden. Die Erhebung wird hierbei auf Basis der Kartenblätter durchgeführt. In Abbildung 1.3 ist dieses schrittweise Vorgehen gut erkennbar.

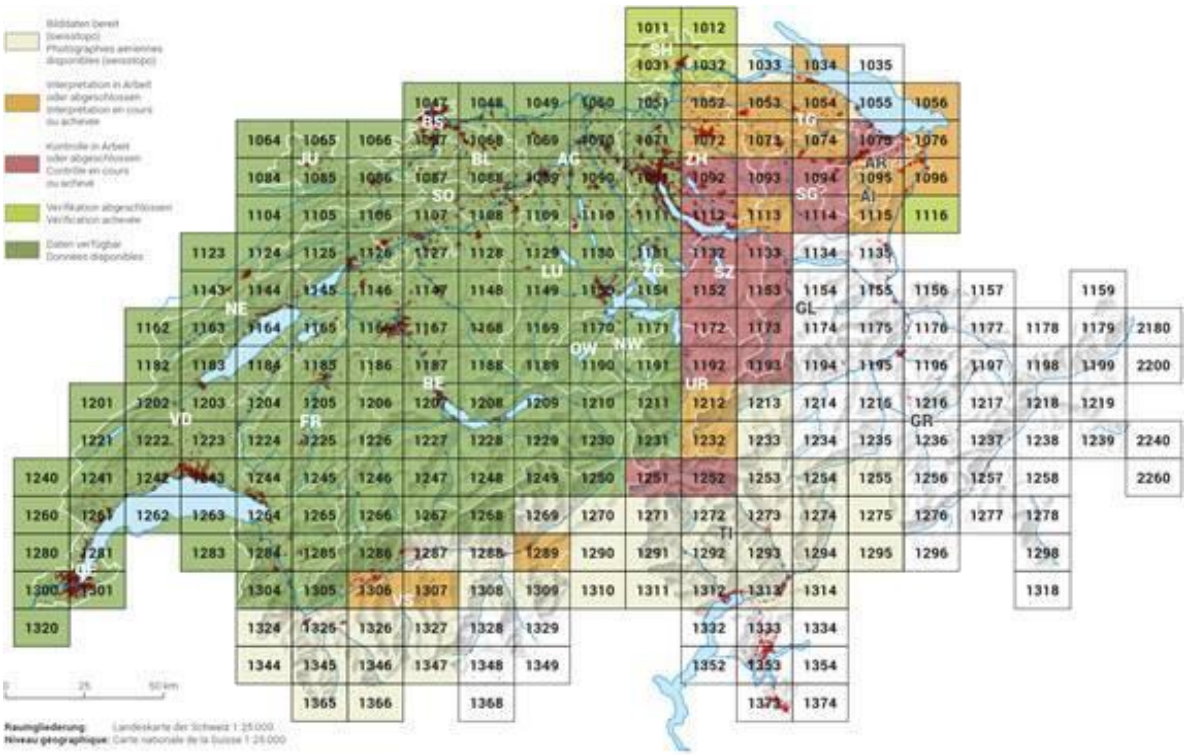


Abbildung 1.3. Erfassungsstand 01.06.2018 der Arealstatistik 2013/18 (Bundesamt für Statistik 2018a).

Eine Referenz bzw. ein „*Ground Truth*“ ist für den Prototypen AI von zentraler Bedeutung, da die Klassifikatoren auf der Basis dieser Referenz angelernt, trainiert und auch validiert werden. Im Prozess der Arealstatistik wird hierbei angenommen, dass jeweils die vorgängigen Erhebungen der Arealstatistik die Referenz bilden. Zur Überprüfung der Nachvollziehbarkeit wurden in AREA4 2’093’964 Stichprobenstandorte durch zwei unabhängige Interpreten/-innen klassiert, dokumentiert und die Ergebnisse verglichen (Abbildung 1.4). Die Ergebnisse zeigten, dass die Zuverlässigkeit der Klassifikation je nach Klasse variierte. Klassen, welche eine hohe thematische Ähnlichkeit aufwiesen, so genannte ‘ähnlichen’ Klassen, waren in ihrer Unterscheidung schwieriger und führten dementsprechend zu grösseren Unterschieden in den Interpretationen. Zudem konnten gemäss den Auswertungen des BFS systematische Abweichungen bzw. personenspezifische Wahrnehmungen zwischen den Interpreten/-innen festgestellt werden.

Bei der Interpretation dieser Ergebnisse muss berücksichtigt werden, dass die Stichprobenanzahl je nach Klasse stark variiert (vgl. Kapitel 1.3.3) und bei gewissen Klassen sehr gering ist. Diese Klassen weisen besonders hohe Streuungen auf. Die klassenspezifische Übereinstimmung der unterschiedlichen Interpretierenden ist in Abbildung 1.5 visualisiert und beträgt mehrheitlich 90-100 %.

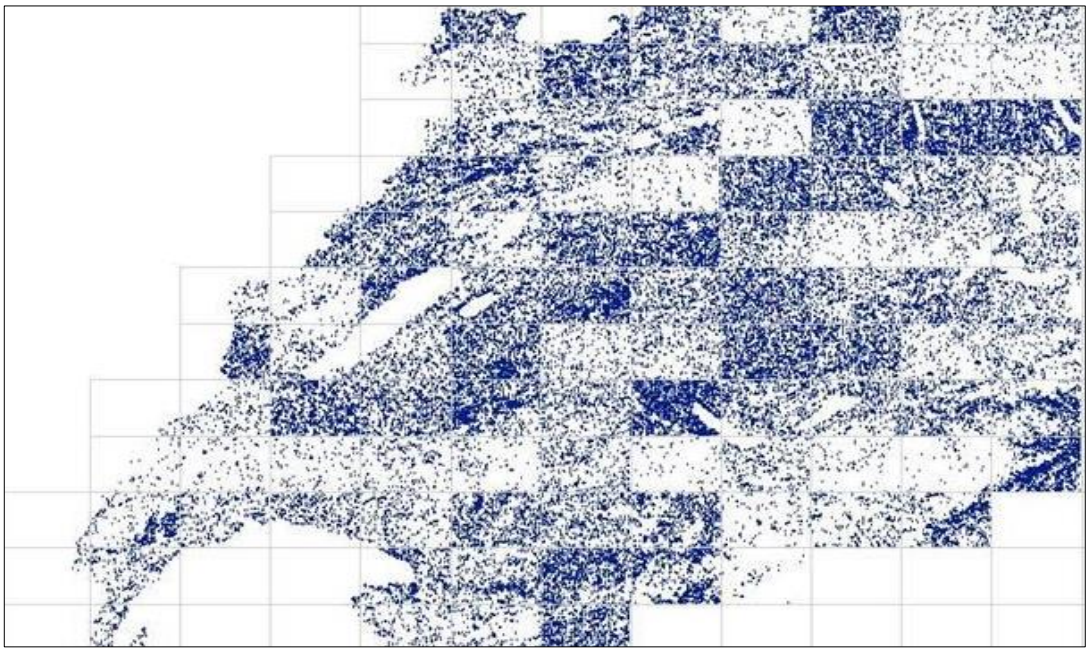
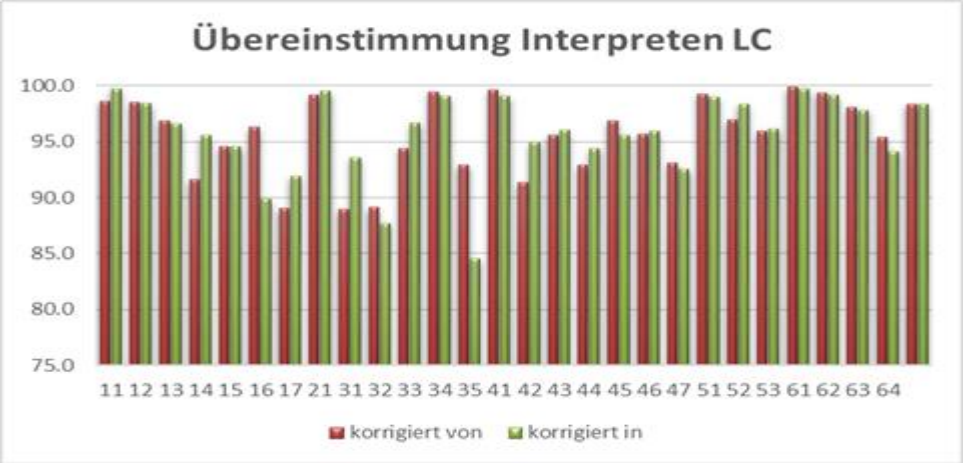


Abbildung 1.4. Stichprobenpunkte mit Abweichungen (Stand Ende 2017). Je nach Kombination der beiden Interpreten (pro Kartenblatt) bestehen Unterschiede in der Zahl der unterschiedlich beurteilten Stichproben.



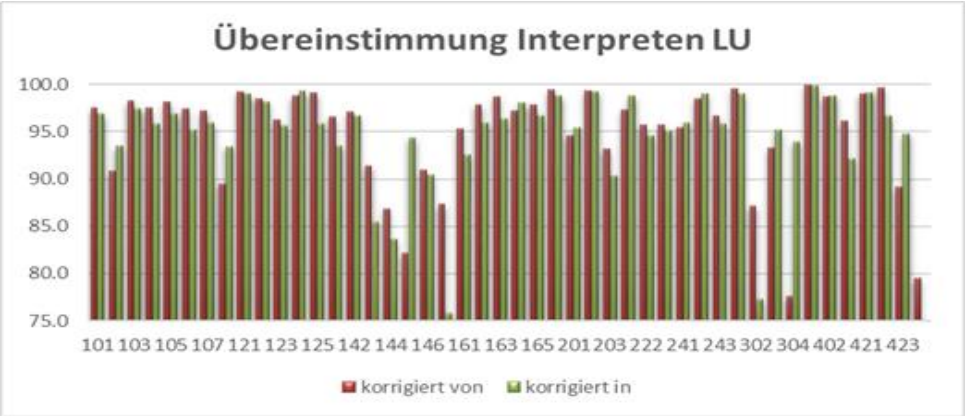


Abbildung 1.5. Übereinstimmung der Interpreten bei der Zuweisung der LULC-Kategorien.

Bei den Klassenbeschreibungen ist zu berücksichtigen, dass einige der Definition zu einem gewissen Grad offen gehalten sind, was einen über alle Klassen anwendbaren, generellen Merkmalskatalog basierend auf den Geodaten erschwert. Ein Beispiel wäre „LC 14 – Beetstrukturen“, wo der Flächenbezug bzw. die Anzahl/ die Grösse der Beete nicht festgelegt ist („Flächen mit klein strukturierter Beeteinteilung, heterogener Bepflanzung und vielen Gehwegen.“). Vergleichbares gilt für die Klasse „LC 15 – Rasen“, mit den Bedingungen, dass der Stichprobenpunkt auf der homogenen Rasenfläche liegen muss und es keine Mindestfläche gibt. Trotzdem werden über den Zugehörigkeitskatalog bzw. die Ausschlüsse Kontextinformationen hinzugezogen, die aber in ihrer räumlichen Dimension nicht fixiert sind. So zählen begrünte Flachdächer nicht zu dieser Kategorie, obwohl grundsätzlich der Stichprobenpunkt auf der entsprechenden Rasenfläche auf dem Dach den Bedingungen entsprechen würde. Dementsprechend können solche Sonderfälle natürlich die Klassifikation bzw. die Klassenzuweisungen deutlich erschweren.

1.3.2 Grundlagedaten Arealstatistik

Für die Arealstatistik stehen eine Vielzahl verschiedener Daten zur Verfügung. Hierbei wird unterschieden zwischen der primären Datenbasis, welche als Interpretationsgrundlage dient, und der sekundären Datenbasis, welche die Interpretation oder Kontrolle ersetzt bzw. zum Vorentscheid dient (vgl. Dokument “Grundlagedaten AREA5 - Stand 13.12.2018). Darüber hinaus finden Hilfsdatensätze Verwendung, die unterstützend im Rahmen des Interpretationsprozesses herangezogen werden können (vgl. Kapitel 1.3.1). Tabelle 1.2 zeigt eine Übersicht der für die Arealstatistik relevanten aktuellen und auch potenziellen Grundlagedaten (mit Hinblick auf AREA5).

Tabelle 1.2. Aktuell für AREA4 zur Verfügung stehende und mögliche zukünftige Grundlagedaten (für AREA 5 - in Kursiv).

Primäre Datenbasis	Sekundäre Datenbasis	Hilfsdatensätze
SWISSIMAGE (stereo) SWISSIMAGE FCIR (stereo) SWISSIMAGE RS (stereo)	AREA 1-3 TLM3D *AREA 4* *ADELE*	Amtliche Vermessung Parzellen und Gemeindegrenzen Differenzkarte Bauzonen Bundesinventare nationaler Schutzgebiete ÖREB-Kataster Swissimage (Farb und FCIRorthofotomosaik) Swissimage FCIR (FCIR-Orthofotomosaik) Pixelkarte LK25 Eidgenössisches Gebäude- und Wohnregister Betriebs- und Unternehmensregister swissALTI3D Landwirtschaftliche Zonengrenzen Vegetationshöhe TLM Bodenbedeckung Bemerkungen aus AREA 1-3 Digitale Fotothek Google Maps/Streetview (Webviewer) Open Street Map (Webviewer) Geoportal Bund & Kantonale Geoportale Internet *Satellitendaten (Sentinel-Missionen)*

Die Verwendung dieser Datengrundlage in einem automatisierten und operationellen Verfahren ist insofern eine Herausforderung, da eine hohe Heterogenität bezüglich Datenstruktur, Datenkonsistenz und Informationsgehalt vorhanden ist. So ist beispielsweise für eine Integration der Hilfsdatensätze in den Klassifikationsprozess ein zeitlicher Erhebungszeitpunkt der Informationen notwendig, der möglichst identisch oder mindestens zeitnah zu den Aufnahmen der primären Datenbasis liegt. Zudem ist im Rahmen von möglichen Datenverschneidungen und Plausibilisierungen eine strukturierte Harmonisierung der thematischen Kategorien unter Berücksichtigung der Qualität der Informationen unabdingbar.

1.3.3 Infrastruktur Hardware/Software

Der Prototyp AI wurde unter Windows 10 entwickelt. Die verwendete Programmiersprache ist Python in der Version 3.6. Mögliche externe Abhängigkeiten (z.B. spezifische Py-Module und Bibliotheken) wurden gekapselt, womit der Prototyp AI als Stand-Alone Software funktioniert. Für das Deep Learning-Modul "*Convolutional Neural Network*" (CNN) kam die Tensorflow-Umgebung in der Version 1.10 zum Einsatz. Die Ansteuerung erfolgte mittels der Highlevel-API *Keras* (Version 2.2.4). Für die Berechnung mittels Nvidia GPUs kam das CUDA-Framework (CUDA Toolkit 9.0), sowie das cuDNN-Framework (cuDNN v7.0) zur Unterstützung parallelisierbarer Machine-Learning Operationen auf CUDA-Cores zum Einsatz. Für das "*Random Forest*" (RF) - Modul wurde der parallelisierbare und hocheffiziente Algorithmus der Bibliothek *ranger* eingesetzt.
Zur Entwicklung des Prototypen AI wurden Consumer-Grade Einzelplatz-Rechner des Instituts Geomatik eingesetzt. Die Weiterentwicklung bestehender Netzwerke geschah auf einem Ubuntu/Fedora-Linux Dell Precision Tower 7910, CPU Intel Xeon E5-2630 2.2GHz 10-Core, 32GB RAM, GPU Nvidia GTX 1080Ti 11GB DDR5 mit 3584 CUDA Cores. Das für den Prototypen AI zu verwendende Netzwerk wurde unter Windows 10 mit einem Consumer-Grade Dell Precision Tower 5810, CPU Intel Xeon E5-1660 3.0GHz Octacore, 32GB RAM, GPU Nvidia Quadro K5200 8GB DDR5 mit 2304 CUDA Cores aufgebaut. Auf diesem Rechner erfolgte auch die Entwicklung der Klassifikatoren. Als minimale Grafiklösung gilt eine Nvidia Quadro K4000 3GB DDR5 mit 768 CUDA Cores; es sind gegenüber der K5200 jedoch zwei- bis dreimal längere Trainingszeiten für das DL CNN zu erwarten. RF profitiert hingegen in erster Linie von einer schnellen CPU. Da die ranger-Bibliothek mit einem hocheffizienten Multithreading-Algorithmus arbeitet, begünstigt vor allem eine grössere Anzahl Hauptprozessoren (z.B. 16-Core statt Octacore) die Berechnungszeit.

Das Training der Klassifikatoren kann semi-automatisch in für notwendig erachteten Zeitabständen neu ausgeführt werden. Die Datengrundlage liegt lokal vor bzw. kann bei Bedarf für zukünftige Erhebungen, durch die im Rahmen des Prototyps entwickelten Komponenten, automatisiert berechnet und heruntergeladen werden (z.B. in Hinblick auf die Metriken basierend auf den Satellitendaten oder die zukünftigen Swissimage-Bilddaten).

2 Prototyp AI

2.1 Ziele

Das Ziel des Prototyps AI ist es, über transparente und reproduzierbare Methoden des maschinellen Lernens und automatisierten Plausibilisierungsregeln eine Effizienz- und Qualitätssteigerung der Arealstatistik zu erreichen. Eine Steigerung der Effizienz kann hierbei insbesondere dadurch erreicht werden, dass die Anzahl derjenigen Hektarpunkte, die interpretiert (Vorinterpretation) oder kontrolliert (Vorkontrolle) werden müssen, reduziert wird. Dies würde dementsprechend zu einer Beschleunigung (d.h. Produkterstellung alle 6 oder 9 Jahre) und/oder Vereinfachung des Interpretationsprozesses führen. Eine Qualitätssteigerung kann vor allem durch zusätzliche bzw. optimierte Plausibilisierungen und erweiterter Interpretationshilfen umgesetzt werden. Der Fokus liegt hierbei vor allem auf den LULC-Kategorien, die einen sehr hohen Anteil an den Hektarpunkten haben (z.B. Kategorien mit Bezug zu Wald, dem unproduktiven Gebiet und der Landwirtschaft).

Als Ausgabe des Prototypen AI ist eine LULC-Klassifikation für alle möglichen Kombinationen aus LU und LC angestrebt. Als Teil dieses Klassifikationsergebnisses wird für jeden Hektarpunkt die Wahrscheinlichkeit der LULC-Zuweisung quantifiziert. In diesem Zusammenhang wird auch eine Veränderungsanalyse durchgeführt, die für jeden Hektarpunkt die Wahrscheinlichkeit für eine Klassenverschiebung für LU und/oder LC bestimmt. Darüber hinaus werden Leitfäden für die Anwendung und strukturierte Optimierung der jeweiligen Module erarbeitet.

2.2 Konzept

Für die Umsetzung der Prototypenentwicklung wurde eine iterative Vorgehensweise gewählt. Im Hinblick auf die AI-Komponenten (CNN/RF) bedeutet dies, dass die entsprechenden Module einzeln entwickelt wurden und jedes für sich hohe Zuverlässigkeiten und eine gute Performance aufweisen musste, bevor sie in die Komplettlösung "Prototyp AI" integriert wurden. In Bezug auf die Datengrundlage lag die höchste Priorität auf den Luftbilddaten, da diese auch gleichzeitig die primäre Datenbasis für den Entscheidungsprozess im Rahmen der Interpretation am BFS darstellen. Zusätzlichen Daten wurden in Abstimmung mit den Projektpartnern anschliessend schrittweise eingebunden. Abbildung 2.1 zeigt die Struktur des Prototyps AI inklusive Schnittstellen, Vorverarbeitungskomponenten und Auswertemodule.

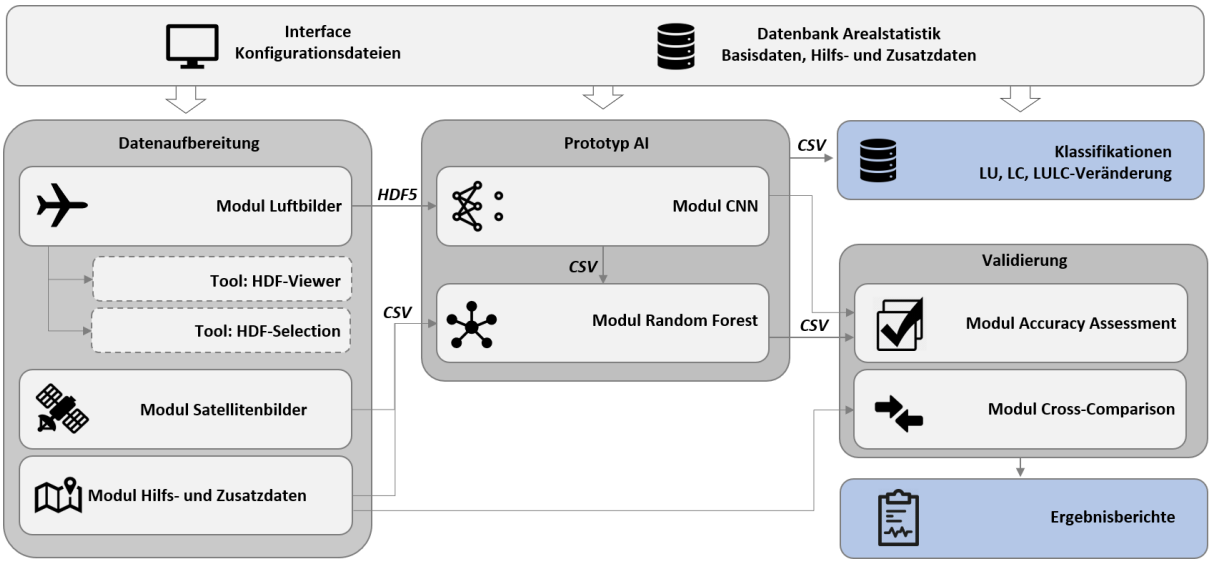


Abbildung 2.1. Übersicht über die Struktur des Prototyp AI sowie der zugehörigen Schnittstellen und Module für die Datenaufbereitung und Auswertung.

Der Prototyp AI setzt sich aus den Modulen Datenaufbereitung, den CNN- und RF-Modulen, sowie einem Validierungsmodul zusammen. Das entwickelte und optimierte CNN wurde auf der Grundlage von Orthofotos (SWISSIMAGE) für den Erhebungszeitraum AREA3 und AREA4 umgesetzt und gibt, neben diversen Qualitätsangaben und statistischen Kennwerten, die Wahrscheinlichkeiten für die Klassenzuordnung separat für LU und LC aus. Diese werden anschliessend im RF-Modul mit ausgewählten Hilfsdatensätzen (z.B. AV, swissALTI3D oder Satellitendaten) zusammengeführt, woraus eine optimierte LULC-Klassifikation resultiert. Die Ergebnisse aus CNN und RF wurden separat bezüglich der thematischen Zuverlässigkeiten und eventueller räumlicher Korrelationen und Muster ausgewertet (Kapitel 4). Diese Auswertung bildet die Basis, um über strukturierte Sensitivitätsanalysen (Kapitel 3) eine optimale Parametrisierung sowie die am besten geeignete Kombination an Eingangsdaten zu ermitteln.

2.3 Datenaufbereitung

In diesem Kapitel werden die Module zur Aufbereitung der primären, sekundären und Hilfsdaten als Eingangsgrössen für den Prototypen AI in ihrer Funktion und Implementierung beschrieben. Die Ausgangssituation ist hierbei in Abbildung 2.2 visualisiert.

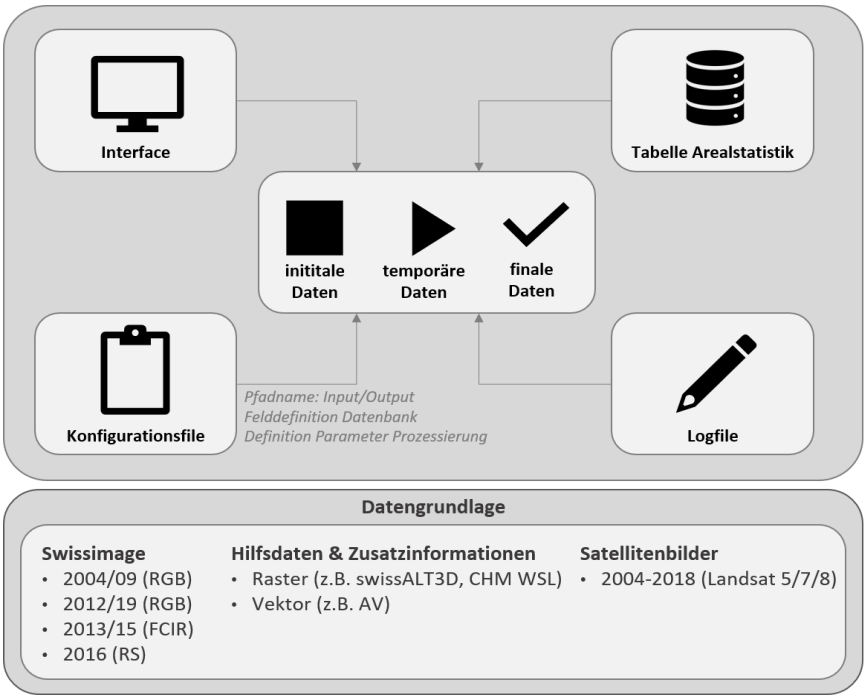


Abbildung 2.2. Konzeptionelle Architektur der Datenvorverarbeitung.

Basierend auf einem Interface kann über eine Konfigurationsdatei auf die einzelnen Module zugegriffen und diese parametrisiert werden. Die Module verknüpfen dann die entsprechenden initialen Basis-, Hilfs- und Zusatzdaten mit der Tabelle/Datenbank der Arealstatistik, führen spezifische Vorverarbeitungsschritte durch und generieren die finalen Daten für die nachfolgenden Klassifikationsschritte. Jeder Ablauf und die jeweilige Parametrisierung werden hierbei über ein Logfile dokumentiert.

Alle Module wurden in der Programmiersprache Python umgesetzt, wobei einige Prozessierungsschritte betreffend der Satellitendaten in der *Google Earth Engine* durchgeführt wurden (dies erfolgte aber ebenfalls über eine Python-Schnittstelle). Die Stichprobendaten von AREA3/4 wurden in eine Datenbank basierend auf *PostgreSQL/PostGIS* überführt, die speziell für räumliche Abfragen ausgelegt ist. Die jeweiligen Module generieren teilweise temporäre Dateien, die mehrmals für bestimmte Prozessierungsschritte notwendig sind – der Fokus liegt aber auf der Generierung gebrauchsfertiger Daten und Datenstrukturen für die CNN/RF-Module als auch für die davon unabhängigen (post-) Klassifikationsansätze (Abbildung 2.3). Die detaillierte Beschreibung der einzelnen Module erfolgt in den nachfolgenden Kapiteln.

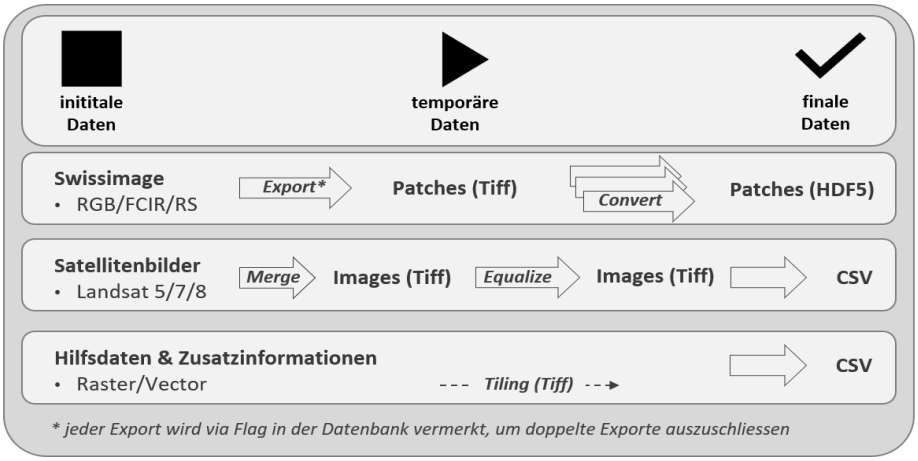


Abbildung 2.3. Datenarchitektur für die jeweiligen Module.

2.3.1 Datenaufbereitung Luftbilder

Die Luftbild-Grundlage basiert auf den Orthofotos von SWISSIMAGE von 2004-2018, der in regelmässigen Abständen durchgeführten landesweiten Befliegung vom Bundesamt für Landestopografie (swisstopo - https://shop.swisstopo.admin.ch/en/products/images/ortho_images). Diese Daten variieren in der räumlichen (50 cm vs. 25 cm), spektralen (RGB vs. FCIR) und radiometrischen (8-bit vs. 16-bit) Auflösung sowie im zugrunde liegenden Koordinatensystem und im Aufnahmezeitpunkt (vgl. Tabelle 2.1).

Tabelle 2.1. Spezifikationen der für die Arealstatistik zur Verfügung stehenden Luftbilder.

Produkt	Räumliche Auflösung	Spektrale Auflösung	Radiometrische Auflösung	Koordinatensystem	Aufnahmezeitpunkt
SWISSIMAGE 50 cm	50 x 50 cm	3-Kanal (R-G-B)	3 x 8-bit	LV03	bis 2005
SWISSIMAGE 25 cm	25 x 25 cm	3-Kanal (R-G-B)	3 x 8-bit	LV03 LV95	2006-2007 ab 2007
SWISSIMAGE FCIR	25 x 25 cm	3-Kanal (NIR-R-G)	3 x 16-bit	LV95	2010-2015
SWISSIMAGE RS	25 x 25 cm	4-Kanal (NIR-R-G-B)	3 x 16-bit	LV95	ab 2016

In der Vorprozessierung wurden dementsprechend verschiedene Schritte der Harmonisierung durchgeführt: Für die Anpassung der räumlichen Auflösung von *SWISSIMAGE 50 cm* wurde ein Resampling auf 25 x 25 cm basierend auf dem Lanczos-Filter durchgeführt. Bei einigen der Luftbilder war das Bezugssystem nicht definiert und/oder wurde nicht korrekt erkannt, weshalb eine Qualitätskontrolle der Metadaten erfolgte und fehlende Informationen ergänzt wurden.

Im nächsten Schritt erfolgte die Extraktion der 50 x 50 m Referenzflächen (= 200 x 200 Pixel) für alle Stichprobenpunkte von AREA3 und AREA4. Hierbei musste ein Vorgehen gewählt werden, welches die Stichprobenpunkte der AREA4 berücksichtigt, welche noch nicht durch Luftbilddaten abgedeckt werden. Im Rahmen der Extraktion der Referenzflächen wird auch die Zuweisung zum zugehörigen Stichprobenpunkt in Form eines eineindeutigen Dateinamens (<ID>_raw.tiff) umgesetzt. Anschliessend erfolgt die Konvertierung in das HDF5-Datenformat, welches das effiziente Speichern und Auslesen von mehrdimensionalen Daten ermöglicht und gleichzeitig auch die Speicherung von Metadaten unterstützt. Die Referenzflächen werden hierbei getrennt für AREA3 und AREA4 sowie für LU, LC und LULC-Veränderung in Unterordnern abgelegt. In Abhängigkeit

von der Zugehörigkeit zu bestimmten Klassen (Klassen mit < 100k Stichprobenpunkten), wurden diejenigen Referenzflächen entfernt, die mittig auf der Grenze zwischen zwei Luftbildaufnahmen lagen (ca. 3 % der Stichprobenpunkte). Ebenfalls ausgeschlossen wurden Testflächen, die nur zum Teil von Luftbildern abgedeckt waren. Dies ist ein notwendiger Schritt zur Qualitätssicherung, da zum Teil bedeutende Unterschiede in den spektralen Signaturen auftraten (Abbildung 2.4) und diese als ungeeignet für das Maschinelle Lernen angesehen wurden.

Abbildung 2.4. Referenzfläche AREA 4 mit nur teilweiser Luftbildabdeckung (links) sowie Referenzflächen im Grenzbereich unterschiedlicher Befliegungszeitpunkte.

Die Prozessierung für SWISSIMAGE FCIR und SWISSIMAGE RS ist leicht abweichend von der oben beschriebenen Vorgehensweise. Da im Gegensatz zum SWISSIMAGE (RGB) aus den Einzelfotos kein Mosaik erstellt wird, kann es bei der Erstellung der Kacheln zu starken Inhomogenitäten kommen. Die damit verbundene Unsicherheit bezüglich der Qualität der Daten wird zusätzlich noch dadurch verstärkt, dass laut swisstopo weder eine radiometrische noch eine geometrische Anpassung der Luftbildstreifen vorgenommen wurde (z.B. Zunahme der planimetrischen Ungenauigkeit im Bereich der Streifenränder). Letztendlich kommt es durch die verschiedenen Bildstreifen zu Bildüberlappungen, weshalb für eine Auswahl an Stichprobenpunkten zwei Aufnahmezeitpunkte zur Verfügung stehen und dementsprechend eine Entscheidung bezüglich der Referenzaufnahme getroffen werden muss.

Im Rahmen der Luftbild-Prozessierung wurde auch ein Algorithmus implementiert, welcher basierend auf den codierten Reflektanzwerten (d.h. die Digital Numbers im 8bit bzw. 16bit Farbraum) und einen anpassbarem Schwellwert die Pixel pro Referenzfläche zählt, die sehr geringe Werte in allen spektralen Kanälen aufweisen. Ist der Anteil dieser Pixel grösser 50% der Referenzfläche, wird für diese Referenzfläche/ für diesen Hektarpunkt ein Flag über eine zusätzliche Spalte in der Datenbank eingetragen. Diese Information kann verwendet werden, um Referenzflächen mit einem grossen Anteil von Schatteneffekten im Rahmen der Auswertung gesondert zu handhaben. Mit Abschluss der Luftbild-Prozessierung liegen für alle Stichprobenpunkte die jeweiligen Referenzflächen (RGB/FCIR/RS) im HDF5-Dateiformat vor. Da die Anzahl an Dateien in einem Ordner limitiert ist, erfolgt die Ablage der Dateien in eine hierarchische Ordnerstruktur (Abbildung 2.5).

Abbildung 2.5. Ordner- und Datenstruktur nach Abschluss der Prozessierung der Luftbilddaten.

Die Vorprozessierung und Prozessierung der Luftbilddaten ist hierbei vollständig in einem Stand-Alone Python-Modul integriert, was sowohl eine eigenständige Funktionsweise als auch eine einfache Schnittstelle zum AI-Prototypen gewährleistet. Die verwendeten externen Bibliotheken sind hierbei gekapselt in das Modul eingebunden (eine Übersicht der Bibliotheken befindet sich in der Code-Dokumentation). Der konzeptionelle Aufbau des Python-Modules ist in Abbildung 2.6 visualisiert.

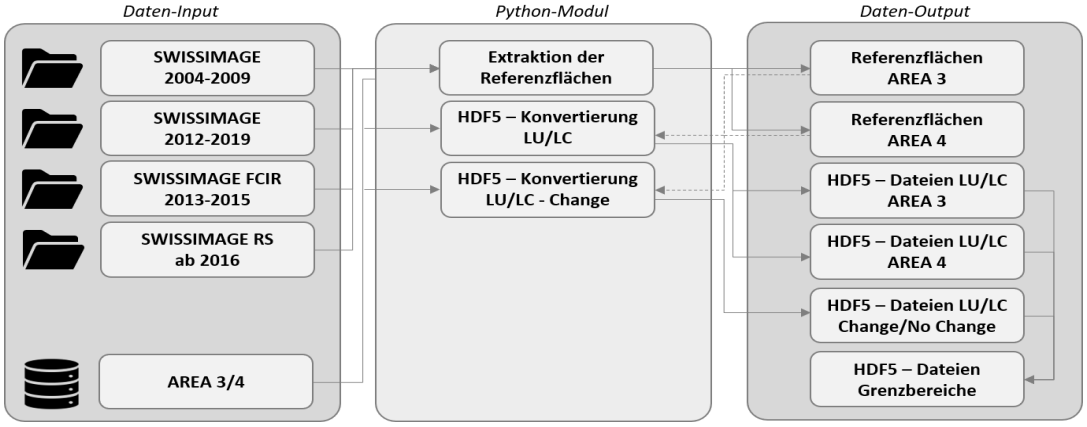


Abbildung 2.6. Funktionsweise des Python-Modules für die Prozessierung der Luftbilddaten.

Zusätzlich wurde noch ein weiteres Skript entwickelt, welches die Auswahl von Referenzflächen basierend auf einer CSV-Tabelle ermöglicht. Dies erleichtert die Auswahl der Referenzflächen im Rahmen eines Sampling-Verfahrens oder die gezielte Kontrolle der Daten für ausgewählte Stichprobenpunkte. Im Rahmen des Samplings für das CNN können die Stichproben (via ID oder Random) aus dem HDF5-Format im PNG-Format ausgegeben werden, wobei automatisch eine fürs CNN geeignete Ordnerstruktur (d.h. Unterscheidung in Training-/Validierungsdaten) vorgenommen wird. Für die CNN-Trainings sind Bildausschnitte mit drei Kanälen (RGB/FCIR) und einer Fläche von 50 x 50 m (200x200 Pixel) vorgesehen. Die Bilddimension ist hierbei frei einstellbar und kann für weitere Tests variiert werden (vgl. Kapitel 3).

2.3.2 Datenaufbereitung Satellitenbilder

Die Satellitendaten wurden primär dafür herangezogen, um den Mehrwert temporaler Eigenschaften, wie zum Beispiel der saisonalen Vegetationsdynamik, für den Klassifikationsprozess zu evaluieren. Für dieses Vorhaben sind nur zeitlich und räumlich hoch aufgelöste Satellitenmissionen geeignet. Weitere Einschränkungen sind durch die Notwendigkeit der Vergleichbarkeit der Dateninhalte, der freien Datenverfügbarkeit und der Kontinuität der Datenaufnahme gegeben. Dementsprechend beschränkte sich die Verwendung der Satellitendaten auf die in Tabelle 2.2 gelisteten Missionen.

Tabelle 2.2. Spezifikationen der verwendeten Satellitendaten.

Mission	Dateninhalt	Zeitliche Auflösung, Räumliche Auflösung	Zeitraum
Landsat 5	R-G-B-NIR-SWIR1/2-TIR	16 Tage, 30x30 m bzw 120x120 m	1984 - 2013
Landsat 7	R-G-B-NIR-SWIR1/2-TIR	16 Tage, 30x30 m bzw 60x60 m	seit 1999
Landsat 8	R-G-B-NIR-SWIR1/2-TIR	16 Tage, 30x30 m bzw 60x60 m	seit 2013

Da eine Analyse und Nutzung der Satellitendaten aufgrund der benötigten Datenmengen und der Anforderungen an die Rechenleistung lokal nur für kleinere Testgebiete möglich wäre, wurde die (Vor-)Prozessierung der Daten mittels der *Google Earth Engine*, einer Cloud-Computing-Plattform, umgesetzt (https://earthengine.google.com/). Der Zugriff auf das entsprechende Skript ist hierbei über eine Python-Schnittstelle gewährleistet.

In einem ersten Schritt wurde sich ausschliesslich auf die *Surface Reflectance* - Daten der Landsat-Missionen (https://landsat.usgs.gov/landsat-missions-timeline) konzentriert, da nur diese eine zeitliche Abdeckung gewährleisteten, die für eine LULC-Veränderungsanalyse notwendig ist. Um die Auswertung möglichst unvoreingenommen durchzuführen, erfolgte die Auswahl der zu untersuchenden Informationsebenen vor allem unter Berücksichtigung der Vergleichbarkeit zwischen den einzelnen Missionen. Die Auswahl der Indizes/Ratios erfolgte basierend auf einem intensiven Literaturstudium, wobei der Fokus auf einer guten Identifikation verschiedener LC-Klassen lag, insbesondere auf der Unterscheidung unterschiedlicher Vegetationstypen (Tabelle 2.3).

Tabelle 2.3. Spezifikationen der extrahierten Variablen aus den Landsat-Datensätzen.

Spektrale Bänder					
• R	• G	• B	• NIR	• SWIR1/2	• TIR
Indizes/Ratios					
NDVI	*(NIR-R)/(NIR+R)*		NDWI	*(G-NIR)/(G+NIR)*	
MNDWI	*(G-SWIR)/(G+SWIR)*		EVI2	*2.5*(NIR-R)/(R+2.4*R+1)*	

Für jede dieser Informationsebenen erfolgte dann für jede Landsat-Mission die Datenanalyse für den Zeitraum 2000 – 2017. Ziel der Analyse war es, höherdimensionale beschreibende Variablen der temporalen Charakteristiken zu erhalten. Zu diesem Zweck wurde eine Ausgleichsrechnung (*Curve Fitting*) über alle Aufnahmezeitpunkte durchgeführt und für jedes Jahr und in regelmässigen Zeitabständen (10 Zeitpunkte) der entsprechende Wert berechnet (vgl. Abbildung 2.7). Die so in der *Google Earth Engine* berechneten Variablen wurden dann als Kacheln heruntergeladen und bilden die Grundlage für die weiteren Prozessierungs- und Klassifikationsschritte.

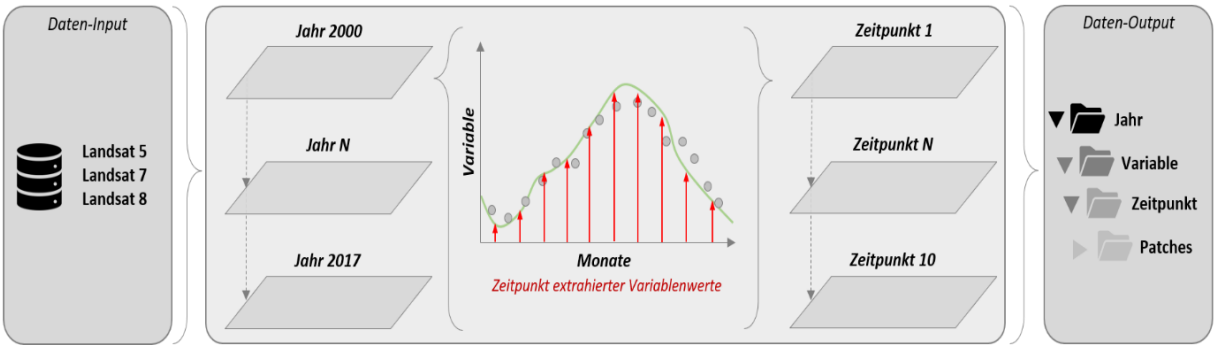


Abbildung 2.7. Berechnung der Variablen in der Google Earth Engine.

Die Verwendung der so generierten Daten ist primär als Eingangsgrösse für den RF – Klassifikator gedacht und damit Teil der *Late Fusion* im Prototypen AI. Damit die Daten optimal in das RF – Verfahren eingebunden werden können, erfolgte die Zuweisung der jeweiligen Pixelwerte (für jede berechnete Variable) zu jedem Stichprobenpunkt der Arealstatistik in Form einer CSV-Tabelle. Darüber hinaus wurden auch CSV-Tabellen erstellt, die die Pixel in der Nachbarschaft des zentralen Pixels berücksichtigen (= Umgebungspixel). Somit können bei Bedarf auch Strukturinformationen bzw. Kontextinformationen eines grösseren räumlichen Ausschnittes in der Klassifikation berücksichtigt werden.

Die Vorprozessierung und Prozessierung der Satellitendaten ist vergleichbar zu den Luftbildern (Kapitel 2.3.1) vollständig in einem Stand-Alone Python-Modul integriert. Über eine einfache UI kann der Nutzer eine Konfigurationsdatei lesen und schreiben, die eine individuelle Anpassung der Parametrisierung an die jeweiligen Bedürfnisse ermöglicht. Der konzeptionelle Aufbau des Python-Modules ist in Abbildung 2.8 visualisiert.

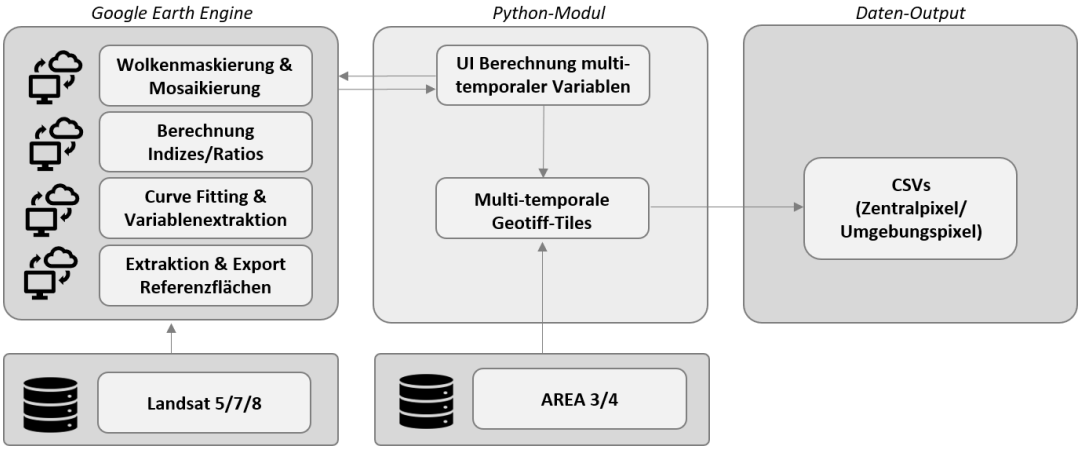


Abbildung 2.8. Funktionsweise des Python-Modules für die Prozessierung der Satellitendaten.

2.3.3 Datenaufbereitung Hilfs- und Zusatzdaten

Die Hilfs- und Zusatzdaten können sowohl im CNN- als auch im RF-Modul eingebunden werden, wobei im Rahmen der vorliegenden Studien nur eine Verwendung im Rahmen der RF-Klassifikation im Kontext der *Late Fusion* vorgenommen wurde. Grundsätzlich würden sich als CNN–Input aber vor allem rasterbasierte Informationen eignen, da Vektorinformationen von der Datencharakteristik nur bedingt den Anforderungen des CNN entsprechen. An Rasterdaten wurden Gelände- bzw. Terrainmodelle (DTM), Oberflächenmodelle (DOM/DSM) und Objekt- bzw. Vegetationshöhenmodelle (CHM) berücksichtigt. Die jeweiligen, für eine automatisierte Prozessierung notwendigen, Spezifikationen sind in Tabelle 2.4 gegeben.

Tabelle 2.4. Ausgewählte Spezifikationen der berücksichtigten rasterbasierten Daten.

Name	Räumliche Auflösung	Zeitmarke	Zeitraum
DOM swisstopo	2 x 2 m	„Flugjahr“ (Footprint Shapefile) „XXX_[Flugjahr].asc“ (ASC Filename)	2000 - 2007
DOM WSL	1 x 1 m	nicht vorhanden (aber: basierend auf SWISSIMAGE)	2007 - 2012 2011 - 2016
DTM AV	2 x 2 m	„Flugjahr“ (Footprint Shapefile) „XXX_[Flugjahr].asc“ (ASC Filename)	2000 - 2008
swissALTI3D	2 x 2 m	nicht vorhanden (aber: basierend auf DTM AV, DTM-TLM, kantonalen Daten, SWISSIMAGE)	seit 2010 (jährliche Nachführung)
CHM LFI	1 x 1 m	nicht vorhanden (aber: basierend auf SWISSIMAGE, DOM WSL, swissALTI3D)	2007 - 2012 2011 - 2016

Für eine zuverlässige Nutzung im Klassifikationsprozess, ist vor allem eine möglichst ähnliche bzw. zeitnahe Abdeckung von Luftbilddaten und den rasterbasierten Hilfs- und Zusatzdaten notwendig. Da dies für jeden Hektarpunkt einzeln bestimmt werden muss, ist die Verfügbarkeit einer auslesbaren Zeitmarke in den Metadaten/ im Dateinamen unabdingbar. Dies ist zum Teil kritisch, da Produkte wie zum Beispiel das CHM LFI eine Kombination von swissALTI3D, dem DOM WSL und dem SWISSIMAGE FCIR ist und dementsprechend unterschiedliche Aufnahmezeitpunkte kombiniert werden. Die Eignung und der Mehrwert solcher Produkte bedarf folglich einer Abklärung von Seiten des BFS, unter anderem auch in Bezug auf die Qualitätssicherung.

Basierend auf diesen Rasterdaten ist die Ableitung weiterer, die Oberfläche beschreibender Variablen möglich. Diese umfassen zum Beispiel die Hangneigung, die Geländeausrichtung oder Struktur/Textur-beschreibende Metriken (z.B. die Homogenität der Variablenwerte für einen definierten Ausschnitt). Die Berechnung dieser Variablen und deren Verwendung ist hierbei abhängig von dem jeweiligen Datensatz und den vom BFS vorgegebenen Rahmenbedingungen.

An Vektordaten stehen eine grosse Auswahl an Produkten zur Verfügung und diese werden bereits umfangreich im Rahmen der Plausibilisierung und der Interpretation der Stichprobenpunkte am BFS verwendet (vgl. Kapitel 1.3.2). Bei den Vektordaten liegt die Herausforderung vor allem in der thematischen Struktur (d.h. Klassentiefen und -definitionen) und der Rückverfolgung der Zeitmarken bei nachgeführten Datensätzen (z.B. AV - vgl. Tabelle 2.5).

Tabelle 2.5. Ausgewählte Spezifikationen der berücksichtigten vektorbasierten Daten.

Name	Inhalte	Zeitmarke	Zeitraum
AV	LULC: 52 Klassen Einzelobjekte: 42 Klassen Eigentümerschaft: 27 Klassen (keine vollständige Abdeckung)	„state_of“ (nur teilweise vorhanden, keine vorhergehenden Versionen sichtbar)	digital seit 1993 (monatliche Nachführung)
swissTLM3D	LULC: 76 Klassen Einzelobjekte: 86 Klassen (keine vollständige Abdeckung, kontinuierliche Anpassung der Klassen)	„Datum_Erstellung“ „Datum_Aenderung“ „Erstellung_Jahr“ „Erstellung_Monat“ „Revision_Jahr“ „Revision_Monat“	seit 2008 (jährliche Nachführung)
swissBuildings3D 2.0	LULC: 30 Klassen (keine vollständige Abdeckung, kontinuierliche Anpassung der Klassen)	„Datum_Erstellung“ „Datum_Aenderung“ „Erstellung_Jahr“ „Erstellung_Monat“	unbestimmt (Nachführung im 6-Jahreszyklus)
STATENT	Einzelobjekte: 794 Klassen (vgl. NOGA 2008)	„STATENT_Year“	seit 2011 (jährliche Nachführung; rückberechnete Daten für die Erhebungsjahre 2005/2008)
GWR-STATPOP	Einzelobjekte: > 70 Klassen (vgl. EUROSTAT)	„GBAUJ“ „GBAUP“ „GABBJ“	seit 2002/2010 (jährliche Nachführung)

Dementsprechend wurde neben der Extrahierung der Informationen als Input für den RF – Klassifikator vor allem der Mehrwert von Schnittmengen zwischen den Produkten analysiert, was eine vorhergehende Harmonisierung der jeweiligen Inhalte und verfügbarer Zeitmarken notwendig machte. Die Ergebnisse der Schnittmengenanalyse zeigten jedoch, dass die thematischen Komponenten der Daten einen direkten Vergleich bzw. einen direkten Nutzen im Rahmen einer Plausibilisierung nur sehr eingeschränkt erlauben (vgl. Kapitel 4).

Um die Nutzung der Hilfs- und Zusatzdaten so offen und flexibel wie möglich zu gestalten, wurde ein Modul entwickelt, welches die die Hilfs- und Zusatzdaten entweder im Raster- oder im Vektorformat einliest, diese mit der Datenbank der Arealstatistik verschneidet und jedem Hektarpunkt den entsprechenden Wert des Hilfs- oder Zusatzdatensatzes an der Position des Stichprobenpunktes zuweist. Für die Daten im Rasterformat steht hierbei die Option offen, auch zusätzliche, flächenbezogene Variablen zu berechnen und zuzuweisen. Die resultierende Datei ist dementsprechend eine CSV-Tabelle, die direkt in den RF-Klassifikator eingelesen werden kann. Zusätzlich können für die Rasterdaten auch Patches, ähnlich wie bei der Luftbildprozessierung (vgl. Kapitel 2.3.1),

generiert werden, wenn deren Verwendung im CNN-Klassifikator von Interesse ist. Die Vorprozessierung und Prozessierung der Hilfs- und Zusatzdaten ist vergleichbar zu den bereits beschriebenen Modulen vollständig in einem Stand-Alone Python-Modul integriert. Über eine einfache UI kann der Nutzer eine Konfigurationsdatei lesen und schreiben, die eine individuelle Anpassung der Parametrisierung an die Bedürfnisse ermöglicht. Der konzeptionelle Aufbau des Python-Modules ist in Abbildung 2.9 visualisiert.

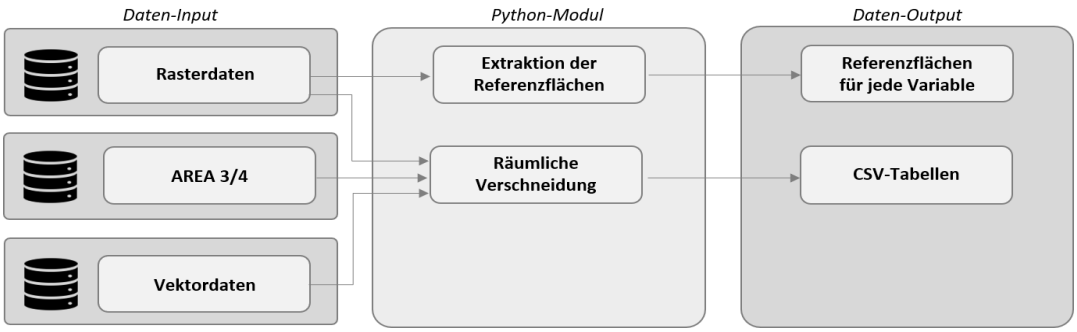


Abbildung 2.9. Funktionsweise des Python-Modules für die Prozessierung der Hilfs- und Zusatzdaten

2.3.4 Performance der Datenvorbereitungsmodule

Für eine erfolgreiche Einbindung der Datenvorbereitungsmodule in die operationellen Abläufe der Arealstatistik, wurden Abschätzungen für AREA 4 bezüglich Prozessierungszeiten und benötigtem Speicherplatz durchgeführt. In Tabelle 2.6 ist dies unterteilt für die jeweiligen Module der Luftbildprozessierung, Satellitenbildprozessierung und der Prozessierung von Hilfs- und Zusatzdaten aufgelistet.

Tabelle 2.6. Abschätzung von Prozessierungszeiten und benötigten Speicherplatz für die Datenaufbereitung.

Modul	Prozess	Prozessierungszeit	Speicherplatz
Modul Luftbilder	Extraktion Referenzflächen (für swissimage FCIR, 50x50m) HDF5-Konvertierung (für swissimage FCIR, 50x50m)	ca. 10 Tage ca. 4 Tage	ca. 5 TB ca. 2.5 TB
Modul Satellitenbilder	Google Earth Engine (für Landsat) Extraktion Referenzflächen (für Landsat)	ca. 4 Stunden ca. 10 Tage	- ca. 750 GB
Modul Hilfs- & Zusatzdaten	Extraktion Referenzflächen (für einen Raster-Datensatz 2 x 2 m) Räumliche Verschneidung (für einen Vektordatensatz)	ca. 4 Tage ca. 2 - 8 Stunden	ca. 500 GB ca. 5 GB

2.4 CNN

Dieses Kapitel beschreibt das CNN-Modul des Prototyps AI. welches in Python 3.6 mit den ML - Bibliotheken Tensorflow und Keras implementiert wurde. Die Quellcode-Dokumentation sowie die Installations- und Gebrauchsanleitungen befinden sich im Dokument “Prototype AI Documentation” (basierend auf der Sphinx-Toolbox).

2.4.1 Systemvoraussetzungen

CPU/GPU

Da ein CNN-Training ein rechenintensiver Prozess ist, erfordert es eine leistungsfähige Hardware. Moderne ML-Bibliotheken wie Tensorflow bieten die Möglichkeit, CNN-bezogene Berechnungen entweder auf CPU oder auf

GPU auszuführen. Die Laufzeit ist auf dem Grafikprozessor 10 bis 100 mal kürzer als auf der CPU, dabei hängt der Beschleunigungsfaktor von der jeweiligen Hardware ab. Auch wenn es möglich ist, dass entwickelte CNN-Modul auf der CPU zu betreiben, so wird doch eine Durchführung auf dem Grafikprozessor empfohlen, um das CNN-Training und die nachfolgende Klassifizierung zu beschleunigen. Es werden eine NVIDIA GPU-Karte mit CUDA Compute Capability 3.5 oder höher sowie die entsprechenden GPU-Treiber und CUDA-Bibliotheken benötigt. Die detaillierten Hard- und Softwareanforderungen sind im Bericht "Prototype AI Documentation" beschrieben.

Speicher

Um ein effizientes Training auf einer lokalen Maschine zu gewährleisten, müssen die Input-Daten lokal vorliegen, da ansonsten Netzwerkkapazitäten den limitierenden Faktor darstellen. Der letztendlich erforderliche Speicherplatz hängt von der Menge der Bildkacheln ab, für die das Netzwerk trainiert oder angewendet werden soll. Hierzu wurden Trainings-Testläufe mit verschieden umfangreichen Datensätzen durchgeführt, um einen möglichst optimalen Kompromiss aus Präzision und Berechnungsdauer zu finden. Mehr Informationen dazu befinden sich in Kapitel 2.4.3. Für das Inferencing - also die Anwendung des Netzwerkes - muss ein kompletter Datensatz an Kacheln im HDF5-Format bereitstehen (ca. 2.5TB bei einer Kachelgrösse von 200x200px, vgl. Tabelle 2.6).

2.4.2 Funktionsweise

Das Convolutional Neural Network ist das Kernstück des CNN-Moduls. Sein Hauptzweck ist die Durchführung der Klassifizierung von RGB- und/oder FCIR-Luftbildern, die vom Modul 'Preprocessing' erzeugt werden, in Bezug auf LULC. Die Ausgaben des CNN können direkt für eine LULC-Klassifikation verwendet oder weiterverarbeitet werden, z.B. als Eingabe für den RF- Klassifikator. Im Rahmen des Projektes wurde die *Xception* CNN-Architektur verwendet, die auf ImageNet vortrainiert wurde (Chollet 2017, Hochuli 2018). Ein vereinfachtes Netzwerkdiagramm ist in Abbildung 2.10 dargestellt.

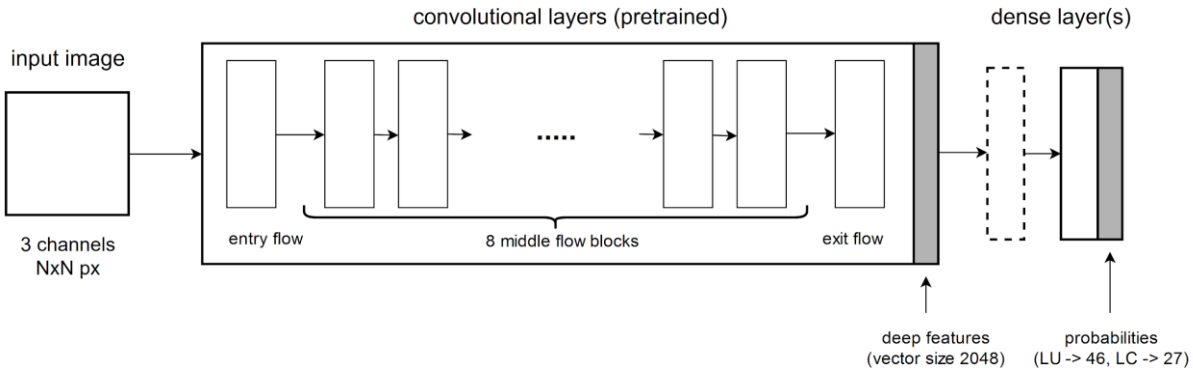


Abbildung 2.10. Vereinfachte Darstellung der Xception CNN - Architektur.

Das *Xception CNN* - Netzwerk besteht aus mehreren Faltungsschichten (Convolutions), die als Deep Feature Extraktoren fungieren. Es basiert auf der sogenannten 'depthwise separable convolution' (Deutsch: tiefenweise getrennten Faltung). Dabei wird je eine räumliche Convolution auf jeden Kanal der Eingangsdaten getrennt angewendet und über eine 1x1-Convolution wieder zusammengeführt. Das Netz 'Xception' basiert auf den getrennten Faltungsschichten sowie residualen Verbindungen (Chollet 2017). Auf die Faltungsschichten folgen dichte Schichten (Dense Layers), welche die Klassifizierung durchführen. Eine Ausgabemöglichkeit stellt eine Wahrscheinlichkeitsverteilung über vordefinierte Klassen dar. Die Klasse mit der höchsten Wahrscheinlichkeit wird einem Eingangsbild als vorhergesagtes Label zugeordnet. Das im CNN-Modul verwendete Netzwerk wurde auf dem ImageNet-Datensatz vortrainiert. Dabei handelt es sich um einen umfangreichen Datensatz von RGB-Alltagsbildern mit 1000 unterschiedlichen Kategorien. Die Nutzung eines vortrainierten CNN ist eine häufig verwendete ML-Technik, die als Transfer-Learning bezeichnet wird. Die Idee hinter dieser Methode ist, ein bestehendes CNN, welches auf einem grossen Datensatz trainiert wurde, für einen eigenen Datensatz zu optimieren

(fine-tune). Das vortrainierte Netzwerk ist bereits fähig, nützliche Bildmerkmale (z.B. Kanten, Ecken, bestimmte Muster usw.) zu extrahieren, und kann so im Vergleich mit einem zufällig initialisierten Netzwerk in kürzerer Zeit eine bessere Leistung erzielen.

Um das Transfer-Learning zu nutzen ist es notwendig, den letzten Dense Layer des ursprünglichen Xception-Netzwerks zu entfernen. Dieser Layer gibt einen Wahrscheinlichkeitsvektor der Grösse 1000 für ImageNet aus – und kann durch einen neuen, zufällig initialisierten Dense Layer z.B. der Grösse 46 für die LU- oder 27 für die LC-Klassifizierung ersetzt werden. Es ist aber auch möglich, einen zusätzlichen Dense Layer beliebiger Grösse zwischen dem Convolutions-Teil des Netzwerks und dem endgültigen Dense Layer hinzuzufügen (als gestricheltes Rechteck in Abbildung 2.10 dargestellt). Da die Dense Layer mit Zufallswerten initialisiert werden und der Rest des Netzwerks vortrainiert ist, wird das Training auf einem neuen Datensatz in zwei Phasen unterteilt: Vortraining der Dense Layer und Feineinstellung (Finetuning). Beim Vortraining der Dense Layers sind alle gefalteten Schichten des Netzwerkes eingefroren ('frozen') und nur die Dense Layer werden trainiert. Nach den Dense Layers wird ein Klassifikator auf Basis des generierten Deep Feature Vectors eingesetzt. Im Falle von Xception handelt es sich um eine spezialisierte Implementierung logistischer Regression. Im Rahmen der Feineinstellung werden der letzte gefaltete Block ('exit flow') und einige der mittleren gefalteten Blöcke ('middle flow blocks') nicht eingefroren ('unfrozen') sondern gemeinsam mit den 'Dense Layern' trainiert. Die Anzahl der an der Feineinstellung beteiligten, mittleren Blöcke ist hierbei ein einstellbarer Parameter. Das Training läuft in kleinen Schritten ab: eine geringe Auswahl von Bildern ('mini-batch') wird ins Netzwerk eingespeist und die Gewichte der Layer werden angepasst, so dass der Wert der Verlustfunktion ('loss function') sinkt. Dieser Prozess wird über mehrere sogenannten Epochen fortgesetzt. Dabei bedeutet eine Epoche, dass alle Bilder einmal ins Netzwerk eingespeist wurden. Am Ende jeder Epoche wird der Wert für die Verlustfunktion ('cross-categorical entropy') und die Gesamtgenauigkeit des Trainings und der Validierung berechnet. Das Training wird beendet, sobald die Gesamtgenauigkeit des Validierungsdatensatzes sich nicht mehr verbessert.

2.4.3 Datenvorbereitung und Sampling

Die Anzahl Stichproben pro Kategorie in der Arealstatistik variiert sehr stark - von einigen Kategorien mit weniger als 1000 Stichproben zu anderen mit über 1 Million Stichproben. Solche starken Unterschiede in der Stichprobenanzahl können negative Einflüsse auf den Trainingsvorgang haben und verlängern darüber hinaus die Trainingszeit. Um diese Effekte zu vermindern, wurden ausgeglichene Datensätze generiert. Hierzu wurden aus den zu den jeweiligen Klassen korrespondierenden HDF5-Files zufällig eine gleiche Anzahl Stichproben extrahiert. Waren in einer Klasse weniger Stichproben als in der benötigten Anzahl vorhanden, so wurden so viele Stichproben wie möglich verwendet. Für einen Basis-Datensatz wurden maximal 10'000 Stichproben pro Kategorie zufällig ausgewählt. Davon werden 80% (d.h. maximal 8000 Stichproben) für das Training und je 10% für die Validierung und als Testdatensatz eingesetzt. Um den Einfluss der Stichprobenanzahl im Trainingsdatensatz auf die Leistung des Neuronalen Netzwerkes beurteilen zu können, wurden noch zusätzliche Datensätze generiert mit maximal 4000 bzw. 18'000 Stichproben pro Kategorie im Trainingsdatensatz. Dies war natürlich nur für die Klassen möglich, für die auch ausreichend Stichproben vorhanden waren. Validierungs- und Testdatensätze wurden dabei nicht geändert, um die Resultate der verschiedenen Trainingsvorgänge vergleichen zu können. Dabei zeigte sich, dass das Training mit grösseren Datensätzen weniger gut ausbalanciert war, als wenn das Training mit kleineren und damit über die Klassen ausgeglicheneren Stichprobenmengen durchgeführt wurde. Am deutlichsten zeigte sich dieser Effekt in den deutlich niedrigeren Genauigkeiten für die Klassen mit zu kleinen Stichprobenmengen. Der Vorteil der grösseren Stichprobenmengen ist hingegen, dass eine grössere Breite möglicher Zustände beschrieben und damit auch klassifiziert werden können. Ein zweistufiges Verfahren, bei welchem zunächst sehr kleine Klasse mit einer grossen ähnlichen Klasse gepoolt werden und anschliessend ein zweites Netzwerk zur Differenzierung eingesetzt wird, wäre optional vorstellbar (siehe Kapitel 5).

2.4.4 Mögliche Einstellungen/Handling

Das Training des CNN ist ein stochastischer Prozess. Die endgültige Netzwerkleistung (d.h. die Endwerte der Verlustfunktion und der Gesamtgenauigkeit siehe Abb. 2.11) hängt von vielen Faktoren ab, z.B. der Mini-Batchgrösse, Grösse der extra Dense Layer oder Grösse des Trainingsdatensatzes. Diese Faktoren werden als Hyperparameter bezeichnet. Die Werte der Hyperparameter werden empirisch ermittelt. Das Training wird hierbei

zunächst mit einem Satz Hyperparameter durchgeführt und die Leistung auf dem Validierungssatz berechnet. Dieses erste Experiment dient als Baseline. Im folgenden Trainingslauf wird jeweils ein Wert des Hyperparameters geändert und die resultierende Leistung mit der Baseline verglichen. Führt die Änderung des Parameters zur Leistungssteigerung (d.h. Steigerung der Gesamtgenauigkeit), wird dieser neue Wert in den nachfolgenden Experimenten verwendet, andernfalls wird der Parameterwert auf den der Baseline zurückgesetzt. Dieser Prozess wird wiederholt, bis die Werte der optimalen Hyperparameter gefunden sind. Die Implementierung des CNN-Moduls ermöglicht es, die folgenden Hyperparameter einzustellen:

- Bildgrösse in Pixel (sollte der Kachelgrösse des ‘Preprocessing’ Moduls entsprechen)
- Mini-Batchgrösse: Zweierpotenzen, z.B. 16, 32, 64
- Extra Dense Layer Grösse: es wird empfohlen (aber keine Notwendigkeit) die Zweierpotenzen zu verwenden (z.B. 1024, 2048, 4096); 0 (Null) bedeutet, dass der zusätzliche Dense Layer vollständig weggelassen wird.
- Dense Layer(s) Dropout: Prozentsatz der Dense Layer - Knoten, die während des Trainings zufällig abgeschaltet werden. Dies hilft, ein Overfitting zu vermeiden. Werte von 0 bis 1 (empfohlen: 0.4-0.6)
- Anzahl der "nicht eingefrorenen" Faltungsblöcke, die an der Feinabstimmung teilnehmen: es handelt sich dabei um die ‘Middle Flow Blocks’ (Abbildung 2.10). Mögliche Werte: 0 bis 8. 0 (0 = alle Blöcke sind während der Feineinstellung eingefroren, 8 = keine Blöcke sind eingefroren).
- Optimizer (mehr Information dazu: https://keras.io/optimizers/)
  - Unterschiedliche Optimizer für ‘Dense Layer Training’ und Feineinstellung können verwendet werden.
  - Optimizer Parameter: mögliche Parameter hängen vom Typ des Optimizers ab (siehe Keras Dokumentation für weitere Informationen).
  - Falls der Optimizer den Parameter ‘Learning Rate’ (lr) unterstützt, ist es möglich, die Cyclic Learning Rate zu verwenden, um das Training zusätzlich zu beschleunigen. Es handelt sich um eine neue Methode zur Optimierung der Lerngeschwindigkeit (Morris 2018).
  - Bild-Augmentationen (Default-Einstellungen der Optimizer: z.B. Rotation, Zoom, Spiegelung)

Es wurden verschiedene Optimizer ausgetestet darunter *Nadam* (default - siehe Abb. 2.11) und der auf der Cyclic Learning Rage basierende SGD (siehe Kapitel 4.2).

```
1   augmentation: {}
2   batch_size: 32
3   dataset: {name: area14_fcir_lu_10000, type: LU}
4   dense:
5     dense_layer_size: 0
6     dropout: 0.5
7     early_stop_patience: 0
8     epochs: 1
9     optimizer:
10      name: nadam
11      params: {}
12  finetune:
13    early_stop_patience: 4
14    epochs: 20
15    optimizer:
16      clr: 1
17      clr_params: {base_lr: 0.0004, epochs_per_half_cycle: 2, max_lr: 0.002}
18      name: sgd
19      params: {momentum: 0.9}
20    trainable_blocks: 6
21  im_size: [200, 200]
22  model: {experiment: 01_block6, name: area14_fcir_lu_10000}
23  plot: 1
24  validation:
25    plot: 1
26    threshold: 0.5
27    top_n: [1, 5]
28  verbose: 2
```

Abbildung 2.11. Beispiel der Einstellungen eines Testdurchlaufes mit dem Nadam Optimizer.

Die Resultate der Trainingsdurchläufe werden pro Trainingsvorgang in einem separaten Ordner abgespeichert. In einer Exceltabelle wird dabei der Genauigkeitsverlauf festgehalten und eine Konfusionsmatrix erstellt, welche sich als sehr wertvoll für Interpretation der Ergebnisse erwiesen hat. Hilfreich für die Auswertung sind auch die ausgegebenen Statistiken pro Kategorie. Diese wurden im Rahmen des MS3 in digitaler Form (inkl. der dazugehörigen Modelle) dem BFS übergeben. Die Zusammenfassung der Trainingsdurchläufe und daraus

folgende Erkenntnisse sind im Kapitel 4.2 zusammengefasst. Hinsichtlich der Reproduzierbarkeit ist es grundsätzlich möglich, das CNN mit einem fixierten Seed-Value auszustatten. Jedoch ist die Framework-Seitige Implementierung der doppelten Seed-Funktion (Globaler Graph und Operational Level) unter Keras nur für Trainingsdurchläufe vorgesehen, die auf der CPU statt auf der GPU durchgeführt werden, da insbesondere die Optimizer-Modelle auf GPUs nicht deterministisch berechnet werden. CPU-basiertes Training würde aber bei den vorliegenden Datenmengen prohibitive Berechnungszeiten erreichen.

2.5 Post-Klassifikator: Random Forest

Das RF-Modul wurde direkt in der „R“-Umgebung ausgeführt, weshalb die direkte Dokumentation der Eingabeparameter als Kommentare im R-Script vorliegen. Das Preprocessing der Eingabedaten muss flexibel an die jeweiligen Bedürfnisse angepasst werden und wird daher über die interaktive IDLE-Umgebung durchgeführt.

2.5.1 Methodenwahl

Es existieren mehrere mögliche Modelle, komplexe Datensätze zu klassifizieren. Zu den verbreitetsten Methoden, welche regelmässig gute Klassifikationsergebnisse erbringen, gehören Random Forest, K-Nearest-Neighbor (kNN), Generalized Linear Model (GLM) und Support Vector Machine (SVM). Die Wahl fiel auf die RF-Methode vor allem deshalb, weil diese eher für Multiklassenprobleme grosser Datensätze geeignet ist, während SVM eher für binäre Klassifikationsprobleme konzipiert wurde. kNN kann zwar auch für grosse Datensätze verwendet werden, benötigt aber wesentlich mehr Rechenzeit als ein RF bei gleichbleibender Präzision. GLM ist schnell und einfach, weist im Vergleich mit anderen Modellen aber den grossen Nachteil im Workflow auf, dass die Verteilung der Sample-Datensätze zuerst geplottet werden muss, um die Variablen anschliessend manuell in eine hierarchische Interaktions-Ordnung zu bringen. Geschieht dies nicht, sind häufig substanzielle Verluste in der Präzision, insbesondere bei der Klassifikation komplexer Datensätze, hinzunehmen. Für GLMs muss ein (linearer) funktionaler Zusammenhang der Variablen zumindest vermutet werden können, während ein Random-Forest-Modell weniger Voraussetzungen an die Variablenzusammensetzung stellt. Für eine AIC-basierte Modellselektion befindet sich das Projekt noch in einer explorativen Phase, in welcher an anderer Stelle noch massgeblicheres Optimierungspotential vorhanden ist (siehe Sensitivitätsanalyse).

Um Multiklassenprobleme mit SVM umzusetzen, müsste der Klassifikationsvorgang zunächst auf mehrere binäre Klassifikationen reduziert werden, was die Trainingsdauer verlängert und die Präzision verringert. Darüber hinaus maximiert SVM den „margin“ (dt. in etwa „Spielraum“) zwischen verschiedenen multidimensionalen Punkten und nutzt deren euklidische Distanz zur Errechnung einer Klassifikationsentscheidung. Dies verursacht einen starken Bias hin zu kategorisierten Variablen, welche zunächst über das sog. One-Hot-Encoding einzeln in 0/1-Entscheidungen aufgetrennt werden müssen. Alle numerisch kodierten Variablen müssen im Preprocessing ausserdem einer Min-Max-Skalierung unterworfen werden, um homogen in den Entscheidungsprozess einfliessen zu können. Um die multidimensionale Matrix aus Samples multipliziert mit den Features aufzubauen, steigt der Speicherbedarf der SVM bei grossen Datensätzen dabei sehr stark an. Als Faustregel kann eine Nutzbarkeitsgrenze bei etwa 10‘000 Samples im Training gezogen werden. Für direkte Bilddaten-Interpretation ist SVM gut geeignet, da die Pixel eines Bildes gut in die homogene Datenstruktur passen, für komplexe Datensätze jedoch eher weniger.

RF hingegen ist vielen dieser Limitationen nicht unterworfen: Im Verlauf dieser Studie konnten Datensätze mit bis zu 665‘000 Samples auf 32GB RAM trainiert werden. Ausserdem funktioniert RF sehr gut mit einer Mischung aus numerischen und kategorischen Variablen (solange diese ansatzweise aussagekräftig sind) und das Training funktioniert wesentlich schneller als bei Verwendung eines kNN-Algorithmus. Ein weiterer Vorteil ist die grössere Interpretierbarkeit der Resultate: RF-Berechnungen erlauben im Gegensatz zu kNN und SVM sowohl eine direkte Berechnung der Variablengewichte und gestatten somit eine Einschätzung, wie sich das Klassifikationsergebnis zusammensetzt. Zudem können trotzdem, wie bei anderen Methoden auch, unkompliziert Prädiktions-Probabilistiken für jedes einzelne Sample generiert werden. Zu bemerken ist, dass die Aussage der Variablengewichte qualitativ der Ausgabe einer PCA vergleichbar ist, mit der die Anteile an die Varianz einzelner Komponenten erklärt wird. Die Variablengewichtungen werden jedoch im RF anhand der Gini-Koeffizienten als normierte Häufigkeit berechnet, mit der eine Variable für eine Baumteilung und die Folge dieser Teilung verantwortlich ist. Teilungen sind als einzelne Entscheidungen innerhalb eines Entscheidungsbaums aufzufassen.

Die Abschätzung von Variablengewichten ermöglicht es, durch die Selektion einflussreicher Variablen eine Optimierung des Klassifikators hinsichtlich Rechenleistung herbeizuführen.

2.5.2 Systemvoraussetzungen

Die Umsetzung basierte auf einer systemnahen C++ Binary, welche über für den Nutzer verständliche Kommandos in der frei verfügbaren Statistikumgebung „R“ aufgerufen werden kann. Als ausführende ML-Metabibliothek wird ranger über das caret-Package angesprochen. Die verschiedenen RF-Implementierung weisen hierbei keine grundlegenden funktionellen Unterschiede in der erreichbaren Präzision auf. Vielmehr ist die erreichbare Präzision von der Performanz abhängig, da eine grosse Anzahl an Bäumen und Splits erzeugt werden soll. Aufgrund der eigenständig laufenden C++ Binary von ranger sind bei diesem prozeduralen Aufruf gegenüber dem direkten Aufruf keine Leistungseinbussen zu erwarten. Das Script kann so auch direkt in einer Umgebung wie R Studio ausgeführt werden. Ranger ist insbesondere für hochdimensionale Datensätze optimiert und zeigt gegenüber anderen RF-Implementierungen eine sehr hohe Geschwindigkeit bei gleichzeitig hoher Speichereffizienz. Für die volle Funktionalität sind zusätzlich die Bibliotheken e1071, lubridate, magrittr, mlbench und tidyverse notwendig. Ein moderner Prozessor (Multicore CPU) beschleunigt den Trainingsvorgang. In Verbindung mit einem sehr grossen Arbeitsspeicher (Empfehlung 64 GB+) können ausserdem grössere Datensätze „gleichzeitig“ statt batchweise in Training und Prädiktion einfliessen.

2.5.3 Funktionsweise

Ein RF ist aus vielen Entscheidungsbäumen mit numerischen Werten aufgebaut. Diese unkorrelierten Entscheidungsbäume werden dabei zufällig initialisiert. Während mehrfachen Iterationen soll dann die Entropie minimiert werden. Dabei darf jeder Baum für eine Klassifikation eine Entscheidung treffen, das Ergebnis mit den meisten «Stimmen» entscheidet am Ende über die endgültige Klassifikation. Insgesamt sind mindestens 1000 Entscheidungsbäume (plus vom Datensatz abhängig dynamisch generierte Extrabäume) im eingesetzten Random Forest Post Klassifikator eingebunden.

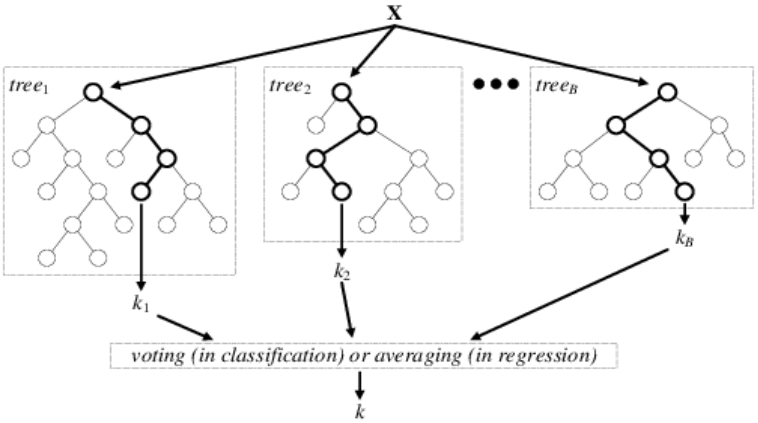


Abbildung 2.12. Schematischer Ausschnitt aus einem Random Forest

2.5.4 Datenvorbereitung und Sampling

Für die Auswertungen im Random Forest wurden bisher die in Tabelle 2.7 aufgelisteten Daten integriert und getestet. Die Integration weiterer Datensätze ist möglich, hängt jedoch von deren Vollständigkeit und Aktualität ab. Insbesondere das Gebäude- und Wohnungsregister GWR wäre für die Klassifikation interessant, wurde jedoch aufgrund unbekannter oder nicht synchroner Zeitstände innerhalb des Datensatzes bisher nicht vorprozessiert. Die Verkettung der Daten zu trainierbaren CSV-Dateien erfolgt mittels kurzer Python-Scripts und interaktiver Codezeilen unter Verwendung der Bibliothek *pandas*. Als Eingabedatensatz können dazu grundsätzlich alle knapp 4 Mio. möglichen Samples der Schweiz verwendet werden, indem die verschiedenen Eingangsdatensätze anhand der RELI-ID verkettet werden. Nachfolgend werden die verwendeten Daten im Detail beschrieben.

Tabelle 2.7. Daten, die im Rahmen des Projektes im RF Verwendung fanden (% = Abdeckung des Datensatzes, falls keine 100% Abdeckung).

Eingangsdaten Bezeichnung	Datentyp	AREA3	AREA4
CNN-Klassifikationswahrscheinlichkeiten RGB	Numerisch (0.00-1.00)	Ja	Ja (67%)
TOP1 CNN-Klassifikationsresultat RGB	Faktor	Ja	Ja (67%)
CNN-Klassifikationswahrscheinlichkeiten FCIR	Numerisch (0.00-1.00)	Nein	Ja (37%)
TOP1 CNN-Klassifikationsresultat FCIR	Faktor	Nein	Ja (37%)
Zeitreihen auf Satellitenbasis (Landsat): 12 spektrale Bänder und Indizes in 10 Punkten verteilt über ein Jahr	Numerisch (0-255)	Ja	Ja
Höhendaten (DEM) und Ableitungen	Numerisch (-1000,00 - +1000,00)	Nein	Ja
Vegetationshöhenmodell (Canopy Height) und Ableitungen	Numerisch (-1000,00 - +1000,00)	Nein	Ja
Bodenbedeckung aus der Amtlichen Vermessung	Faktor	Nein	Ja
Koordinaten (geographische Länge/Breite)	Numerisch (float)	Ja	Ja

Die Integration weiterer Datensätze ist möglich, hängt jedoch von deren Vollständigkeit und Aktualität ab. Insbesondere das Gebäude- und Wohnungsregister GWR wäre für die Klassifikation interessant, wurde jedoch aufgrund unbekannter oder nicht synchroner Zeitstände innerhalb des Datensatzes bisher nicht vorprozessiert. Die Verkettung der Daten zu trainierbaren CSV-Dateien erfolgt mittels kurzer Python-Scripts und interaktiver Codezeilen unter Verwendung der Bibliothek *pandas*. Als Eingabedatensatz können dazu grundsätzlich alle knapp 4 Mio. möglichen Samples der Schweiz verwendet werden, indem die verschiedenen Eingangsdatensätze anhand der RELI-ID verkettet werden. Nachfolgend werden die verwendeten Daten im Detail beschrieben.

Output CNN

Für die Szenarien AREA3-RGB und AREA4-RGB, sowie AREA4 und -FCIR wurde das jeweils optimale DL-Modell verwendet (dasjenige mit der höchsten Overall Acuuracy), um per Inferencing einen symmetrischen Datensatz möglichst über die ganze Schweiz zu generieren, der jeweils identische RELI-IDs (Stichprobenpunkt-Zuweisungen) enthält. Die probabilistischen Klassifikationswerte werden dabei für jede Klasse einzeln angegeben. Zusätzlich wird die TOP1-Klasse als faktorisierte Variable und die zugehörige probabilistische Confidence ins Training miteinbezogen. Dies kommt einer verdoppelten Gewichtung des jeweils stärksten Ergebnisses aus der DL-Klassifikation gleich. Für das RF-Modell mit Zusatzdaten werden AREA4-RGB- und AREA4-FCIR-Klassifikationsergebnisse aus dem DL verkettet, für das RF-Modell der Veränderungsanalyse AREA3-RGB- und AREA4-RGB-Klassifikationsergebnisse.

AV-Daten

Die AV-Daten wurden auf eine faktorisierte Variable mit 26 möglichen Levels kombiniert (lc_lcsurface). Faktorisierte Variablen werden für das Training im Random Forest durch das ranger-Modul automatisch anhand ihrer Levels in Spalten aufgetrennt und per «One-hot-Encoding» in die Werte 0 und 1 konvertiert. Für den RF standen also 26 Variablen auf Basis der Bodenbedeckung der AV zur Verfügung.

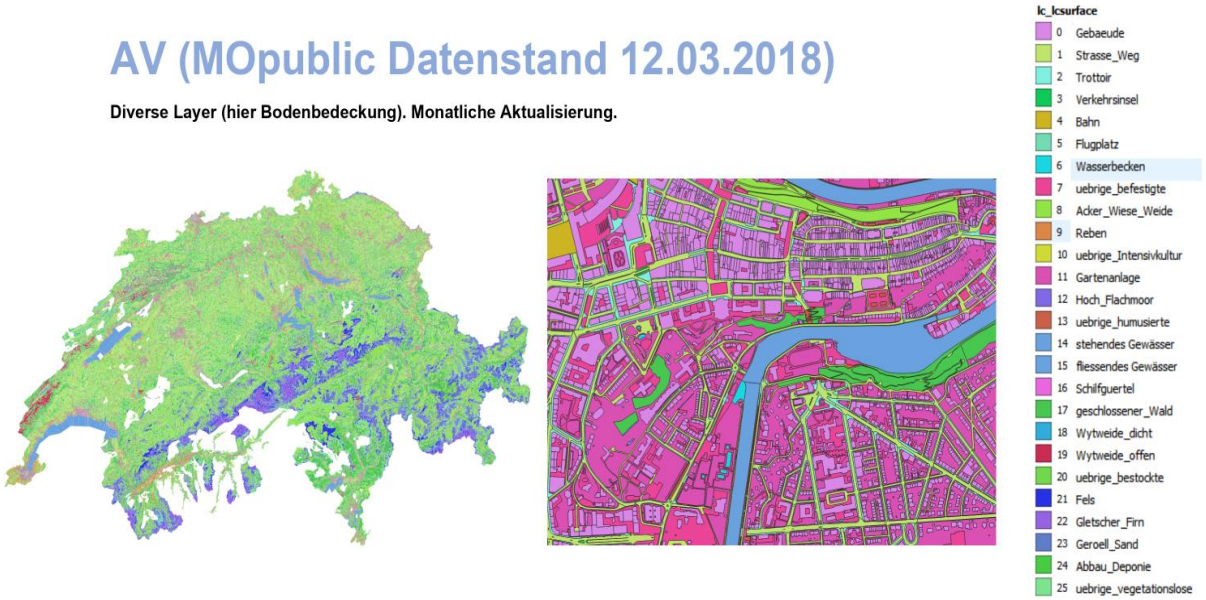


Abbildung 2.13. Aufteilung der AV-Bodenbedeckung in 26 Klassen.

Satellitendaten

Die Satellitendaten werden primär dafür herangezogen, um den Mehrwert temporaler Eigenschaften, wie zum Beispiel der saisonalen Vegetationsdynamik, in den Klassifikationsprozess zu integrieren. Dazu wurden 12 Informationsebenen im RF integriert (vgl. Kapitel 2.3), welche jeweils 10 interpolierte Zeitstände verteilt über das Jahr der Aufnahme beinhalten. Insgesamt stehen für jede Stichprobe also 120 Variablen im 8-Bit-Wertebereich (ganzzahlig, 0-255) aus der temporalen Satellitenklassifikation zur Verfügung (siehe Kap. 2.3.2).

Höheninformationen (DEM und CHM)

Die für den Random Forest verwendeten Höhendaten basieren grösstenteils auf SwissALTI3D (DEM, siehe Kap. 2.3.3) und dem Vegetationshöhenmodell der WSL (CHM – Canopy Height Model). Die Daten wurden in einem möglichst aktuellen Stand für die Klassifizierung von AREA4-Modellen verwendet. Als Trainingsvariablen wurden folgende Grössen (jeweils in Bezug zur der Bildkachel) verwendet: Mittlere Höhe, Standardabweichung der Höhe, Minimale Höhe, Maximale Höhe, Slope (Steilheit) und Aspect (Ausrichtung der Hangneigung) (vgl. Kapitel 2.3.3). Der verwendete Wertebereicht ist double, z.B. eine DEM-Höhenangabe in Meter mit Nachkommadezimalen: 417,63. Für die Variable Aspect (-180° bis 180°) sind auch negative Werte möglich.

Koordinaten

Ein Einsatz der direkten Koordinaten als Klassifikationsinput wurde getestet, verschlechterte jedoch das Klassifikationsergebnis marginal, da keine Korrelationen mit anderen Features gefunden wurden. Dies bedeutet auch, dass der Einfluss horizontal-räumlicher Zusammenhänge auf die Klassifikationsergebnisse vernachlässigt werden kann. Der Ansatz wurde daher nicht weiterverfolgt.

Inputschema der verwendeten Daten

In Abbildung 2.14 ist der Header einer typischen Multisensor-Trainingstabelle für den RF dargestellt. Das One-Hot-Encoding der faktorisierten Variablen wird durch ranger/caret vorgenommen und muss nicht in das CSV geschrieben werden. Sind faktorisierte Variablen (insbesondere Klassenzugehörigkeiten aus dem BFS-Schema) jedoch als Zahlwerte gespeichert, so muss ein Buchstabe (z.B. „x“) hinzugefügt werden, um eine Interpretation der Variable als Faktor/String zu erzwingen.

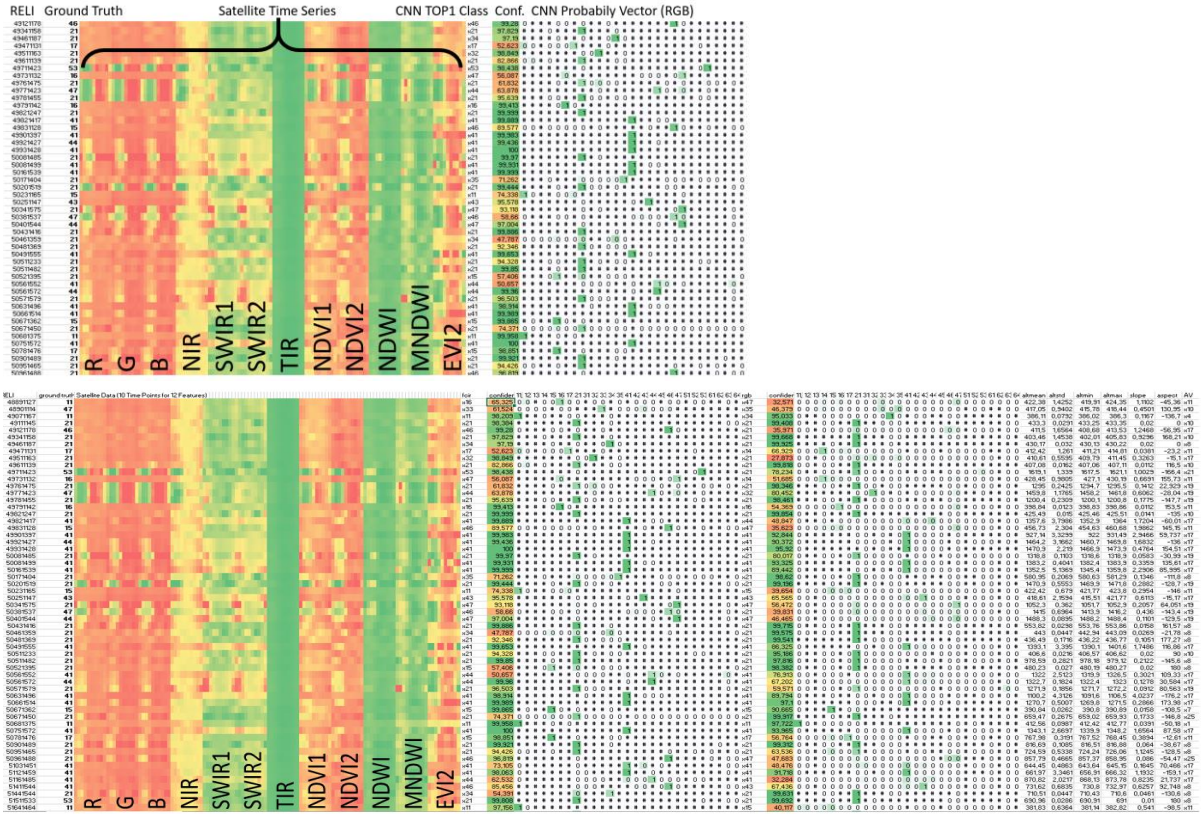


Abbildung 2.14. Beispielkomposition möglicher Eingabedatensätze für den RF: Konfiguration Satelliten-Zeitreihen und CNN (RGB) ohne Zusatzdaten (oben); Konfiguration Satellitenzeitreihen, CNN (RGB), CNN (FCIR), DEM und AV (unten).

2.5.5 Mögliche Einstellungen/Handling

Grundsätzlich können mit dem «*caret*»-Modul unter R eine Vielzahl verschiedener RF-Klassifikationsalgorithmen genutzt werden. Als die in vielen Bereichen performanteste Lösung gilt derzeit «ranger». Bei allen wachsenden relevanten Parametern wie Anzahl Features, Anzahl Samples, Anzahl Bäume und Anzahl Splitting Features wird eine höhere Performanz hinsichtlich Prozessierungszeit als bei ähnlichen Packages erreicht (siehe Abb. 2.15). Wichtig zu erwähnen ist, dass sich hingegen die unterschiedlichen Algorithmen in Bezug auf das Klassifikationsverhalten bei gleichbleibender Parameterbesetzung kaum unterscheiden. Deshalb kann der zeitliche Faktor für die Trainingszeiten als wesentlicher Teil der Algorithmenwahl und losgelöst von der Klassifikationsleistung betrachtet werden.

Bei der «caret»-Implementierung von «ranger» werden per Default n=25 Bootstrap-Samples gezogen. Ohne Einstellung des bootstrap-Parameters dauert ein Trainingsvorgang deshalb 25x so lange wie ein einmaliges Training mit dem vollständigen Datensatz. Von allen Merkmalen M (Features oder Dimensionen, im höherdimensionalen Model Fit befinden sich regelmässig M>200 Variablenelemente im Datensatz) werden für die Knotenpunkte eine Auswahl (Subset) an Merkmalen zufällig ausgewählt, die für die Aufspaltung (Split) infrage kommen. Per Default werden jeweils drei M-abhängige Modelle verglichen (Variable mtry):

- Nur 2 Variablen dürfen sich pro Baum ändern.
- die Hälfte aller Variablen dürfen sich pro Baum verändern.
- alle Variablen dürfen sich in allen Bäumen verändern.

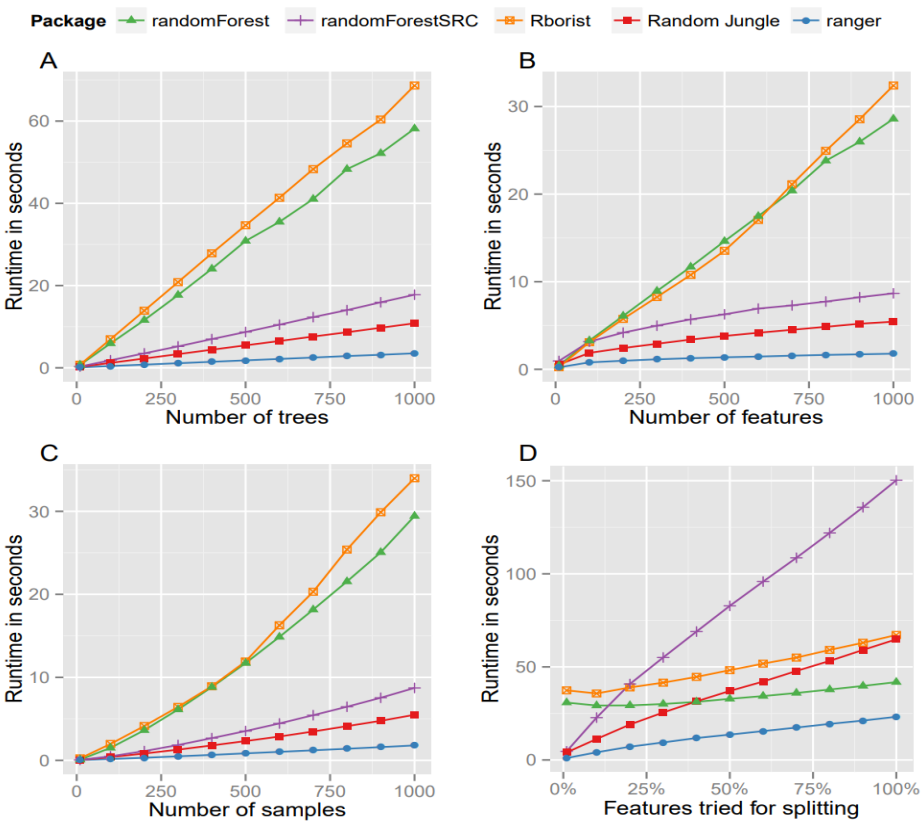


Abbildung 2.15. Vergleich verschiedener RF-Implementierungen, ranger in blau (Wright & Ziegler 2015).

Mit jeder Iteration kann die Entropie minimiert werden, bis ein Baum voll ausgebaut ist. Wird der Parameter M nicht eingestellt, resultiert dies in einer Verdreifachung der Prozessierungszeit. Stellen sich bestimmte Variablen als besonders wichtig heraus, so können an diesen Splitting Points nach vordefinierten Regeln auch neue Bäume erzeugt werden. Wird nur eine dieser Erzeugungsregeln („gini“ oder „extratrees“) eingesetzt, so halbiert sich die Trainingsdauer. Im Falle einer Klassifikation wird die wahrscheinlichste Klasse einer Stichprobe durch eine Mehrheitswahl aus allen Entscheidungsbäumen bestimmt. In diesem Verfahren wird auch eine probabilistische Einschätzung der Klassifikationspräzision sowohl nach Stichproben als auch nach Klassen oder Features (Variable Importance) möglich. Gegenüber anderen Random-Forest-Implementierungen ist ranger/caret in der Literatur gegenüber ähnlichen CPU-basierten Umsetzungen als hochperformantes Tool im Benchmarking aufgefallen. Für alle vier Kerngrössen (Anzahl Bäume, Anzahl Variablen, Anzahl Samples, Anzahl Variablen für Knotensplit) benötigt ranger/caret dabei weniger Berechnungszeit als vergleichbare Packages (Abbildung 2.15).
Die Anzahl Samples in Training und Prädiktion beeinflussen Genauigkeit und Trainingsgeschwindigkeit des Modells. Mehr Samples resultieren grundsätzlich in höherer Genauigkeit, sowie längerer Trainingsdauer. Bei 32 GB verfügbarem Arbeitsspeicher beträgt das Limit einlesbarer Samples je nach Anzahl Variablen ca. 1 – 1,5 Mio Zeilen. Diese Grenze kann für die Prädiktion durch batchweises Vorgehen umgangen werden. Für die Anzahl maximal trainierbarer Samples ist die Obergrenze jedoch hardwareabhängig. Für das Training wurde in der Regel zwischen 250‘000 und 700‘000 Stichproben randomisiert aus der Grundgesamtheit gezogen. Bereits im CNN für Training oder Validierung genutzte Samples wurden zuvor aus der Grundgesamtheit ausgenommen.

2.5.6 Parameter-Optimierung

In der frühen Projektphase wurden alle Modelle mit drei verschiedenen Variabel-Knotensplits (M) und unter Austesten der Split-Rule trainiert. Dabei zeigte sich, dass der Wechsel der Splitting Rule einen maximalen Einfluss von ±1% an der Gesamt-Accuracy ausmacht. Zumeist schnitt dabei das „extratrees“-Modell leicht besser ab, als das „gini“-Modell. In der späteren Projektphase wurde daher die Splitting Rule auf „extratrees“ festgelegt, was die Berechnungszeit halbierte.

Für M zeigte sich nach den ersten Trainingsdurchläufen, dass aufgrund der grossen Anzahl Variablen im Training Modelle mit nur zwei sich ändernden Variablen pro Baum wesentlich schlechter abschnitten, als Modelle mit mehr Variablen. Die besten Gesamt-Accuracy-Werte wurden mit ca. 60% der Anzahl Inputvariablen erreicht, wobei die Verbesserungen oberhalb von 50% Variableneinsatz mit maximal ca. ±2% Verbesserung ebenfalls marginal ausfielen. Es ist daher nicht notwendig, M weiter zu iterieren, sofern es bei einem Wert von ca. 50-80% der Inputvariablensatzbreite festgelegt wird. Beispiel: Werden 200 Variablen trainiert, so kann M=120 gesetzt werden.

Das Bootstrapping hat keinen Einfluss auf die Gesamtgenauigkeit, erhöht durch das Resampling allerdings die statistische Aussagekraft der finalen Metriken. Für Parameter- oder Variablentests muss es deshalb nicht verwendet werden, für ein produktives Modell empfiehlt sich jedoch eine Evaluation mit Bootstrapping.

2.5.7 Experimentelles Vorgehen

Für Land Cover und Land Use wurden jeweils getrennt folgende Klassifikationsmodelle mit dem Random Forest berechnet:

- AREA 3: CNN RGB + Satellite Time Series
- AREA 4: CNN RGB + Satellite Time Series
- AREA 4: CNN RGB + CNN FCIR + Satellite Time Series + DEM + CHM + AV

Dabei wird für jedes Modell zunächst der passende Inputdatensatz durch relationale Joins der Daten-CSVs mit pandas erzeugt. Anschliessend wird der Datensatz mit dem R-Script mit zufälliger Verteilung (ein „Seed“ zum Fixieren des Zufallsgenerators ist möglich) in einen Test- und einen Trainingsdatensatz aufgetrennt. Während des Trainings berechnet das R-Script kontinuierlich Permutationsmatrizen, aus welchen letztlich die Variablengewichte abgeleitet und exportiert werden. Nach dem Training wird das generierte Klassifikationsmodell gespeichert und es werden Prädiktionen durchgeführt, mit welchen eine Konfusionsmatrix erstellt und verschiedene statistische Auswertungen vorgenommen werden.

Für die Veränderungsanalyse (Change Detection) ist das Vorgehen ähnlich, es gibt jedoch je nach Basisvoraussetzung drei verschiedene Ansätze:

- Change Detection durch direkte Anwendung des bestmöglichen Klassifikationsmodells

  Hierbei wird das Klassifikationsresultat des Random Forest mit der letzten Arealstatistik-Erhebung tabellarisch verglichen. Unter der Annahme, dass die Resultate der letzten Erhebung korrekt sind, entspricht die Overall Accuracy dieses Ansatzes somit der erreichbaren Overall Accuracy im Klassifikationsmodell.

- Change Detection AREA3-AREA4: Tabellarischer Ansatz

  Für den tabellarischen Ansatz werden die Resultate der AREA3- und der AREA4-Klassifikationsmodelle aus dem Random Forest verglichen. Es ist also nicht notwendig, die Ground-Truth-Informationen (insbesondere des ersten Durchlaufs) über den Trainingsdatensatz hinaus kennen zu müssen.

- Change Detection AREA3- AREA4: Zusätzliches dichotomes Random-Forest-Modell

  Für ein zusätzliches Random-Forest-Modell wird der Datensatz mit einer neuen Information „Change“ / „No Change“ versehen und auf dieses Merkmal traininert. Der Datensatz muss in diesem Fall jedoch ausgeglichen sein (50%/50%), da typischerweise ungleich mehr Stichproben (>90%) keinen Klassenwechsel aufweisen. Dies limitiert die Anzahl der zur Verfügung stehenden Trainingsdaten stark.

Alle vorgestelllten Change-Detection-Ansätze basieren derzeit auf den Klassifikationsresultaten des CNN. Eine Alternative wäre eine frühere Fusion der Inputparameter aus beiden Zeitschnitten z.B. durch eine gemeinsame Klassifikation des Deep Feature Vectors. Ein entsprechender Versuch ohne Zusatzdaten unter Zuhilfenahme einer Support-Vector-Machine (SVM) erreichte jedoch nur unzureichende Ergebnisse. Der Einbezug der Zusatzdaten verbessert die Klassifikationsleistung so stark, dass die Verwendung eines reinen CNN-Ansatzes zu diesem Zeitpunkt hinfällig erscheint. Ein Experiment für eine Change Detection unter Verwendung der Deep Feature Vektoren plus den Zusatzdaten ist vorstellbar.

Für den Fall der dichotomen Klassifikation ist die Verwendung einer SVM sicherlich noch untersuchenswert. GLM kann als Option vernachlässigt werden, da wiederum möglichst lineare, bekannte Zusammenhänge der Eingabedaten bekannt sein sollten. Die Wahl für den Random Forest gegenüber anderen Modellen fiel demnach aus Zeit- und Flexibilitätsgründen, da weniger Annahmen in der explorativen Phase getätigt werden müssen.Eine Erweiterung des direkten, sowie des dichotomen Ansatzes wäre darüberhinaus noch durch den Einbezug der Ground-Truth-Daten der vorangehenden Arealstatistik-Erhebungen denkbar. Für die Anwendung des bestmöglichen Klassifikationsmodells in der Veränderungsanalyse gelten die Evaluationsparameter des jeweiligen Modells, weshalb sie im Kapitel 4.3 nicht gesondert aufgeführt sind. Die notwendige Grundannahme für die Anwendung ist, dass die Ground-Truth-Ergebnisse der vorherigen Arealstatistikerhebung bekannt sind. Die beiden anderen Modelle kommen ohne diese Annahme aus.

3 Sensitivitätsanalysen

Die erfolgreiche Einbindung des Prototyps in einem operationellen Ablauf und die Qualität bzw. der Nutzen der daraus resultierenden Ergebnisse sind von der Parametrisierung des Systems und der verwendeten Eingangsdaten in hohem Masse abhängig. Aufgrund der Komplexität des Systems und der Vielfalt der zur Verfügung stehenden Eingangsdaten ist eine umfangreiche Analyse der entsprechenden Sensitivitäten notwendig. Die Sensitivitätsanalysen sind hierbei sowohl in Hinblick auf die System-Performance (z.B. die Effizienz vom CNN/RF in Bezug auf die Laufzeiten und Ressourcennutzung - vgl. Kapitel 2.4.4 und Kapitel 2.5.5) als auch für die Beurteilung möglicher Abhängigkeiten der Klassifikationsgenauigkeiten von Bedeutung. Um ein diesbezüglich strukturiertes Vorgehen umzusetzen, sind sowohl die sensitiven Eingangsparameter zu spezifizieren als auch die für die Evaluierung notwendigen Kennzahlen zu bestimmen (vgl. Tabelle 3.1).

Tabelle 3.1. Szenarien möglicher Sensitivitätsanalysen, wobei weder die Eingangsparameter noch die Kennzahlen einen Anspruch auf Vollständigkeit haben.

Kategorie	Eingangsparameter	Erläuterung	Kennzahl
System	CNN	Parametrisierung CNN (z.B. Lernrate, Anzahl eingefrorene Blöcke, Augmentationen)	Performance-Metriken
	RF	Parametrisierung RF (z.B. Bootstrap-Samples, Anzahl Bäume)	Performance-Metriken
Sampling	Klassenspezifisch	Training CNN und/oder RF nur mit ausgewählten Klassen	Genauigkeitsmetriken
	Hierarchisch	Training CNN und/oder RF nur basierend auf Zuverlässigkeitskategorien	Genauigkeitsmetriken
	Kombination aus klassenspezifisch und hierarchischen Sampling		
Daten	Datenauswahl Zusatzdaten (allgemein)	Training mit einer bestimmten Auswahl an Zusatzdaten (iterative Selektion in Abhängigkeit des Einflusses)	Genauigkeitsmetriken
	Datenauswahl Zusatzdaten (spezifisch)	Training mit ausgewählten Kategorien eines Datensatzes (z.B. ausgewählte Informationsebenen der AV)	Genauigkeitsmetriken
	Datenauswahl Luftbild	Grösse der Samplefläche, Anzahl Samples	Genauigkeitsmetriken Performance-Metriken
	Kombination aus allgemeiner und spezifischer Datenauswahl		Genauigkeitsmetriken Performance-Metriken
Validierung	Vorhersage-wahrscheinlichkeit	Durchführung der Validierung für bestimmte Kategorien der Vorhersagewahrscheinlichkeiten	Genauigkeitsmetriken

Um die Herausforderung bzw. die Komplexität der Sensitivitätsanalysen zu veranschaulichen, wird nachfolgend das Vorgehen für das Beispiel “Anzahl Samples” skizziert (Abbildung 3.1).

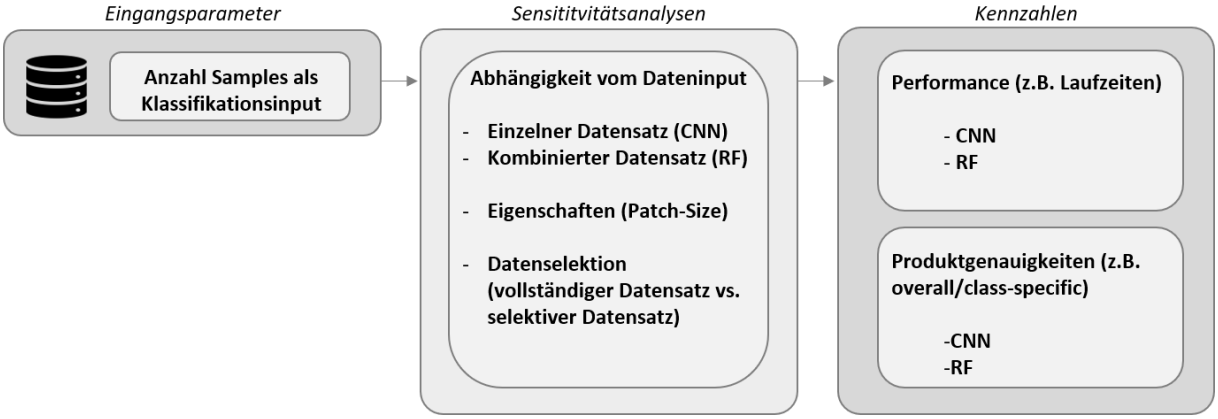


Abbildung 3.1. Mögliche Sensitivitätsanalysen für den Eingangsparameter Anzahl Samples.

Wenn der Einfluss der Sampleanzahl analysiert werden soll, also welche Grundmenge an Daten optimal für das Training des Klassifikator und/oder die Güte der Produkte ist, wären folgende Auswahl an Sensitivitätsanalysen für eine umfassende Betrachtung möglich (nicht vollständig):

1) Welchen Einfluss hat die Menge an Training-Samples auf die Performance von CNN und RF?
2) Welchen Einfluss hat die Menge an Training-Samples auf die Klassifikationsergebnisse (differenziert für die Gesamtgenauigkeiten und die spezifischen Klassengenauigkeiten)?
3) Wir verändern sich die Kennzahlen unter 1) und 2), wenn die Samples in ihrer Art verändert werden, z.B. Anpassung der Grösse der Referenzflächen?
4) Wie verändern sich die Kennzahlen unter 1) und 2) in Abhängigkeit von der Dimensionalität des Sample-Datensatzes (einzelner Datensatz, z.B. nur RGB Luftbilder, vs. kombinierter Datensatz, z.B. Luftbilder + Satellitenbilder + Hilfsdatensatz)
5) Wie verändern sich die Kennzahlen unter 1) und 2), wenn mit einem selektiven Datensatz trainiert werden soll, z.B. Verwendung des CHM ausschliesslich für Vegetationsklassen oder Berücksichtigung nur derjenigen Samples, die in einem vorgängigen Schritt hohe Vorhersagewahrscheinlichkeiten hatten?

Die jeweiligen Analysen sind hierbei miteinander verbunden. Dies bedeutet, sobald man z.B. die Analyse 3) durchführen möchte, man dies idealerweise für die jeweiligen Varianten von 1), 2), 4) und 5) umsetzt. Dementsprechend hätte man allein für den Eingangsparameter "*Sample Size*" bereits 25 für sich stehende Analysen, unabhängig davon, wie differenziert man den Eingangsparameter variiert (d.h. wie viele "Kategorien" an Sample Sizes man untersucht, z.B. 5k, 10k, 20k Samples). Von daher wäre eine vorhergehende Plausibilisierung bzw. Gewichtung der Sensitivitäten angebracht, um den zeitlichen Aufwand in einem durchführbaren Rahmen zu belassen. So kann zum Beispiel bei sehr redundanten Datensätzen (SWISSIMAGE RGB vs. SWISSIMAGE FCIR) oder nur geringfügigen Veränderungen der Sampleeigenschaften (z.B. Veränderung der Grösse der Referenzflächen von 50x50 m auf 40x40 m) von einem sehr ähnlichen Verhalten bezüglich der Sensitivitäten ausgegangen werden. Zudem sind einige der oben beschriebenen Kennzahlen auto-korreliert, da zum Beispiel die Produktgenauigkeiten vom CNN die RF-Produktgenauigkeiten direkt beeinflussen, und können unter Umständen entsprechend von den Analysen ausgeschlossen werden.

4 Ergebnisse

Im Rahmen der vorliegenden Studie wurden verschiedene Klassifikationen ausgeführt, wobei deren Fokus primär auf der Evaluierung der Funktionsweise des Prototyps AI und der Plausibilität der Ergebnisse lag. Dementsprechend sind die Interpretationen der Ergebnisse zum jetzigen Zeitpunkt als vorläufig einzustufen. Die folgende Übersicht zeigt die unterschiedlichen Klassifikationen mit den jeweiligen Basis-Konfigurationen (Zielklassen, Inputdaten & Klassifikator).

Klassifikation	CNN	RGB	FCIR	SAT	AV	CHM	ALT	AREA3	AREA4
LC CNN		x						x	
		x							x
			x						x
LU CNN		x						x	
		x							x
			x						x
LC RF	x	X (CNN)		x				x	
	x	X (CNN)		x					x
	x		X (CNN)	x	x	x	x		x
LU RF	x	X (CNN)		x				x	
	x	X (CNN)		x					x
	x		X (CNN)	x	x	x	x		x
LULC RF	x	X (CNN)	X (CNN)	x	x	x	x		x
CD RF (tabellarisch)	x	X (CNN)		x				x	
CD RF (dichotom)	x	X (CNN)		x				x	

4.1 Validierung

Die Validierung erfolgt basierend auf den Ausgabetabellen der CNN- und/oder RF-Klassifikation und ist im Validierungs-Modul als vollautomatisierter Prozess integriert. Neben der normierten/nicht-normierten Konfusionsmatrix (Abbildung 4.1 und 4.2) werden die darauf basierenden Metriken Overall Accuracy (OA), User's und Producer's Accuracy UA/PA) bzw. Error of Omission/Comission, Kappa-Koeffizient (KK) und das F-Mass als Tabelle und farblich kodierte Abbildung ausgegeben.

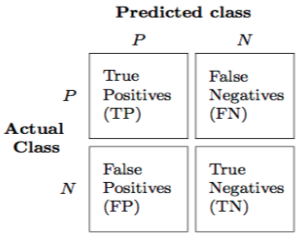


Abbildung 4.1. Konzept der Konfusions- bzw. Fehlermatrix.

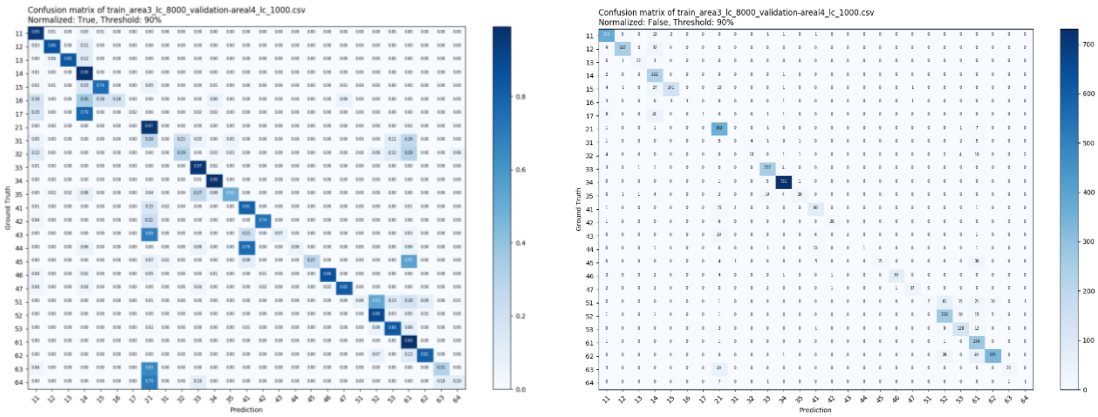


Abbildung 4.2. Beispiel für eine normierte (links) und nicht normierte Konfusions-Matrix (rechts).

Das F-Mass (*F*) kombiniert Genauigkeit (*Precision*) und Trefferquote (*Recall*) mittels des gewichteten harmonischen Mittels:

$$Precision = \frac{tp}{tp + fp}$$

$$Recall = \frac{tp}{tp + fn}$$

$$F = 2\frac{Precision * Recall}{Precision + Recall}$$

Die Berechnung von OA, UA, PA und dem KK orientiert sich an den Formeln in Abbildung 4.3. Die OA wird gebildet, indem die Summe der richtig klassifizierten Stichprobenpunkte durch die Gesamtanzahl der Stichprobenpunkte dividiert wird. Die UA berechnet sich aus den richtig klassifizierten Stichprobenpunkte durch die Gesamtzahl der Stichprobenpunkte der jeweiligen Klasse. Die PA berechnet sich aus den richtig klassifizierten Stichprobenpunkte (eine Zahl aus der Hauptdiagonalen) durch die Gesamtzahl der Referenz-Stichprobenpunkte der jeweiligen Klasse (Spaltensumme). So erhält man die Wahrscheinlichkeit, mit der die Stichprobenpunkte der Klassifikation mit den Referenzdaten übereinstimmen.

Predicted Class \ Actual Class	A	B	C	D	Σ
A	n_{AA}	n_{AB}	n_{AC}	n_{AD}	n_{A+}
B	n_{BA}	n_{BB}	n_{BC}	n_{BD}	n_{B+}
C	n_{CA}	n_{CB}	n_{CC}	n_{CD}	n_{C+}
D	n_{DA}	n_{DB}	n_{DC}	n_{DD}	n_{D+}
Σ	N_{+A}	N_{+B}	N_{+C}	N_{+D}	n

$$OA = \frac{\sum_{i=1}^{q} n_{ii}}{n} \times 100$$

$$UA = \frac{n_{ii}}{n_{i+}} \times 100$$

$$PA = \frac{n_{ii}}{n_{+i}} \times 100$$

$$Kappa\ Coefficient = \frac{n\sum_{i=1}^{q} n_{ii} - \sum_{i=1}^{q} n_{i+}n_{+i}}{n^2 - \sum_{i=1}^{q} n_{i+}n_{+i}}$$

q = number of classes
n = total data samples
n_{ii} = amount of agreeing samples
n_{i+} = value of row totals
n_{+i} = value of column totals

Abbildung 4.3. Berechnung der Klassifikations-Metriken basierend auf der Konfusions-Matrix.

Der KK ist ein Maß für die Beurteilung der Gesamtgenauigkeit. Im Unterschied zur OA geht er davon aus, dass die erstellte und die Referenzklassifizierung von gleicher Zuverlässigkeit sind. Dementsprechend werden bei seiner Berechnung auch die Unter- und Überlassungsfehler berücksichtigt - d.h., man kann darüber Aussagen treffen, wie gut die Klassifizierungsergebnisse mit den Referenzdaten übereinstimmen. Der KK berechnet sich aus der Differenz zwischen den richtig klassifizierten Stichprobenpunkten und einer nicht korrekten Übereinstimmung. Das Ergebnis wird durch den theoretisch möglichen Anteil, der über den Zufall hinausgehenden Übereinstimmungen dividiert. In der folgenden Formel sieht man, dass der KK alle Elemente der Konfusionsmatrix enthält. Der Kappa-Koeffizient kann Werte in einem Intervall von -1 bis 1 annehmen, wobei eine komplette Übereinstimmung von Klassifizierung und Referenzdaten in den Wert 1 und bei keiner Übereinstimmung in den Wert -1 resultieren würden. In der Literatur werden Werte von über 0,8 als ein sehr hohes Maß der Übereinstimmung gewertet. Durch die hohe Abhängigkeiten der Klassifikationsgenauigkeiten von den verschiedenen Faktoren bezüglich System und Daten (vgl. Kapitel 3), liefern die oben beschriebenen Metriken lediglich Annäherungswerte der Klassifikationsgüte, die jedoch nicht als quantitative Masse Gültigkeit haben.

Als Basis-Output des Validierungsmoduls werden ein Übersichtsreport und eine Übersichtsgrafik ausgegeben (Abbildung 4.4). Zusätzlich besteht die Möglichkeit, falls vorhanden, die Klassenwahrscheinlichkeit der LU- und LC-Klassifikationen zu verknüpfen und eine Abbildung der mittleren LULC-Genauigkeiten zu generieren (Abbildung 4.5), wobei aber zu beachten ist, dass die Interpretation dieser Werte unter Berücksichtigung der jeweiligen spezifischen LU und LC Werte erfolgen sollte und diese Art der Berechnung mehr als ein Anhaltspunkt und weniger als ein statistisch aussagekräftiges Genauigkeitsmass zu verstehen ist. Darüber hinaus werden die

kumulierten Verteilungsfunktionen der Klassenwahrscheinlichkeiten für jede Klasse ausgegeben (Abbildung 4.6). Da die räumliche Auswertung keine erkenn- oder messbaren Muster zeigte (vgl. Abbildung 4.7), wurde auf eine Integration dieser prozessierungsintensiven Methode verzichtet, um die Leistung des Modules nicht unnötig zu verringern. Dies kann jedoch bei Bedarf problemlos integriert werden.

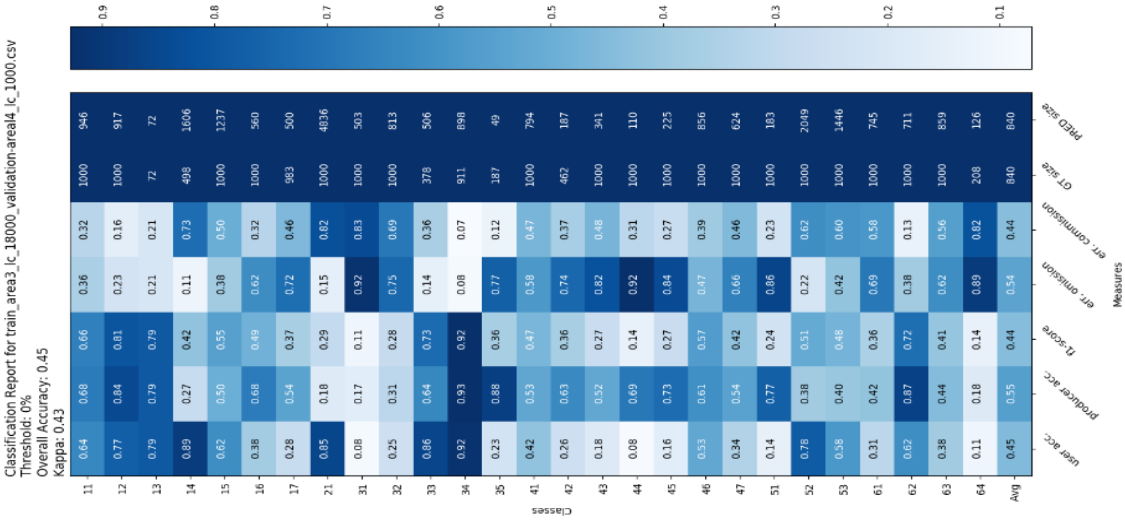


Abbildung 4.4. Beispiel der Übersichtsgrafik als Ausgabe des Validierungsmodules.

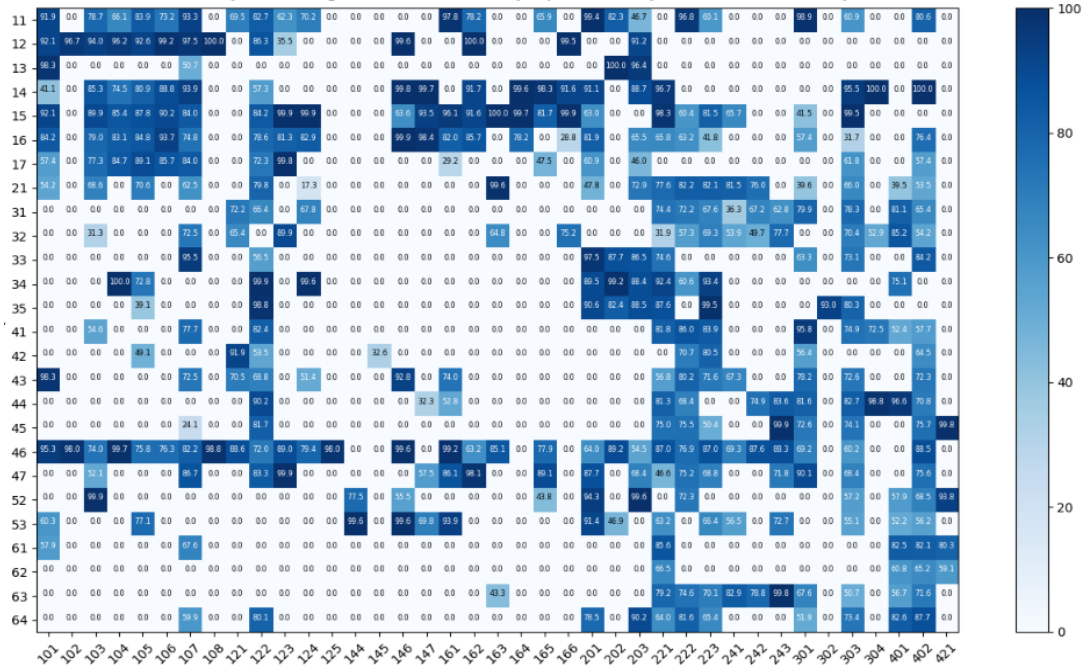


Abbildung 4.5. Beispiel für die Verknüpfung von LULC - Produktgenauigkeiten.

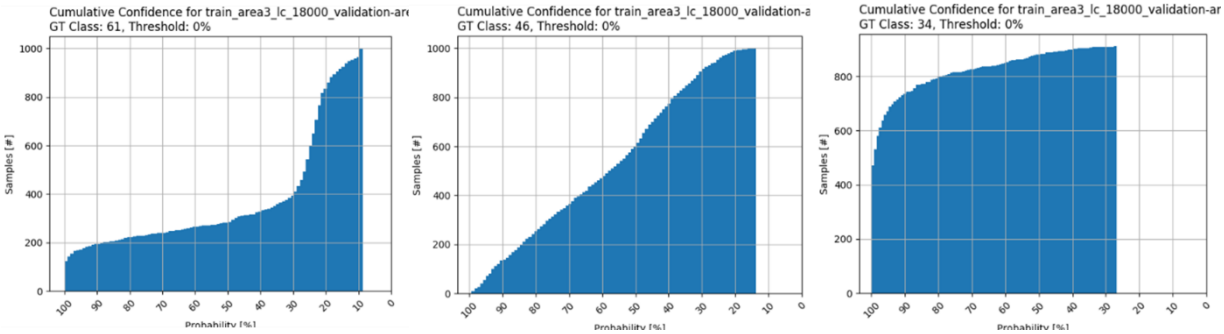


Abbildung 4.6. Beispiel für kumulierte Verteilungsfunktionen der Klassenwahrscheinlichkeiten.

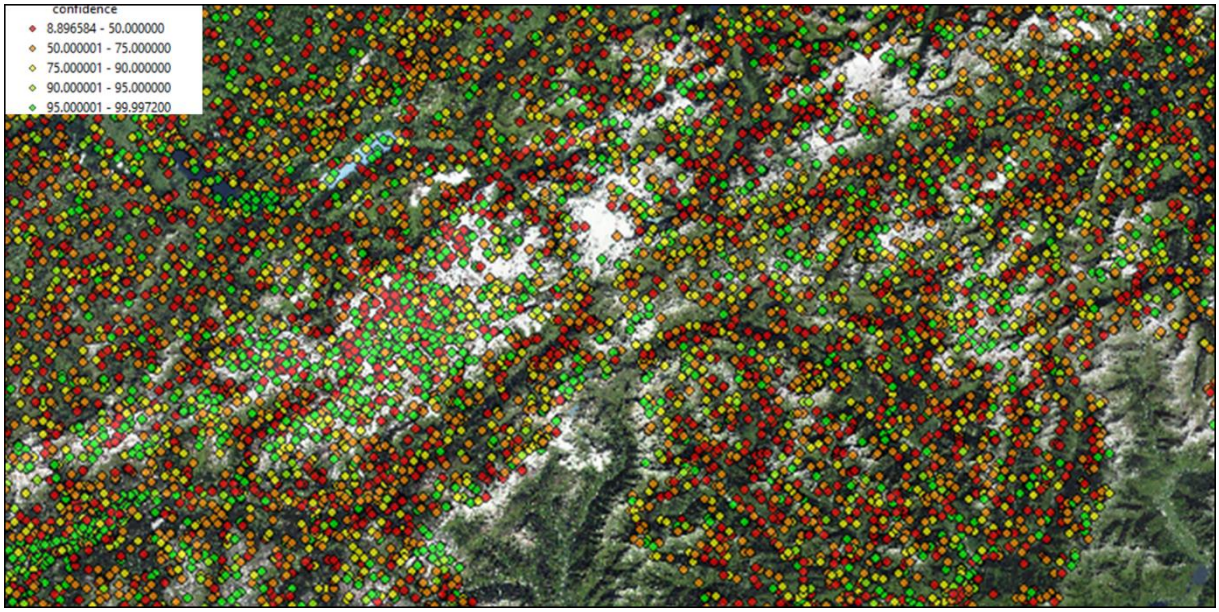


Abbildung 4.7. Beispiel für die räumliche Verteilung der Vorhersagewahrscheinlichkeiten.

Zudem erlaubt das Validierungsmodul auch Kreuzvergleiche (Cross-Comparison) zwischen verschiedenen Eingangsdatensätzen (z.B. zwischen der Arealstatistik und der AV). Dies ist insofern von Bedeutung, um das Potential einer Einbindung von Zugangsdaten in die Klassifikation und/oder in das Plausibilisierungsvorgehen abschätzen zu können (Abbildung 4.8). So kann man am Beispiel der AV sehen, dass diese nur für sehr wenige Klassen (z.B. Reben/Rebbau) in einer direkten Plausibilisierung Verwendung finden könnte und bestimmte Klassen der AV, wie zum Beispiel Acker_Wiese_Weide, starke Überlappungen mit fast allen LU-Klassen der Arealstatistik aufweisen und diese somit nur bedingt unterstützend im Rahmen eines automatisierten Klassifikationsvorgehen eingebunden werden können. Dementsprechend ist es empfehlenswert, über vorhergehende Sensitivitätsanalysen und gezielte Kreuzvergleiche nur die thematischen Informationen zu verwenden, die einen deutlichen Klassenbezug zur Arealstatistik aufweisen (z.B. die Verwendung eines reduzierten AV-Datensatzes).

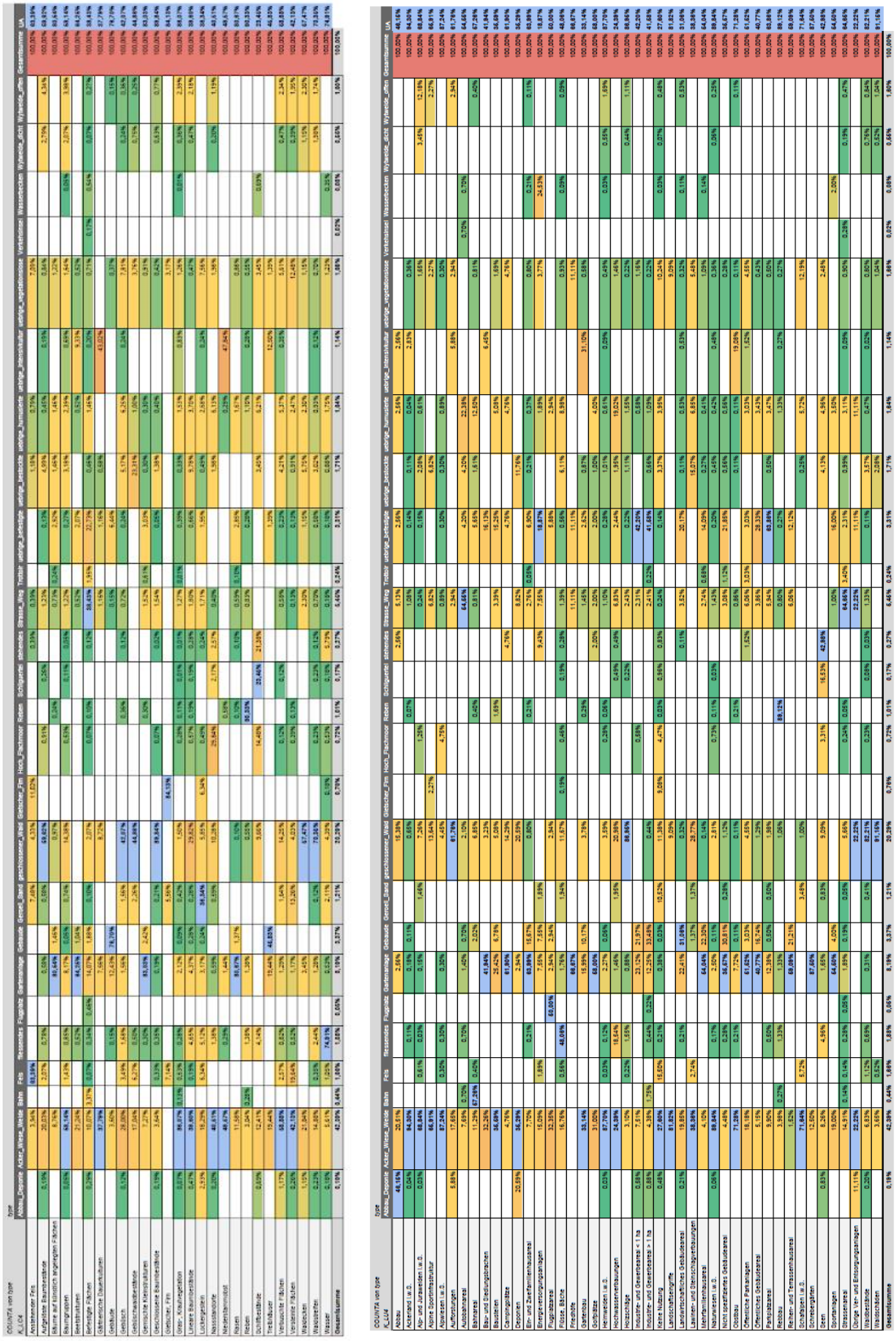


Abbildung 4.8. Beispiel für die Cross-Comparison von AV und Arealstatistik.

Für eine strukturierte Validierung ist ein entsprechend systematisches Vorgehen notwendig. Ähnlich zu den Sensitivitätsanalysen können hierbei unterschiedliche Ansätze gewählt werden, wobei grundsätzlich zwischen der Validierung des Systems und der Validierung der Klassifikationsgenauigkeiten unterschieden werden muss (Abbildung 4.9).

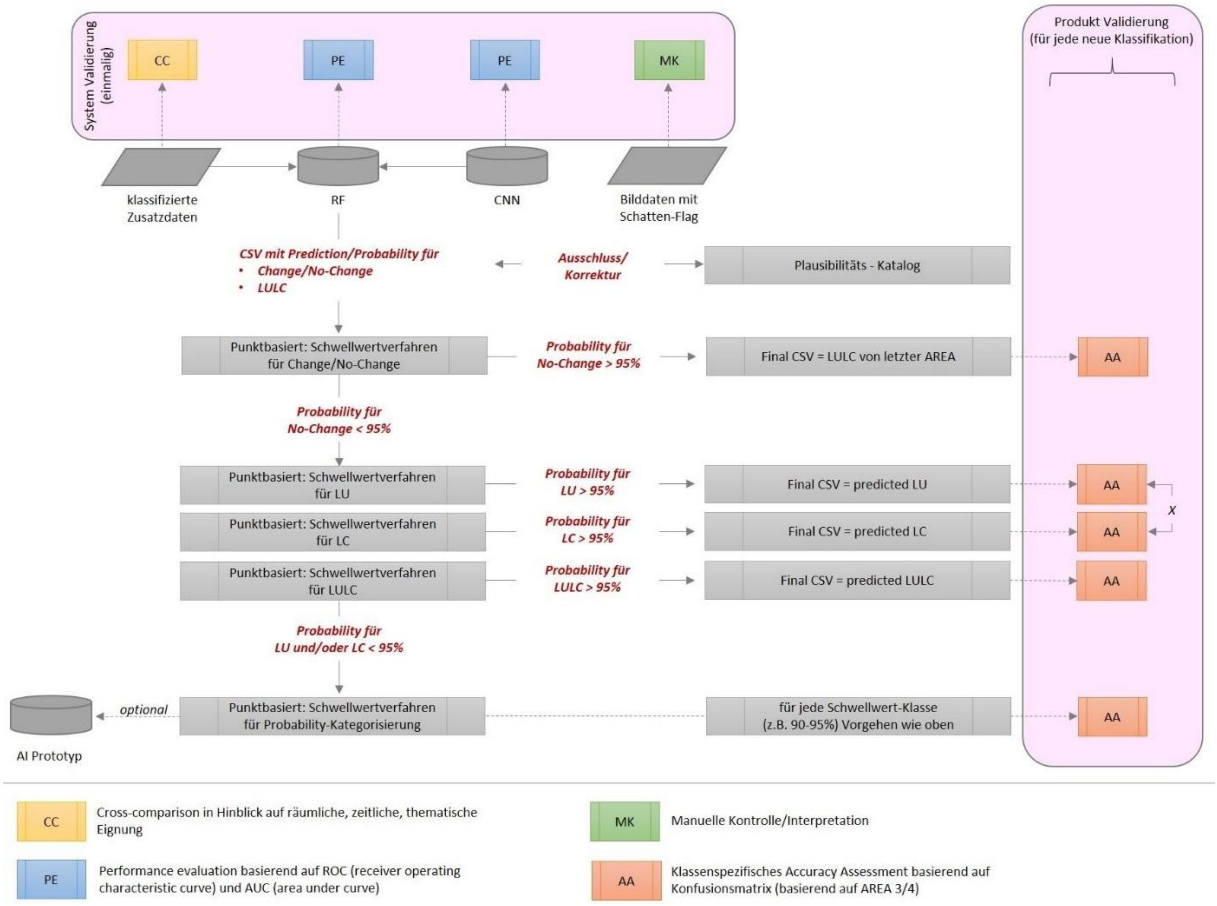


Abbildung 4.9. Mögliche Strategie für ein strukturiertes Vorgehen bei der Validierung.

Wie in Abbildung 4.9 ersichtlich ist, kann grundsätzlich zwischen einer klassen- und einer stichprobenpunktspezifischen Validierung unterschieden werden. Klassenspezifisch bedeutet, dass Aussagen über die thematische Genauigkeit für die jeweilige LULC-Klasse ausgegeben werden, wobei dies in Form der mittleren Genauigkeiten über alle Stichprobenpunkte erfolgt. Eine stichprobenpunktspezifische Validierung beruht auf der jeweiligen Vorhersagegenauigkeit für jeden einzelnen Punkt unabhängig der thematischen Zuweisung und einer anschliessenden Analyse der klassenspezifischen Verteilungsmuster/Eigenschaften. Dies hätte den Vorteil, dass man nur diejenigen Stichprobenpunkte in der Validierung berücksichtigt, die vom System her mit hohen Vorhersagewahrscheinlichkeiten klassifiziert wurden (vgl. Kapitel 3). Damit können zum Beispiel klassifizierte Stichprobenpunkte mit hohen Vorhersagewahrscheinlichkeiten im Rahmen des Erhebungsprozesses berücksichtigt werden, obwohl diese vielleicht zu Klassen gehören, die eher geringe mittlere Genauigkeit aufweisen.

Um eventuelle Fehlinterpretationen auszuschliessen, ist zudem eine sorgfältige Trennung von Daten, die im Rahmen der Klassifikation als Trainingsdaten verwendet wurden, und Daten, die ausschliesslich im Rahmen der Validierung Verwendung finden, notwendig. Dies kann insbesondere bei den Klassen problematisch werden, die nur mit sehr geringen Häufigkeiten auftreten.

4.2 CNN

4.2.1 Experimentelle Ergebnisse

Die ausführlichste Untersuchung zu den Hyperparametern wurde für die RGB LC Klassifikation durchgeführt. Die reduzierten Trainingsdaten wurden für die meisten Trainings verwendet (z.B. für einen Datensatz mit über 4000 Stichproben pro Kategorie). Trainingsdurchläufe für kleinere Datensätze benötigen weniger Zeit (vgl. Tabelle 4.1 und 4.2) und erlauben schnellere Iterationen über alle möglichen Hyperparameter.

Tabelle 4.1. Laufzeiten für Trainingsdurchläufe.

DS	Type	1 epoch	Full run (8 epochs)
4000	LU	~ 20 min	~ 3 h
	LC	~ 15 min	~ 2 h
8000	LU	~ 35 min	~ 5 h
	LC	~ 30 min	~ 4 h
18000	LU	~ 60 min	~ 9 h
	LC	~ 50 min	~ 8 h

Tabelle 4.2. Inference – Laufzeiten für den vollständigen Datensatz.

Data	Probabilities	Deep Features
AREA3 RGB	6.5 h	9 h
AREA4 RGB	4 h	5.5 h
AREA4 FCIR	3 h	4 h

Die aufgrund von reduzierten Datensätzen gefundenen, besten Einstellungen für Hyperparameter wurden für das Training von grösseren Datensätzen angewendet. Die Resultate sind in Tabelle 4.3 ersichtlich. Die Einstellungen für die Basislinie sind in Tabelle 4.4 aufgelistet. Grundsätzlich wurde für jedes Experiment jeweils nur ein einzelner Parameterwert verändert. Sobald diese Änderung zu einer besseren Performance führte, wurde der Wert beibehalten. Zum Beispiel wurde das Experiment mit '01-03 mini-batch size' für die Werte 64, 31 und 128 getestet. Da mit einer Batchgrösse von 32 die beste Performance erzielt wurde, sind alle folgenden Experimente mit dieser Einstellung weitergeführt worden. Die Parameterwerte, welche zur Steigerung der Performance beitragen, sind in Orange herausgehoben, die beste, erzielte Performance in Grün. Die besten für die Klassifikation RGB LC gefundenen Hyperparameter sind in Tabelle 4.5 aufgelistet (nur Parameter mit Abweichung von der Baseline sind aufgelistet).

Die gleichen Experimente wurden für die RGB LU und FCIR LULC - Klassifikationen durchgeführt. Die Ergebnisse sind in den Tabellen 4.4 bis 4.13 dargestellt.

Tabelle 4.3. Hyperparameter-Tuning für die RGB LC - Klassifikation.

DS	Experiment	Hyperparameters	Epochs	Overall accuracy
4000	01 baseline	baseline	8	0.46
	02 batch32	batch size 32	7	0.49
	03 batch128	batch size 128	7	0.45
	04 dense4096	extra dense layer size 4096	11	0.44
	05 dense1024	extra dense layer size 1024	9	0.45
	06 blocks0	trainable blocks 0	8	0.46
	07 blocks1	trainable blocks 1	7	0.46
	08 blocks3	trainable blocks 3	8	0.48
	09 blocks4	trainable blocks 4	7	0.47
	10 blocks5	trainable blocks 5	8	0.51
	11 blocks6	trainable blocks 6	6	0.52
	12 blocks7	trainable blocks 7	11	0.52
	13 blocks8	trainable blocks 8	7	0.45
	14 aug1	image augmentations: horizontal flips	7	0.52
	15 aug2	image augmentations: vertical flips	6	0.48
	16 aug3	image augmentations: rotations up to 90°	8	0.45
	17 aug4	image augmentations: zoom (range=0.2)	6	0.50
8000	01 best4000	dataset size: up to 8000 per category	11	0.54
18000	01 best4000	dataset size: up to 18000 per category	8	0.51

Tabelle 4.4. Baseline Hyperparameter für die RGB LC - Klassifikation.

Dataset size	up to 4000 samples per category in training set
Mini-batch size	64
Extra dense layer size	0 (no extra dense layer)
Dropout	0.5
Dense layer pretraining params	
Epochs	1
Optimizer	Nadam, learning rate=0.005
Finetuning params	
Epochs (max)	20
Early stop patience	4
Trainable blocks	2
Optimizer	SGD, cyclic learning rate (base_lr=0.01, max_lr=0.1, epochs_per_half_cycle: 2), momentum=0.9
Image augmentations	none

Tabelle 4.5. Beste Hyperparameter für RGB LC - Klassifikation.

Dataset size	up to 8000 samples per category in training set
Mini-batch size	32
Trainable blocks	6
Image augmentations	horizontal flips or none

Table 4.6. Hyperparameter - Tuning für die RGB LU - Klassifikation.

DS	Experiment	Hyperparameters	Epochs	Overall accuracy
4000	01 baseline	baseline	7	0.43
	02 batch32	batch size 32	5	0.44
	03 block5	trainable blocks 5	7	0.47
	04 block6	trainable blocks 6	8	0.44
	05 block7	trainable blocks 7	6	0.49
	06 block8	trainable blocks 8	8	0.45
	07 block4	trainable blocks 4	5	0.42
	08 block7 aug	trainable blocks 7 + horizontal flips	8	0.46
	09 block5 aug	trainable blocks 5 + horizontal flips	6	0.47
8000	01 block7 aug	dataset size: up to 8000 per category, trainable blocks 7 + horizontal flips	9	0.48
	02 block5	dataset size: up to 8000 per category, trainable blocks 5	6	0.46
	03 block7	dataset size: up to 8000 per category, trainable blocks 7	7	0.50
18000	01 block7	dataset size: up to 18000 per category, trainable blocks 7	8	0.46

Tabelle. 4.7. Baseline Hyperparameter für die RGB LU - Klassifikation.

Dataset size	up to 4000 samples per category in training set
Mini-batch size	64
Extra dense layer size	0 (no extra dense layer)
Dropout	0.5
Dense layer pretraining params	
Epochs	1
Optimizer	Nadam, learning rate=0.002
Finetuning params	
Epochs (max)	20
Early stop patience	4
Trainable blocks	2
Optimizer	SGD, cyclic learning rate (base_lr=0.005, max_lr=0.05, epochs_per_half_cycle: 2), momentum=0.9
Image augmentations	none

Tabelle 4.8. Beste Hyperparameter für die RGB LU - Klassifikation.

Dataset size	up to 8000 samples per category in training set
Mini-batch size	32
Trainable blocks	7

Tabelle 4.9. Hyperparameter - Tuning für die FCIR LC - Klassifikation.

DS	Experiment	Hyperparameters	Epochs	Overall accuracy
4000	01 blocks5	baseline	7	0.52
	02 blocks6	trainable blocks 6	6	0.54
	03 blocks7	trainable blocks 7	8	0.51
8000	01 blocks6	dataset size: up to 8000 per category	17	0.56

Tabelle 4.10. Baseline Hyperparameter für die FCIR LC - Klassifikation.

Image size	200x200 px
Dataset size	up to 4000 samples per category in training set
Mini-batch size	32
Extra dense layer size	0 (no extra dense layer)
Dropout	0.5
Dense layer pretraining params	
Epochs	1
Optimizer	Nadam, learning rate=0.002
Finetuning params	
Epochs (max)	20
Early stop patience	4
Trainable blocks	5
Optimizer	SGD, cyclic learning rate (base_lr=0.001, max_lr=0.006, epochs_per_half_cycle: 2), momentum=0.9
Image augmentations	none

Tabelle 4.11. Beste Hyperparameter für die FCIR LC - Klassifikation.

Dataset size	up to 8000 samples per category in training set
Trainable blocks	6

Tabelle 4.12. Hyperparameter tuning for FCIR LU classification.

DS	Experiment	Hyperparameters	Epochs	Overall accuracy
4000	01 blocks5	baseline	6	0.51
	02 blocks6	trainable blocks 6	6	0.51
	03 blocks7	trainable blocks 7	6	0.51
8000	01 blocks6	dataset size: up to 8000 per category	5	0.49

Tabelle 4.13. Baseline Hyperparameter für FCIR LU - Klassifikation.

Image size	200x200 px
Dataset size	up to 4000 samples per category in training set
Mini-batch size	32
Extra dense layer size	0 (no extra dense layer)
Dropout	0.5
Dense layer pretraining params	
Epochs	1
Optimizer	Nadam, learning rate=0.002
Finetuning params	
Epochs (max)	20
Early stop patience	4
Trainable blocks	5
Optimizer	SGD, cyclic learning rate (base_lr=0.0004, max_lr=0.002, epochs_per_half_cycle: 2), momentum=0.9
Image augmentations	none

4.2.2 Zusammenfassung Erkenntnisse

Aufgrund der 46 durchgeführten Trainingstestläufe konnte erkannt werden, dass die Bild-Augmentationen (Spiegelungen, Rotationen usw.), womit bei unterrepräsentierten Kategorien die Stichprobenzahl erhöht werden sollte, nur sehr beschränkt helfen, das Gesamtergebnis zu verbessern. Sie führen hingegen zu zeitlichen Verzögerungen des Prozesses.

Es ist üblich, das Training zu beenden, wenn die Gesamtgenauigkeit des Validierungsdatensatzes sich nicht mehr verbessert. Es gibt aber gewisse Schwankungen zwischen den Epochen aufgrund der Zufälligkeiten des Trainingsprozesses: es kann passieren, dass die Gesamtgenauigkeit vorübergehend sinkt und nach einigen Epochen wieder ansteigt. Um einen zu frühen Abbruch des Prozesses zu vermeiden, wird ‘early stopping patience’ (Abbruch-Geduld) verwendet: das Training läuft einige Epochen weiter, um sicher zu stellen, dass die höchstmögliche Gesamtgenauigkeit erreicht ist. Bei Einstellungen ‘Maximal 20 Epochen trainieren’ und ‘Abbruch-Geduld 4 Epochen’ wird das Training meist bereits nach 6-8 Epochen beendet (Abbildung 4.10). Es wird dasjenige Modell verwendet, das den niedrigsten Validation-Loss innerhalb der trainierten Epochen aufweist.

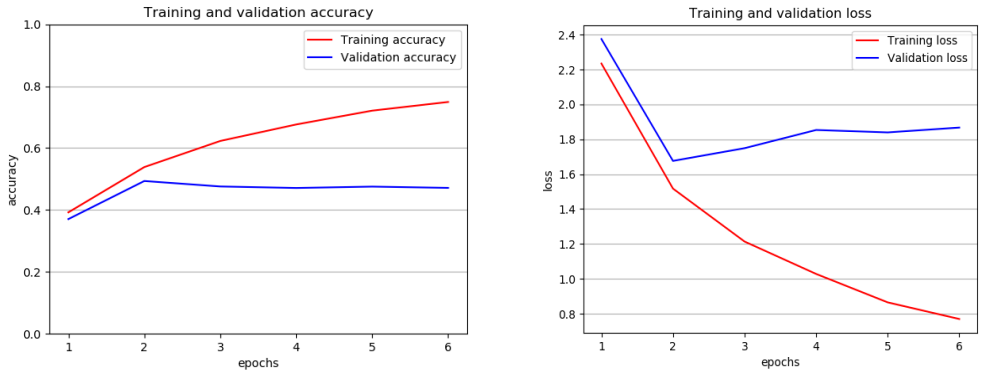


Abbildung 4.10. Gesamtgenauigkeit (links) und ‘Loss’ über sechs Epochen getrennt für den Validierungs- und Trainingsdatensatz.

Das heisst, die maximale Gesamtgenauigkeit wird nach 2-4 Epochen erreicht. Bei einem untrainierten Netzwerk (=Training von Null) würde dieser Prozess viel länger dauern. Da es sich in unserem Fall um ein vortrainiertes Netz handelt, wird die optimale Genauigkeit deutlich früher erreicht, was viel kürzere Rechenzeiten zur Folge hat. Als weitere Möglichkeit zur Auswertung könnten auch Schwellenwerte eingeführt werden und damit die Kategorie 'unklassiert' hinzugefügt werden. Mit dieser Massnahme könnte die Genauigkeit der verbleibenden Kategorien gesteigert werden, weil sich die Wahrscheinlichkeiten nicht mehr auf so viele Kategorien verteilen würden. Dasselbe könnte auch mit einem 'Top5'- oder 'Top3-'Ansatz erreicht werden, wobei die fünf bzw. drei höchsten Klassenwahrscheinlichkeiten für eine Stichprobe betrachtet bzw. zusammengefasst werden. Momentan hat die grosse Anzahl an Kategorien den Effekt, dass die 'Top-Wahrscheinlichkeit' einer Klasse niedriger ausfällt. Bei der Betrachtung der Gesamtwahrscheinlichkeiten muss zudem die grosse Bandbreite an Wahrscheinlichkeiten der verschiedenen Kategorien berücksichtigt werden, da dadurch wenige schlecht klassierbare Stichproben eine sonst gute Gesamtgenauigkeit herabsetzen können (Abbildung 4.11).

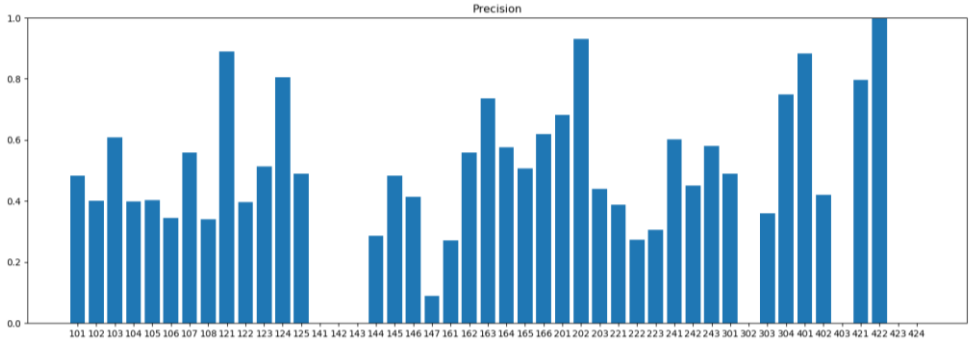


Abbildung 4.11: Beispiel der Verteilung der Klassen-Genauigkeiten aus der CNN-Klassifikation.

Die Genauigkeiten pro Kategorie sind somit beim grössten Teil der Klassen bedeutend höher als die Gesamtgenauigkeit. In der farblich dargestellten Konfusionsmatrix (Abbildung 4.12) fällt dieser Effekt, der sich auf viele Kategorien verteilende Gesamtgenauigkeiten, weniger stark ins Auge und die Ergebnisse weisen eine zu erwartende, gute Verteilung auf.

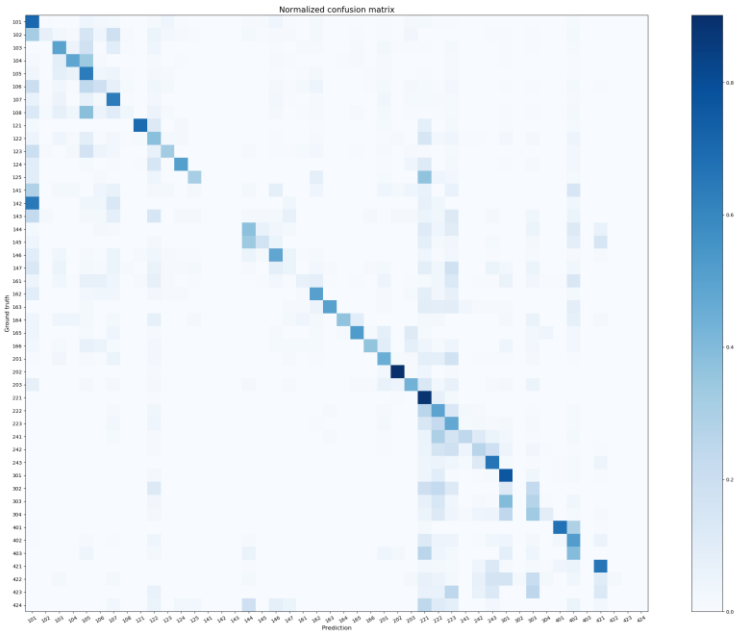


Abbildung 4.12. Beispiel einer normierten Konfusionsmatrix

Der grösste Teil der Trainingsdurchläufe des DL-Netzwerks wurde auf Daten der AREA3 durchgeführt. Das so trainierte Netzwerk zeigt eine gute Übertragbarkeit auf Area4-Daten (Gesamtgenauigkeit nur 1-2% niedriger). In der Umkehr zeigen auf AREA4 trainierte Netzwerke erwartungsgemäss eine weniger gute Übertragbarkeit auf die weniger hoch aufgelösten Daten von AREA3. Training auf weniger hoch aufgelösten Daten ist somit besser als

umgekehrt. Für AREA5 bedeutet dies, dass das Training auf den älteren Jahrgängen möglich ist, jedoch eher auf AREA4 oder auf AREA5 selbst empfohlen wird, da gewisse Effekte aufgrund der unterschiedlichen Auflösung nachweisbar sind. Die erhaltenen Gesamtgenauigkeiten über alle Kategorien belaufen sich am Ende des CNN Moduls zwischen 50% und 56% und dienen als Input für den nächsten Schritt im Random Forest.

4.3 Ergebnisse RF

4.3.1 Land Cover AREA3

Für das LC-Modell auf Basis von AREA3 konnte wie für alle anderen Modelle, in welchen AREA3-Klassifikationen vorgenommen wurden, nur Deep-Learning-Resultate für RGB-Kacheln (vgl. Kapitel 2.4) und die Satellitendaten (vgl. Kapitel 2.3.2) des AREA3-Zeitstandes verwendet werden. DL-Resultate (27x Klassenkonfidenzen + 27x TOP1 One-Hot-Encoded Klassenfaktoren + 1x TOP 1 Konfidenz = 55 Variablen). Trainiert wurden 362‘875 Samples. In der evaluierten Prädiktion befanden sich von ~3,7 Mio theoretisch untersuchbaren Samples 673‘991 Stichproben.

A3LC-Sat	11 - Befestigte Flächen	12 - Gebäude	13 - Treibhäuser	14 - Beetstrukturen	15 - Rasen	16 - Bäume auf künstl. Flächen	17 - Gemischte Kleinstrukturen	21 - Gras-, Krautvegetation	31 - Gebüsch	32 - Verbuschte Flächen	33 - Niederstammobst	34 - Reben	35 - Gärtnerische Dauerkulturen	41 - Geschlossene Baumbestände	42 - Waldecken	43 - Waldstreifen	44 - Aufgelöste Baumbestände	45 - Gebüschwaldbestände	46 - Lineare Baumbestände	47 - Baumgruppen	51 - Anstehender Fels	52 - Lockergestein	53 - Versteinte Flächen	61 - Wasser	62 - Gletscher, Firn	63 - Nassstandorte	64 - Schilfbestände	Prediction
11 - Befestigte Flächen	72.5%	27.5%	17.6%	39.5%	24.0%	21.3%	46.5%	1.0%	0.5%	0.7%	1.0%	1.2%	0.0%	0.1%	0.4%	0.6%	0.3%	0.0%	2.4%	1.3%	0.0%	0.7%	0.2%	0.5%	0.0%	0.1%	0.0%	24284
12 - Gebäude	5.2%	66.4%	0.0%	9.6%	4.0%	1.0%	2.5%	0.1%	0.0%	0.0%	0.0%	0.2%	1.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	6366
13 - Treibhäuser	0.0%	0.0%	29.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	14
14 - Beetstrukturen	0.0%	0.0%	0.0%	8.2%	0.1%	0.1%	0.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	49
15 - Rasen	2.4%	2.3%	0.0%	14.2%	40.1%	16.8%	29.0%	0.2%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.6%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	4139
16 - Bäume auf künstl. Flächen	0.9%	0.3%	0.0%	3.6%	3.5%	32.3%	7.9%	0.0%	0.0%	0.0%	0.3%	0.1%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	1.5%	0.9%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1300
17 - Gemischte Kleinstrukturen	0.0%	0.0%	0.0%	0.4%	0.2%	0.1%	0.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	28
21 - Gras-, Krautvegetation	12.1%	2.5%	50.0%	22.4%	25.1%	7.8%	9.5%	93.4%	23.4%	48.8%	30.1%	17.0%	65.7%	1.1%	34.9%	9.6%	3.7%	5.0%	26.0%	24.9%	1.9%	4.0%	23.3%	3.0%	0.1%	86.1%	59.3%	261736
31 - Gebüsch	0.2%	0.0%	0.0%	0.0%	0.1%	0.1%	0.1%	0.5%	32.3%	14.5%	0.0%	0.1%	7.1%	0.3%	1.1%	0.9%	2.8%	9.8%	1.3%	5.3%	0.4%	0.3%	1.9%	0.1%	0.0%	1.9%	2.3%	12590
32 - Verbuschte Flächen	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%	7.4%	15.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.3%	1.4%	0.0%	1.3%	0.1%	0.1%	1.6%	0.0%	0.0%	0.6%	0.0%	5300
33 - Niederstammobst	0.0%	0.0%	0.0%	0.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	57.5%	0.3%	7.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	233
34 - Reben	0.3%	0.0%	0.0%	0.0%	0.1%	0.1%	0.3%	0.0%	0.0%	0.1%	1.7%	80.1%	1.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1131
35 - Gärtnerische Dauerkulturen	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0
41 - Geschlossene Baumbestände	3.5%	0.3%	0.0%	0.0%	0.5%	3.8%	0.3%	1.0%	21.1%	4.8%	2.7%	0.5%	10.0%	87.2%	51.4%	70.5%	57.9%	24.0%	11.9%	16.7%	1.6%	0.7%	1.2%	1.6%	0.0%	4.5%	17.4%	180378
42 - Waldecken	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	2
43 - Waldstreifen	0.0%	0.0%	0.0%	0.0%	0.0%	0.5%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.7%	6.4%	0.1%	0.0%	2.5%	0.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	655
44 - Aufgelöste Baumbestände	0.2%	0.0%	0.0%	0.0%	0.0%	0.3%	0.2%	0.1%	1.4%	1.4%	0.0%	0.0%	0.0%	0.9%	1.8%	2.5%	25.1%	1.2%	0.6%	7.9%	0.2%	0.2%	0.5%	0.1%	0.0%	0.6%	0.0%	7723
45 - Gebüschwaldbestände	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	6.6%	2.9%	0.0%	0.0%	0.0%	0.4%	0.4%	1.0%	1.7%	54.9%	0.3%	1.5%	0.6%	0.2%	0.8%	0.1%	0.0%	0.1%	0.0%	9104
46 - Lineare Baumbestände	0.7%	0.1%	0.0%	0.0%	0.4%	5.2%	0.3%	0.2%	0.5%	0.2%	0.0%	0.3%	1.4%	0.1%	1.4%	4.3%	0.3%	0.0%	43.6%	2.0%	0.0%	0.0%	0.0%	0.4%	0.0%	0.6%	3.5%	3785
47 - Baumgruppen	0.7%	0.2%	0.0%	1.8%	1.5%	10.1%	2.1%	0.7%	1.9%	2.4%	6.5%	0.1%	5.7%	0.2%	7.4%	3.3%	6.7%	0.8%	8.6%	34.7%	0.3%	0.2%	0.8%	0.2%	0.0%	1.9%	1.2%	10470
51 - Anstehender Fels	0.1%	0.0%	2.9%	0.0%	0.0%	0.0%	0.0%	0.3%	1.3%	1.3%	0.0%	0.0%	0.0%	0.1%	0.4%	0.4%	0.6%	1.9%	0.3%	0.9%	71.2%	10.1%	10.4%	0.6%	8.9%	0.0%	1.2%	34678
52 - Lockergestein	0.6%	0.1%	0.0%	0.0%	0.1%	0.1%	0.0%	0.4%	0.5%	1.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.1%	0.2%	0.0%	0.3%	14.2%	75.1%	11.7%	1.6%	9.2%	0.4%	1.2%	46013
53 - Versteinte Flächen	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1.6%	2.9%	6.5%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.3%	0.6%	0.0%	1.0%	6.1%	6.3%	47.4%	0.4%	0.0%	0.5%	0.0%	22563
61 - Wasser	0.3%	0.1%	0.0%	0.0%	0.2%	0.5%	0.0%	0.0%	0.1%	0.1%	0.0%	0.1%	0.0%	0.0%	0.0%	0.3%	0.1%	0.1%	0.8%	0.1%	0.2%	0.4%	0.1%	90.9%	0.1%	0.3%	11.6%	24615
62 - Gletscher, Firn	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	3.1%	1.6%	0.1%	0.2%	81.7%	0.0%	0.0%	16754
63 - Nassstandorte	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	2.2%	1.2%	78
64 - Schilfbestände	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1.2%	3
Reference	22876	7111	34	281	5413	1533	906	242317	17523	13939	292	1180	70	158574	284	4602	14946	10033	4737	15177	33778	46269	24881	26400	18226	2523	86	673991

Abbildung 4.13. Confusion Matrix (normalisiert) für Land Cover AREA3 - Im Training: CNN-Klassifikationsergebnisse (RGB) und Satelliten-Zeitreihen.

Genauigkeit A3LC-Sat (F1)	Sensitivity	Specificity	Pos. Pred.	Neg. Pred.	Precision	Recall	F1	Prevalence	Det. Rate	Det. Preval.	Bal. Accuracy	Samples	Correct Samples
11 - Befestigte Flächen	72.5%	98.8%	68.3%	99.0%	68.3%	72.5%	70.3%	3.4%	2.5%	3.6%	71.3%	127'092	89'374
12 - Gebäude	66.4%	99.8%	74.2%	99.6%	74.2%	66.4%	70.1%	1.1%	0.7%	0.9%	66.1%	39'506	27'678
13 - Treibhäuser	29.4%	100.0%	71.4%	100.0%	71.4%	29.4%	41.7%	0.0%	0.0%	0.0%	29.4%	189	79
14 - Beetstrukturen	8.2%	100.0%	46.9%	100.0%	46.9%	8.2%	13.9%	0.0%	0.0%	0.0%	8.2%	1'561	218
15 - Rasen	40.1%	99.7%	52.5%	99.5%	52.5%	40.1%	45.5%	0.8%	0.3%	0.6%	39.9%	30'073	13'683
16 - Bäume auf künstl. Flächen	32.3%	99.9%	38.1%	99.8%	38.1%	32.3%	34.9%	0.2%	0.1%	0.2%	32.2%	8'517	2'976
17 - Gemischte Kleinstrukturen	0.8%	100.0%	25.0%	99.9%	25.0%	0.8%	1.5%	0.1%	0.0%	0.0%	0.8%	5'033	75
21 - Gras-, Krautvegetation	93.4%	91.8%	86.4%	96.1%	86.4%	93.4%	89.8%	36.0%	33.6%	38.8%	85.1%	1'346'236	1'208'435
31 - Gebüsch	32.3%	98.9%	44.9%	98.2%	44.9%	32.3%	37.6%	2.6%	0.8%	1.9%	31.2%	97'352	36'590
32 - Verbuschte Flächen	15.3%	99.5%	40.3%	98.2%	40.3%	15.3%	22.2%	2.1%	0.3%	0.8%	14.9%	77'441	17'212
33 - Niederstammobst	57.5%	100.0%	72.1%	100.0%	72.1%	57.5%	64.0%	0.0%	0.0%	0.0%	57.5%	1'622	1'038
34 - Reben	80.1%	100.0%	83.6%	100.0%	83.6%	80.1%	81.8%	0.2%	0.1%	0.2%	80.1%	6'556	5'361
35 - Gärtnerische Dauerkulturen	0.0%	100.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	389	-
41 - Geschlossene Baumbestände	96.7%	94.8%	85.0%	99.0%	85.0%	96.7%	90.5%	23.5%	22.8%	26.8%	91.5%	880'987	797'455
42 - Waldecken	0.4%	100.0%	50.0%	100.0%	50.0%	0.4%	0.7%	0.0%	0.0%	0.0%	0.4%	1'578	11
43 - Waldstreifen	6.4%	99.9%	45.2%	99.4%	45.2%	6.4%	11.3%	0.7%	0.0%	0.1%	6.4%	25'567	2'879
44 - Aufgelöste Baumbestände	25.1%	99.4%	48.6%	98.3%	48.6%	25.1%	33.1%	2.2%	0.6%	1.1%	24.5%	83'035	27'479
45 - Gebüschwaldbestände	54.9%	99.5%	60.5%	99.3%	60.5%	54.9%	57.6%	1.5%	0.8%	1.4%	54.4%	55'740	32'092
46 - Lineare Baumbestände	43.6%	99.7%	54.5%	99.6%	54.5%	43.6%	48.4%	0.7%	0.3%	0.6%	43.3%	26'317	12'748
47 - Baumgruppen	34.7%	99.2%	50.4%	98.5%	50.4%	34.7%	41.1%	2.3%	0.8%	1.6%	34.0%	84'319	34'672
51 - Anstehender Fels	71.2%	98.3%	69.3%	98.5%	69.3%	71.2%	70.3%	5.0%	3.6%	5.1%	69.5%	187'660	131'836
52 - Lockergestein	75.1%	98.2%	75.5%	98.2%	75.5%	75.1%	75.3%	6.9%	5.2%	6.8%	73.3%	257'056	193'529
53 - Versteinte Flächen	47.4%	98.3%	52.3%	98.0%	52.3%	47.4%	49.8%	3.7%	1.8%	3.3%	45.8%	138'231	68'772
61 - Wasser	90.9%	99.9%	97.5%	99.6%	97.5%	90.9%	94.1%	3.9%	3.9%	3.7%	90.8%	146'670	138'013
62 - Gletscher, Firn	81.7%	99.7%	88.9%	99.5%	88.9%	81.7%	85.2%	2.7%	2.2%	2.5%	81.4%	101'258	86'240
63 - Nassstandorte	2.2%	100.0%	70.5%	99.6%	70.5%	2.2%	4.2%	0.4%	0.0%	0.0%	2.2%	14'017	593
64 - Schilfbestände	1.2%	100.0%	33.3%	100.0%	33.3%	1.2%	2.2%	0.0%	0.0%	0.0%	1.2%	478	11
												3'744'480	2'929'049
													78.22%

Abbildung 4.14. Genauigkeitsmetriken für Land Cover AREA3 - Im Training: CNN-Klassifikationsergebnisse (RGB) und Satelliten-Zeitreihen; rechts dargestellt die bei derzeitiger Datenlage maximal vorhersagbare Anzahl Samples mit diesem Modell.

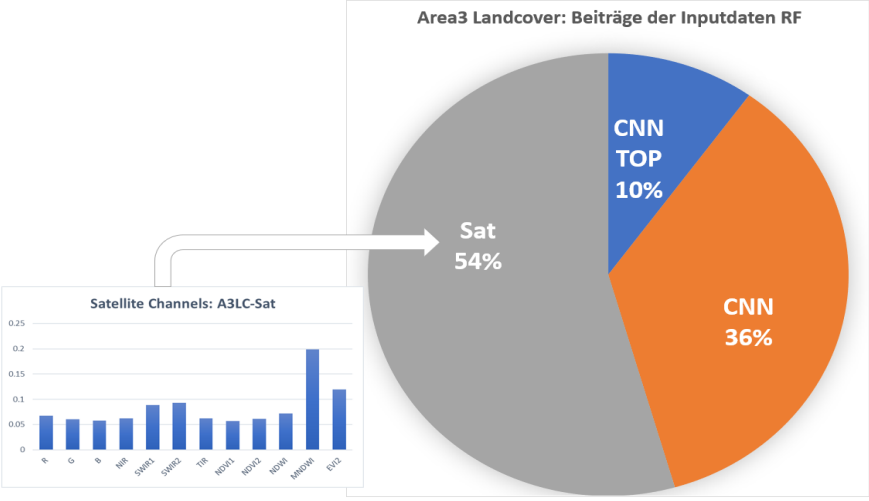


Abbildung 4.15. Variablengewichte für Land Cover AREA3 - Im Training: CNN-Klassifikationsergebnisse (RGB) und Satelliten-Zeitreihen.

Besonders gut funktionierte die Klassifikation bei Klassen mit ausreichend grosser Stichprobengrösse, die homogen in ihrer RGB- und Zeit-Struktur sind. Werte über 85% Genauigkeit (F1) konnten für die Klassen „21 – Gras-/Krautvegetation“, „41 – Geschlossene Baumbestände“, „61 – Wasser“ und „62 – Gletscher/Firn“ erreicht werden. Besonders schlechte Werte traten einerseits in heterogenen Klassen mit niedriger Prävalenz (Besetzung) auf (z.B. „17 – gemischte Kleinstrukturen“), andererseits konnten auch Klassen mit niedriger Prävalenz, die eine hohe visuelle Ähnlichkeit zu sehr grossen Klassen aufweisen nicht erkannt werden: So wurden z.B. die sehr kleinen Klassen „42 – Waldecken“ und „43 – Waldstreifen“ nahezu ausschliesslich der Klasse „41 – Geschlossene Baumbestände“ zugewiesen. Insbesondere der starke Fokus der Arealstatistik auf das Zentrum der Kacheln kann zu Grenzfällen führen, in welchen ein Grossteil der Kachel einer anderen Klasse zugerechnet wird. Hier zeigt sich unter Umständen eine Limitation der starken Gewichtung von Satellitenzeitreihen. Durch die niedrige Auflösung der verwendeten Landsat-Daten ist eine Differenzierung zwischen Waldrandbereichen und Wäldern praktisch nicht möglich. Auch die im Verhältnis niedrige Auflösung der Inputkacheln für das CNN mit 50cm GSD scheint das Erkennen der Randbereiche zu erschweren, sodass kleine Klassen ins Hintertreffen geraten. Trotz dieser Limitationen ist eine Overall Prediction Accuracy von 78,22% bereits ein ausserordentlich guter Wert für das Resultat einer so feinen Land Cover Klassifikation.

4.3.2 Land Cover AREA4

A4LC-Sat	11 - Befestigte Flächen	12 - Gebäude	13 - Treibhäuser	14 - Beetstrukturen	15 - Rasen	16 - Bäume auf künstl. Flächen	17 - Gemischte Kleinstrukturen	21 - Gras, Krautvegetation	31 - Gebüsch	32 - Verbuschte Flächen	33 - Niederstammobst	34 - Reben	35 - Gärtnerische Dauerkulturen	41 - Geschlossene Baumbestände	42 - Waldecken	43 - Waldstreifen	44 - Aufgelöste Baumbestände	45 - Gebüschwaldbestände	46 - Lineare Baumbestände	47 - Baumgruppen	51 - Anstehender Fels	52 - Lockergestein	53 - Versteinte Flächen	61 - Wasser	62 - Gletscher, Firn	63 - Nassstandorte	64 - Schilfbestände	Prediction
11 - Befestigte Flächen	78.6%	32.6%	39.1%	56.7%	26.1%	29.9%	60.7%	1.0%	1.2%	1.5%	2.9%	2.9%	5.0%	0.1%	1.7%	0.8%	0.4%	0.0%	4.1%	2.1%	0.1%	2.0%	0.5%	0.5%	0.0%	0.5%	1.5%	35266
12 - Gebäude	3.1%	63.8%	0.0%	0.9%	1.7%	0.4%	2.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	5694
13 - Treibhäuser	0.0%	0.0%	45.7%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	23
14 - Beetstrukturen	0.1%	0.1%	0.0%	15.7%	0.1%	0.1%	0.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	95
15 - Rasen	2.0%	1.4%	0.0%	7.2%	46.4%	12.7%	21.5%	0.2%	0.0%	0.1%	0.0%	0.0%	0.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.4%	0.0%	0.1%	0.0%	0.1%	0.0%	0.0%	0.0%	4931
16 - Bäume auf künstl. Flächen	0.7%	0.2%	0.0%	0.9%	3.1%	35.2%	6.2%	0.0%	0.1%	0.1%	0.0%	0.0%	0.8%	0.0%	0.7%	0.3%	0.0%	0.0%	1.5%	1.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	1577
17 - Gemischte Kleinstrukturen	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0
21 - Gras-, Krautvegetation	10.4%	1.5%	15.2%	16.6%	20.5%	6.3%	6.5%	80.9%	28.1%	58.4%	27.3%	13.2%	55.8%	0.9%	39.2%	7.2%	4.0%	9.0%	23.3%	32.1%	2.4%	4.3%	28.0%	1.8%	0.1%	84.5%	61.7%	307033
31 - Gebüsch	0.1%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.1%	26.3%	10.0%	0.4%	0.0%	10.0%	0.2%	0.0%	0.2%	2.2%	7.3%	0.7%	2.5%	0.2%	0.2%	1.2%	0.1%	0.0%	0.9%	3.8%	5039
32 - Verbuschte Flächen	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	5.8%	10.1%	0.0%	0.0%	0.0%	0.0%	0.3%	0.1%	0.2%	1.4%	0.1%	0.8%	0.1%	0.1%	0.9%	0.0%	0.0%	0.5%	0.0%	2005
33 - Niederstammobst	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	63.2%	0.2%	0.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	197
34 - Reben	0.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.0%	0.1%	0.0%	0.4%	83.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	638
35 - Gärtnerische Dauerkulturen	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	13.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	29
41 - Geschlossene Baumbestände	2.6%	0.1%	0.0%	0.9%	0.2%	1.7%	0.3%	0.6%	21.1%	3.5%	2.9%	0.2%	5.0%	97.4%	44.6%	74.5%	59.0%	21.8%	9.5%	11.0%	1.1%	0.7%	0.8%	1.0%	0.0%	2.5%	6.8%	193288
42 - Waldecken	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0
43 - Waldstreifen	0.0%	0.0%	0.0%	0.0%	0.0%	0.7%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.1%	1.7%	5.5%	0.1%	0.1%	2.3%	0.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	615
44 - Aufgelöste Baumbestände	0.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	3.9%	2.8%	0.0%	0.0%	0.8%	0.9%	0.3%	2.7%	26.3%	2.8%	0.7%	7.5%	0.3%	0.4%	0.8%	0.2%	0.0%	1.4%	0.0%	7952
45 - Gebüschwaldbestände	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	5.4%	1.6%	0.0%	0.0%	0.0%	0.1%	0.0%	0.2%	0.7%	53.0%	0.1%	0.3%	0.2%	0.0%	0.3%	0.0%	0.0%	0.1%	0.0%	3312
46 - Lineare Baumbestände	0.5%	0.1%	0.0%	0.0%	0.2%	3.0%	0.5%	0.2%	0.8%	0.2%	0.4%	0.0%	1.7%	0.1%	3.0%	4.7%	0.3%	0.1%	45.6%	1.9%	0.0%	0.0%	0.0%	0.4%	0.0%	1.3%	3.0%	4215
47 - Baumgruppen	0.6%	0.1%	0.0%	0.9%	1.4%	9.7%	1.1%	0.7%	1.9%	2.5%	2.5%	0.0%	4.2%	0.1%	8.4%	3.4%	5.8%	0.9%	11.0%	38.3%	0.2%	0.2%	0.7%	0.2%	0.0%	3.5%	2.3%	10198
51 - Anstehender Fels	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	1.8%	1.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.4%	2.4%	0.1%	0.8%	77.1%	12.6%	11.3%	0.3%	9.1%	0.0%	0.0%	26181
52 - Lockergestein	0.6%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.5%	1.1%	0.0%	0.2%	0.0%	0.0%	0.0%	0.0%	0.3%	0.5%	0.1%	0.3%	11.2%	71.7%	8.8%	1.1%	8.2%	0.3%	0.0%	23572
53 - Versteinte Flächen	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.7%	2.6%	6.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.6%	0.0%	0.5%	5.5%	5.8%	46.6%	0.2%	0.0%	0.3%	0.0%	12257
61 - Wasser	0.2%	0.1%	0.0%	0.0%	0.3%	0.2%	0.0%	0.0%	0.2%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.1%	0.0%	0.8%	0.2%	0.0%	0.3%	0.1%	93.8%	0.0%	0.3%	15.8%	32034
62 - Gletscher, Firn	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1.7%	1.4%	0.0%	0.1%	82.6%	0.0%	0.0%	10671
63 - Nassstandorte	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	3.8%	3.8%	137
64 - Schilfbestände	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1.5%	2
Reference	32144	7040	46	319	6679	2233	629	289834	8456	7784	242	583	120	177242	296	4471	14363	4108	5509	13510	25154	24228	13810	33647	11939	2392	133	686961

Abbildung 4.16. Confusion Matrix (normalisiert) für Land Cover AREA4 - Im Training: CNN-Klassifikationsergebnisse (RGB) und Satelliten-Zeitreihen.

Genauigkeit A4LC-Sat (F1)	Sensitivity	Specificity	Pos. Pred.	Neg. Pred.	Precision	Recall	F1	Prevalence	Det. Rate	Det. Preval.	Bal. Accuracy	Samples	Correct Samples
11 - Befestigte Flächen	78.6%	98.5%	71.7%	98.9%	71.7%	78.6%	75.0%	4.7%	3.7%	5.1%	77.1%	100'450	75'329
12 - Gebäude	63.8%	99.8%	78.9%	99.6%	78.9%	63.8%	70.6%	1.0%	0.7%	0.8%	63.6%	22'000	15'521
13 - Treibhäuser	45.7%	100.0%	91.3%	100.0%	91.3%	45.7%	60.9%	0.0%	0.0%	0.0%	45.7%	144	87
14 - Beetstrukturen	15.7%	100.0%	52.6%	100.0%	52.6%	15.7%	24.2%	0.0%	0.0%	0.0%	15.7%	997	241
15 - Rasen	46.4%	99.7%	62.8%	99.5%	62.8%	46.4%	53.4%	1.0%	0.5%	0.7%	46.1%	20'872	11'135
16 - Bäume auf künstl. Flächen	35.2%	99.9%	49.8%	99.8%	49.8%	35.2%	41.3%	0.3%	0.1%	0.2%	35.1%	6'978	2'879
17 - Gemischte Kleinstrukturen	0.0%	100.0%	0.0%	99.9%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	1'966	-
21 - Gras-, Krautvegetation	95.9%	92.7%	90.5%	96.9%	90.5%	95.9%	93.1%	42.2%	40.5%	44.7%	88.6%	905'883	843'748
31 - Gebüsch	26.3%	99.6%	44.2%	99.1%	44.2%	26.3%	33.0%	1.2%	0.3%	0.7%	25.9%	26'425	8'718
32 - Verbuschte Flächen	10.1%	99.8%	39.2%	99.0%	39.2%	10.1%	16.0%	1.1%	0.1%	0.3%	9.9%	24'325	3'901
33 - Niederstammobst	63.2%	100.0%	77.7%	100.0%	77.7%	63.2%	69.7%	0.0%	0.0%	0.0%	63.2%	756	527
34 - Reben	83.4%	100.0%	76.2%	100.0%	76.2%	83.4%	79.6%	0.1%	0.1%	0.1%	83.3%	1'822	1'450
35 - Gärtnerische Dauerkulturen	13.3%	100.0%	55.2%	100.0%	55.2%	13.3%	21.5%	0.0%	0.0%	0.0%	13.3%	375	81
41 - Geschlossene Baumbestände	97.4%	96.0%	89.3%	99.1%	89.3%	97.4%	93.2%	25.8%	25.1%	28.1%	93.4%	553'879	516'191
42 - Waldecken	0.0%	100.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	925	-
43 - Waldstreifen	5.5%	99.9%	39.7%	99.4%	39.7%	5.5%	9.6%	0.7%	0.0%	0.1%	5.4%	13'972	1'341
44 - Aufgelöste Baumbestände	26.3%	99.4%	47.5%	98.4%	47.5%	26.3%	33.9%	2.1%	0.6%	1.2%	25.7%	44'884	15'210
45 - Gebüschwaldbestände	53.0%	99.8%	65.8%	99.7%	65.8%	53.0%	58.7%	0.6%	0.3%	0.5%	52.9%	12'837	7'536
46 - Lineare Baumbestände	45.6%	99.7%	59.5%	99.6%	59.5%	45.6%	51.6%	0.8%	0.4%	0.6%	45.3%	17'216	8'887
47 - Baumgruppen	38.3%	99.3%	50.8%	98.8%	50.8%	38.3%	43.7%	2.0%	0.8%	1.5%	37.6%	42'219	18'445
51 - Anstehender Fels	77.1%	99.0%	74.1%	99.1%	74.1%	77.1%	75.6%	3.7%	2.8%	3.8%	76.1%	78'606	59'421
52 - Lockergestein	71.7%	99.1%	73.7%	99.0%	73.7%	71.7%	72.6%	3.5%	2.5%	3.4%	70.7%	75'712	54'997
53 - Versteinte Flächen	46.6%	99.1%	52.5%	98.9%	52.5%	46.6%	49.4%	2.0%	0.9%	1.8%	45.7%	43'156	21'304
61 - Wasser	93.8%	99.9%	98.6%	99.7%	98.6%	93.8%	96.1%	4.9%	4.6%	4.7%	93.8%	105'146	101'082
62 - Gletscher, Firn	82.6%	99.9%	92.4%	99.7%	92.4%	82.6%	87.3%	1.7%	1.4%	1.6%	82.5%	37'309	32'557
63 - Nassstandorte	3.8%	100.0%	65.7%	99.7%	65.7%	3.8%	7.1%	0.3%	0.0%	0.0%	3.8%	7'475	532
64 - Schilfbestände	1.5%	100.0%	100.0%	100.0%	100.0%	1.5%	3.0%	0.0%	0.0%	0.0%	1.5%	416	12
												2'146'743	1'801'134
													83.90%

Abbildung 4.17. Genauigkeitsmetriken für Land Cover AREA4 - Im Training: CNN-Klassifikationsergebnisse (RGB) und Satelliten-Zeitreihen; rechts dargestellt die bei derzeitiger Datenlage maximal vorhersagbare Anzahl Samples mit diesem Modell.

Das Land-Cover-Modell für AREA4 wurde zunächst ohne Zuhilfenahme des FCIR-Datensatzes oder der Zusatzdaten durchgeführt, um Vergleichbarkeit mit dem AREA3-Modell zu gewährleisten und gleichzeitig als Inputfaktor für die Prädikationsbasierte Change Detection Verwendung zu finden. Im Training befanden sich 369‘902 Samples, in der Prädiktion 686‘961 Samples.

Das bei AREA3 gewonnene Bild setzt sich hier fort, dass vor allem grössere, homogene Klassen mit einer hohen Genauigkeit vorhergesagt werden können (weiterhin 21, 41, 61, 62 > 85% F1), während heterogene Klassen mit sehr niedriger Prävalenz schlecht abschneiden. Die Overall-Accuracy konnte jedoch auf 83,90% gesteigert werden. Da bisher jedoch zum Zeitpunkt des Kachelexports für lediglich 57% der Stichproben eine Kombination aus vorhandenem SWISSIMAGE und AREA4-BFS-Klassifikation vorlag, reduzierte sich die Anzahl maximal integrierbarer Samples.

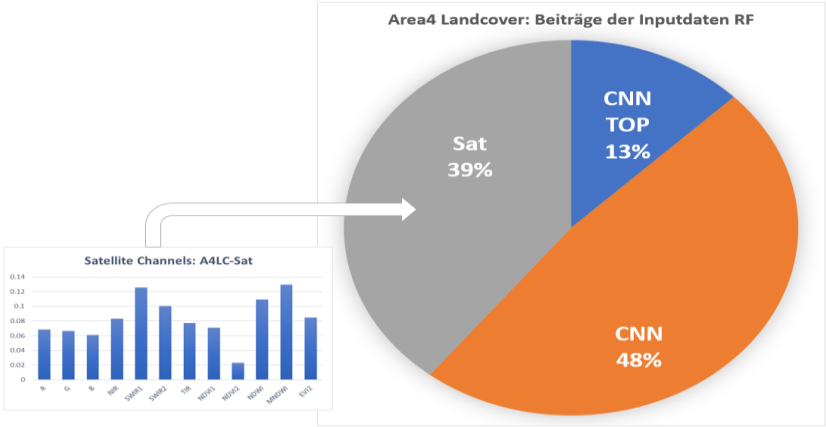
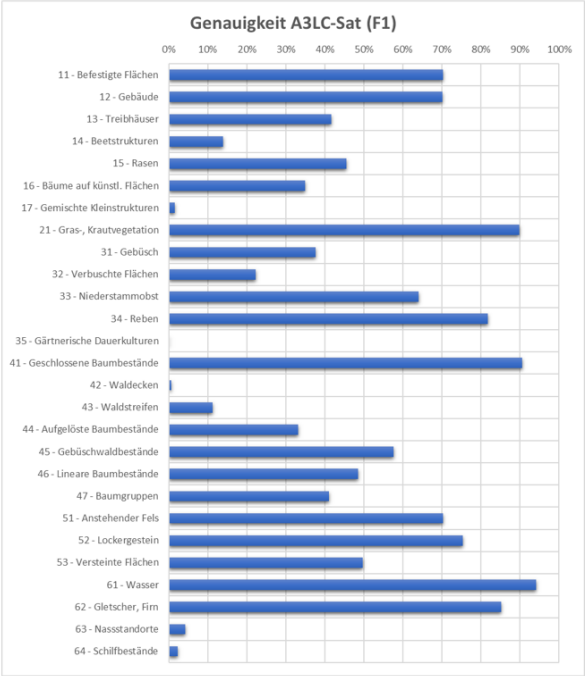


Abbildung 4.18. Variablengewichte für Land Cover AREA4 - Im Training: CNN-Klassifikationsergebnisse (RGB) und Satelliten-Zeitreihen.

Gegenüber AREA3-LC wuchs in diesem Modell insbesondere der Anteil der Deep-Learning-Klassifikationen stark an. Zu einem grossen Teil kann dies vermutlich auf die höhere Ground-Sampling-Auflösung der Kacheldaten und somit einer höheren Präzision der DL-Klassifikation zurückgeführt werden. Bei den Kanälen der Satellitenzeitreihen-Klassifikation fällt die Diversifikation der Variablengewichte auf. Im AREA3-LC-Modell war der Wasserindex MNDWI wesentlich wichtiger als alle anderen Kanäle; im AREA4-LC-Modell geht die Vorrangrolle zugunsten SWIR und EVI2 etwas zurück.



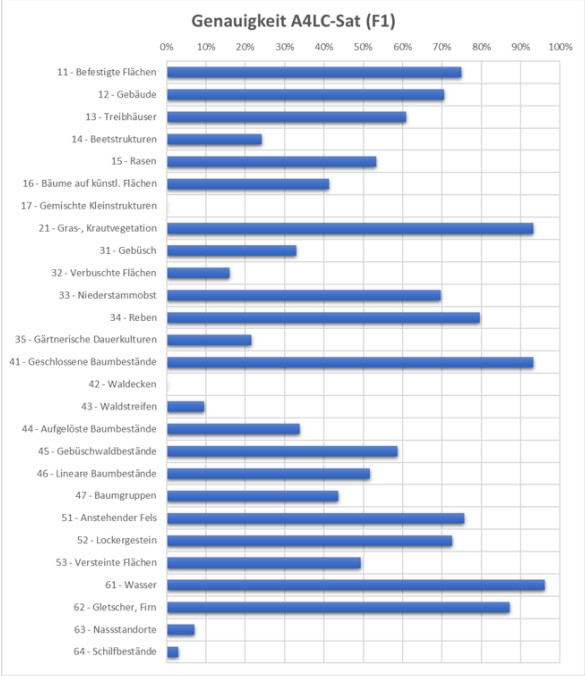


Abbildung 4.19. Vergleich der F1-Score-Genauigkeiten pro Klasse zwischen den RF-Ergebnissen (Land Cover) von AREA3 und AREA4.

Im Vergleich zwischen den LC-Modellen für AREA3 und 4 fällt auf, dass einige Klassen (vermutlich aufgrund der höheren Auflösung der CNN-Inputkacheln) wesentlich besser bei AREA4, als bei AREA3 abschneiden. Substanzielle Rückgänge in der Genauigkeit sind nicht ersichtlich. Die stärksten Steigerungen finden sich bei den folgenden Klassen:

Tabelle 4.14. Klassen mit den stärksten Steigerungen zwischen LC-AREA3 und LC-AREA4.

Klasse	Bezeichnung	Veränderungen gegenüber AREA3 >5%
11	Befestigte Flächen	+5%
13	Treibhäuser	+20%
14	Beetstrukturen	+10%
15	Rasen	+7%
16	Bäume auf künstlichen Flächen	+6%
31	Gebüsch	+5%
51	Anstehender Fels	+6%

Es sind relativ heterogene und trotzdem noch mässig stark besetzte Klassen, die von der Steigerung der Kachelbodenauflösung beim CNN profitieren zu scheinen.

4.3.3 Land Cover AREA4 mit Zusatzdaten

Die Verwendung der Zusatzdaten im voll ausgebauten AREA4-LC-Modell brachte nochmals eine relativ deutliche Verbesserung der Performanz auf eine Gesamtgenauigkeit von 88,51%. Zwar erreicht lediglich zusätzlich Klasse 34 – Reben nun zusätzlich zu den Klassen aus den AREA3/4-LC-Modellen eine Genauigkeit >85% (F1), aber viele kleinere Klassen mit einer Prävalenz unter 5% schneiden nun deutlich besser in einem Genauigkeitsbereich von >70% ab. Gegenüber dem vorherigen AREA4-LC-Modell nimmt jedoch die Anzahl zum derzeitigen Stand vorhersagbarer Punkte jedoch abermals ab, da FCIR-Inputdaten für das CNN lediglich für einen Anteil von 36% aller Stichproben vorhanden sind. Auch die Verknüpfung mit den Zusatzdaten verringert die zur Verfügung stehende Stichprobengrösse nochmals (wenn auch in wesentlich geringerem Rahmen), da die Datensätze noch einzelne Leerstellen enthalten. Im Training befanden sich 251‘550 Samples, in der Prädiktion 467‘166 Samples.

A4LC-Full	11 - Befestigte Flächen	12 - Gebäude	13 - Treibhäuser	14 - Beetstrukturen	15 - Rasen	16 - Bäume auf künstl. Flächen	17 - Gemischte Kleinstrukturen	21 - Gras-, Krautvegetation	31 - Gebüsch	32 - Verbuschte Flächen	33 - Niederstammobst	34 - Reben	35 - Gärtnerische Dauerkulturen	41 - Geschlossene Baumbestände	42 - Waldecken	43 - Waldstreifen	44 - Aufgelöste Baumbestände	45 - Gebüschwaldbestände	46 - Lineare Baumbestände	47 - Baumgruppen	51 - Anstehender Fels	52 - Lockergestein	53 - Versteinte Flächen	61 - Wasser	62 - Gletscher, Firn	63 - Nassstandorte	64 - Schilfbestände	Prediction
11 - Befestigte Flächen	84.8%	21.5%	18.2%	39.3%	22.5%	29.2%	51.9%	0.7%	0.8%	1.0%	1.0%	0.9%	2.7%	0.1%	1.9%	0.5%	0.3%	0.0%	2.8%	1.5%	0.0%	2.7%	0.8%	0.2%	0.0%	0.8%	1.1%	23135
12 - Gebäude	1.6%	76.0%	9.1%	0.9%	0.9%	0.2%	1.5%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	3259
13 - Treibhäuser	0.0%	0.0%	48.5%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	23
14 - Beetstrukturen	0.0%	0.0%	0.0%	20.5%	0.1%	0.0%	0.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	71
15 - Rasen	1.7%	0.7%	0.0%	13.1%	56.3%	19.4%	31.8%	0.2%	0.1%	0.2%	0.0%	0.2%	1.4%	0.0%	0.5%	0.0%	0.0%	0.0%	0.2%	0.7%	0.0%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%	3226
16 - Bäume auf künstl. Flächen	0.2%	0.1%	0.0%	0.0%	0.7%	21.5%	2.9%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%	0.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	258
17 - Gemischte Kleinstrukturen	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	4
21 - Gras-, Krautvegetation	7.1%	1.2%	24.2%	23.1%	17.2%	6.4%	7.7%	97.1%	24.3%	68.2%	26.7%	3.6%	69.9%	0.6%	22.2%	5.2%	3.5%	11.4%	25.1%	33.3%	2.2%	5.4%	32.4%	1.0%	0.0%	88.6%	24.1%	230320
31 - Gebüsch	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	29.8%	8.8%	0.0%	0.0%	2.7%	0.1%	0.0%	0.2%	1.8%	20.7%	0.2%	1.8%	0.1%	0.3%	1.0%	0.0%	0.0%	0.3%	2.3%	2251
32 - Verbuschte Flächen	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	2.4%	4.6%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	2.8%	0.0%	0.4%	0.0%	0.1%	0.7%	0.0%	0.0%	0.5%	0.0%	389
33 - Niederstammobst	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	68.8%	0.0%	1.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	174
34 - Reben	0.1%	0.0%	0.0%	0.4%	0.1%	0.0%	0.2%	0.0%	0.1%	0.0%	1.0%	95.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	608
35 - Gärtnerische Dauerkulturen	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	16.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	14
41 - Geschlossene Baumbestände	2.8%	0.1%	0.0%	0.0%	0.2%	1.1%	0.2%	0.5%	29.1%	4.4%	0.0%	0.0%	1.4%	94.0%	60.6%	81.8%	57.8%	39.4%	12.6%	11.0%	1.2%	1.3%	1.0%	1.2%	0.0%	2.0%	8.0%	145253
42 - Waldecken	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0
43 - Waldstreifen	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1.4%	0.0%	0.5%	2.1%	0.0%	0.0%	1.0%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	121
44 - Aufgelöste Baumbestände	0.1%	0.0%	0.0%	0.0%	0.0%	0.2%	0.0%	0.1%	6.2%	3.1%	0.0%	0.0%	0.0%	0.8%	0.0%	2.4%	28.9%	8.1%	0.7%	7.9%	0.3%	0.4%	1.1%	0.1%	0.0%	2.0%	0.0%	5627
45 - Gebüschwaldbestände	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	11.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	46
46 - Lineare Baumbestände	0.3%	0.0%	0.0%	0.4%	0.4%	5.8%	0.7%	0.1%	0.9%	0.3%	0.0%	0.2%	0.0%	0.1%	4.6%	4.5%	0.3%	0.0%	44.6%	2.0%	0.0%	0.0%	0.1%	0.3%	0.0%	1.5%	4.6%	2528
47 - Baumgruppen	0.5%	0.2%	0.0%	2.2%	1.6%	15.5%	2.0%	0.6%	2.5%	3.7%	2.5%	0.0%	1.4%	0.1%	8.8%	2.8%	6.2%	2.4%	10.8%	39.2%	0.2%	0.3%	1.1%	0.1%	0.0%	2.3%	0.0%	6962
51 - Anstehender Fels	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	1.3%	1.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.4%	2.4%	0.0%	0.9%	82.1%	7.9%	10.0%	0.1%	4.8%	0.0%	0.0%	9552
52 - Lockergestein	0.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.6%	1.0%	0.0%	0.2%	0.0%	0.0%	0.0%	0.0%	0.2%	0.4%	0.1%	0.3%	8.0%	73.9%	9.8%	0.7%	3.3%	0.0%	0.0%	9311
53 - Versteinte Flächen	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.4%	1.3%	3.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.4%	0.0%	0.4%	3.9%	4.8%	41.7%	0.1%	0.0%	0.0%	0.0%	4145
61 - Wasser	0.2%	0.0%	0.0%	0.0%	0.0%	0.6%	0.2%	0.0%	0.5%	0.3%	0.0%	0.0%	0.0%	0.0%	0.9%	0.3%	0.1%	0.0%	1.4%	0.1%	0.1%	1.5%	0.2%	96.1%	0.1%	1.0%	29.9%	16440
62 - Gletscher, Firn	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1.8%	1.2%	0.0%	0.0%	91.9%	0.0%	0.0%	3404
63 - Nassstandorte	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.3%	0.0%	11
64 - Schilfbestände	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.8%	29.9%	34
Reference	**21923**	**3741**	**33**	**229**	**3670**	**530**	**453**	**222954**	**3813**	**2701**	**202**	**560**	**73**	**134922**	**216**	**2486**	**10070**	**246**	**3531**	**9627**	**9303**	**10035**	**5333**	**16649**	**3384**	**395**	**87**	**467166**

Abbildung 4.20. Confusion Matrix (normalisiert) für Land Cover AREA4 - Im Training: CNN-Klassifikationsergebnisse (RGB und FCIR getrennt), Satelliten-Zeitreihen und alle verfügbaren Zusatzdaten (Höhenmodell, Vegetationshöhen, Amtl. Vermessung).

Genauigkeit A4LC-Full (F1)	Sensitivity	Specificity	Pos. Pred.	Neg. Pred.	Precision	Recall	F1	Prevalence	Det. Rate	Det. Preval.	Bal. Accuracy	Samples	Correct Samples
11 - Befestigte Flächen	84.8%	99.0%	80.4%	99.3%	80.4%	84.8%	82.6%	4.7%	4.0%	5.0%	83.8%	62'637	51'713
12 - Gebäude	76.0%	99.9%	87.2%	99.8%	87.2%	76.0%	81.2%	0.8%	0.6%	0.7%	75.9%	10'689	8'679
13 - Treibhäuser	48.5%	100.0%	69.6%	100.0%	69.6%	48.5%	57.1%	0.0%	0.0%	0.0%	48.5%	94	54
14 - Beetstrukturen	20.5%	100.0%	66.2%	100.0%	66.2%	20.5%	31.3%	0.0%	0.0%	0.0%	20.5%	654	205
15 - Rasen	56.3%	99.8%	64.1%	99.7%	64.1%	56.3%	60.0%	0.8%	0.4%	0.7%	56.1%	10'486	6'289
16 - Bäume auf künstl. Flächen	21.5%	100.0%	44.2%	99.9%	44.2%	21.5%	28.9%	0.1%	0.0%	0.1%	21.5%	1'514	438
17 - Gemischte Kleinstrukturen	0.4%	100.0%	50.0%	99.9%	50.0%	0.4%	0.9%	0.1%	0.0%	0.0%	0.4%	1'294	11
21 - Gras-, Krautvegetation	97.1%	94.4%	94.0%	97.3%	94.0%	97.1%	95.6%	47.7%	46.4%	49.3%	91.5%	637'010	608'768
31 - Gebüsch	29.8%	99.8%	50.4%	99.4%	50.4%	29.8%	37.4%	0.8%	0.2%	0.5%	29.5%	10'894	4'078
32 - Verbuschte Flächen	4.6%	99.9%	32.1%	99.4%	32.1%	4.6%	8.1%	0.6%	0.0%	0.1%	4.6%	7'717	624
33 - Niederstammobst	68.8%	100.0%	79.9%	100.0%	79.9%	68.8%	73.9%	0.0%	0.0%	0.0%	68.8%	577	427
34 - Reben	95.0%	100.0%	87.5%	100.0%	87.5%	95.0%	91.1%	0.1%	0.1%	0.1%	95.0%	1'600	1'458
35 - Gärtnerische Dauerkulturen	16.4%	100.0%	85.7%	100.0%	85.7%	16.4%	27.6%	0.0%	0.0%	0.0%	16.4%	209	58
41 - Geschlossene Baumbestände	98.0%	96.1%	91.1%	99.2%	91.1%	98.0%	94.4%	28.9%	28.3%	31.1%	94.1%	385'490	364'017
42 - Waldecken	0.0%	100.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	617	-
43 - Waldstreifen	2.1%	100.0%	43.8%	99.5%	43.8%	2.1%	4.1%	0.5%	0.0%	0.0%	2.1%	7'103	289
44 - Aufgelöste Baumbestände	28.9%	99.4%	51.8%	98.4%	51.8%	28.9%	37.1%	2.2%	0.6%	1.2%	28.3%	28'771	10'682
45 - Gebüschwaldbestände	11.8%	100.0%	63.0%	100.0%	63.0%	11.8%	19.9%	0.1%	0.0%	0.0%	11.8%	703	140
46 - Lineare Baumbestände	44.6%	99.8%	62.3%	99.6%	62.3%	44.6%	52.0%	0.8%	0.3%	0.5%	44.4%	10'089	5'242
47 - Baumgruppen	39.2%	99.3%	54.2%	98.7%	54.2%	39.2%	45.5%	2.1%	0.8%	1.5%	38.5%	27'506	12'522
51 - Anstehender Fels	82.1%	99.6%	79.9%	99.6%	79.9%	82.1%	81.0%	2.0%	1.6%	2.0%	81.6%	26'580	21'523
52 - Lockergestein	73.9%	99.6%	79.7%	99.4%	79.7%	73.9%	76.7%	2.1%	1.6%	2.0%	73.5%	28'671	21'993
53 - Versteinte Flächen	41.7%	99.6%	53.6%	99.3%	53.6%	41.7%	46.9%	1.1%	0.5%	0.9%	41.3%	15'237	7'148
61 - Wasser	96.1%	99.9%	97.3%	99.9%	97.3%	96.1%	96.7%	3.6%	3.4%	3.5%	96.0%	47'568	46'006
62 - Gletscher, Firn	91.9%	99.9%	91.4%	99.9%	91.4%	91.9%	91.7%	0.7%	0.7%	0.7%	91.9%	9'669	8'862
63 - Nassstandorte	0.3%	100.0%	9.1%	99.9%	9.1%	0.3%	0.5%	0.1%	0.0%	0.0%	0.3%	1'129	6
64 - Schilfbestände	29.9%	100.0%	76.5%	100.0%	76.5%	29.9%	43.0%	0.0%	0.0%	0.0%	29.9%	249	107
												1'334'756	**1'181'337**
													88.51%

Abbildung 4.21. Genauigkeitsmetriken für Land Cover AREA4 - Im Training: CNN-Klassifikationsergebnisse (RGB und FCIR getrennt), Satelliten-Zeitreihen und alle verfügbaren Zusatzdaten (Höhenmodell, Vegetationshöhen, Amtl. Vermessung); rechts dargestellt die bei derzeitiger Datenlage maximal vorhersagbare Anzahl Samples mit diesem Modell.

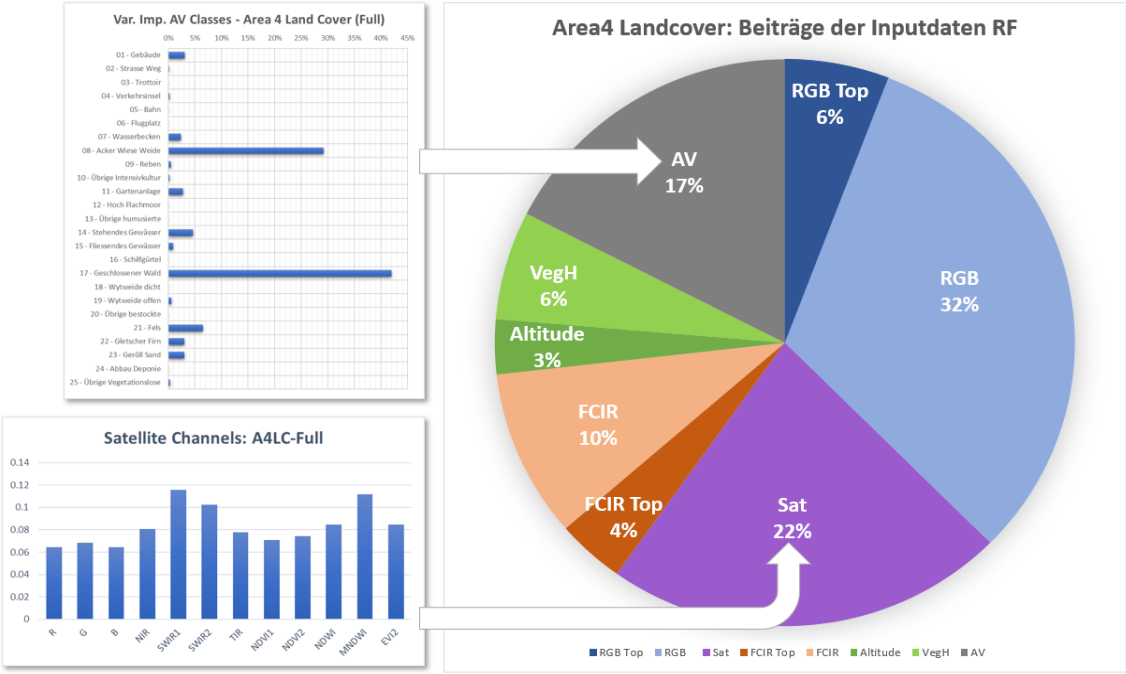


Abbildung 4.22. Variablengewichte für Land Cover AREA4 - Im Training: CNN-Klassifikationsergebnisse (RGB und FCIR getrennt), Satelliten-Zeitreihen und alle verfügbaren Zusatzdaten (Höhenmodell, Vegetationshöhen, AV).

Nach jedem Trainingsdurchlauf werden im Random Forest für jede einzelne Variable Gewichtungen exportiert. Bei den Variablengewichten fällt auf, dass die neu hinzugefügten Zusatzdaten gemeinsam lediglich ein Gewicht von ~40% erhalten. Die Resultate der RGB-CNN-Klassifikation, sowie die Satellitenzeitreihen bleiben die wichtigsten Inputparameter. Auch die Klassifikationsresultate des FCIR-CNN erreichen nicht die Wichtigkeit der RGB-CNN-Klassifikationen. Möglicherweise ist dies darauf zurückzuführen, dass für das CNN ein auf RGB-Daten vortrainiertes Netzwerk eingesetzt wurde, das dementsprechend natürlicherweise besser darin ist, RGB-Muster als FCIR-Muster zu differenzieren.

Einen überraschend hohen Anteil der Variablengewichte nehmen die Daten der amtlichen Vermessung ein. Innerhalb dieser Daten ist die Verteilung jedoch erstaunlich unausgeglichen: Lediglich die zwei Klassen „08 Acker Wiese Weide“ und „17 Geschlossener Wald“ erreichen hohe Variablengewichte. Vermutlich korrelieren diese beiden Variablen sehr stark mit den stark besetzten BFS-Klassen für Acker und Wald und machen gleichzeitig daher auch innerhalb des AV-Datensatzes einen grossen Anteil der Stichproben aus. Hinsichtlich der Variablengewichte ist zu bemerken, dass diese zwischen den betrachteten RF Modellen nicht direkt vergleichbar sind, da die Klassifikatoren auf eine unterschiedliche Variablenselektion basieren. Jedoch erklärt innerhalb eines Modells das Gewicht den Anteil der entsprechenden Variable an der Klassifikationsleistung.

4.3.4 Land Cover – Vergleich Klassifikationsmodelle

Im Vergleich zwischen den RF-Klassifikationsmodellen für Land Cover zeigt sich in allen Performanzmetriken der gleiche Trend. AREA4 mit Zusatzdaten erreicht gegenüber AREA4 ohne Zusatzdaten eine Steigerung der Genauigkeit von ca. 5%, gegenüber AREA3 ohne Zusatzdaten sogar eine Steigerung um ca. 10%.

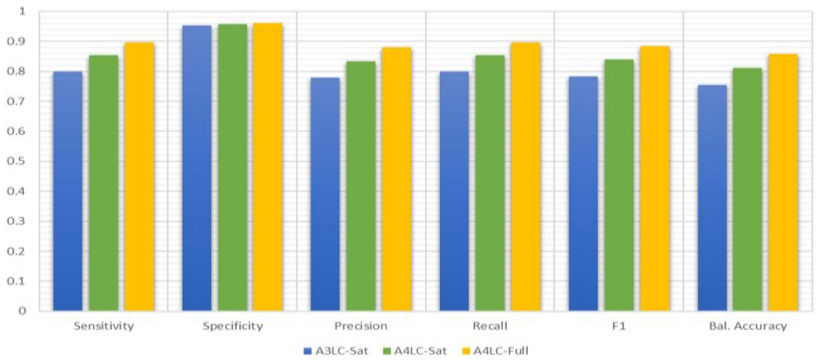


Abbildung 4.23. Vergleich der Modellperformanz (Genauigkeitsmetriken) für Land Cover.

4.3.5 Change Detection Land Cover

CD LC: Direkter Ansatz

Beim direkten Ansatz entsprechen alle Metriken der Klassifikationsresultate denjenigen, die im jeweiligen Klassifikationsmodell vorliegen.

CD LC: Tabellarischer Ansatz

Beim tabellarischen Ansatz entspricht die maximale Anzahl auswertbarer Strichproben derjenigen für welche AREA4-RGB-CNN-Klassifikationsresultate vorliegen. Insgesamt wird so lediglich eine Genauigkeit von 67,51% erreicht, da im Umkehrschluss für etwa 1/3 der Stichproben entweder im AREA3- oder im AREA4-Modell eine Fehlklassifikation vorlag. Die Endgenauigkeit entspricht in erster konservativer Näherung einer Multiplikation der Genauigkeitsmetriken der Klassifikationsmodelle von AREA3 und AREA4. Es können jedoch theoretisch dadurch leicht bessere Werte erreicht werden, dass in beiden Fällen die gleichen Stichproben einer Fehlklassifikation unterliegen. Wie bereits bei den einzelnen Klassifikationen sind es besonders homogene, grössere Klassen die eine gute Klassifikationsgenauigkeit erreichen: 21, 34, 41, 61 und 62.

N_{max} = 2‘146‘743
Accuracy = 67.51 %

11 - Befestigte Flächen	55'228
12 - Gebäude	9'259
13 - Treibhäuser	19
14 - Beetstrukturen	13
15 - Rasen	3'834
16 - Bäume auf künstl. Flächen	788
17 - Gemischte Kleinstrukturen	-
21 - Gras-, Krautvegetation	683'160
31 - Gebüsch	2'139
32 - Verbuschte Flächen	358
33 - Niederstammobst	275
34 - Reben	1'216
35 - Gärtnerische Dauerkulturen	-
41 - Geschlossene Baumbestände	473'229
42 - Waldecken	-
43 - Waldstreifen	48
44 - Aufgelöste Baumbestände	2'826
45 - Gebüschwaldbestände	3'689
46 - Lineare Baumbestände	3'379
47 - Baumgruppen	5'388
51 - Anstehender Fels	41'598
52 - Lockergestein	39'238
53 - Versteinte Flächen	9'033
61 - Wasser	89'536
62 - Gletscher, Firn	25'072
63 - Nassstandorte	6
64 - Schilfbestände	0
	1'449'330

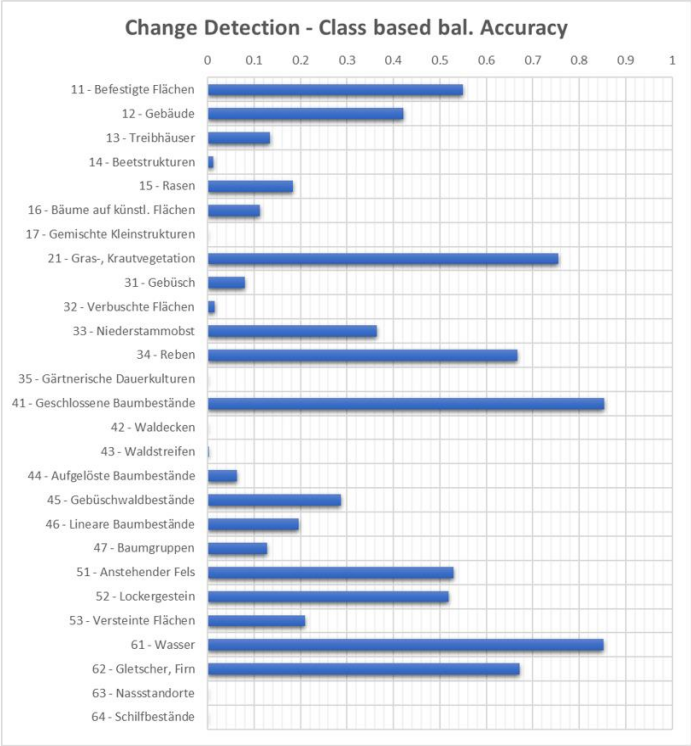


Abbildung 4.24. Change Detection mittels tabellarischem Ansatz; im Vergleich die Land Cover Klassifikationsresultate AREA3 vs. AREA4 aus den Modellen CNN (RGB) + Satellitenzeitreihen.

CD LC: Dichotomer RF-Ansatz

Beim dichotomen Ansatz wird für die Change Detection ein eigenes RF-Modell trainiert. Die Ausgangsvariable ist die Entscheidung Change / No Change für jedes einzelne Sample. Um einen starken Bias für „No Change“ (etwa 97% der Samples) auszuschliessen muss bei diesem Modell der Datensatz ausgeglichen sein. In einem nicht ausgeglichenen Modell wurden praktisch alle Samples als „No Change“ eingestuft, was zwar mathematisch in einer Gesamtgenauigkeit von ~96% resultierte, jedoch praktisch keine eigentlichen Changes detektiert. Dies bedeutet, dass sich nur so viele „No Change“-Samples im Modell befinden dürfen, wie es Change-Samples gibt. Gegenüber dem tabellarischen Ansatz ist dies eine Limitation, die vor allem die Grösse des Trainingsdatensatzes beeinflusst. Im vorliegenden Szenario befanden sich 26‘500 Samples im Training und weitere 10‘284 Samples in der Prädiktion. Da sich zwischen den letzten zwei Erhebungsschritten ca. 100‘000 Stichproben geändert haben, könnte hier also bereits jetzt noch ein etwa im Faktor 3 grösseres Modell berechnet werden. Dabei ist jedoch unklar, ob sich die Werte dadurch noch erheblich verbessern würden. Eine Erweiterung des „Change“-Datensatzes wäre wahrscheinlich nur durch die manuelle Klassifikation neuer „Change“-Punkte in der Arealstatistik plus einer Erzeugung neuer Zusatzdaten möglich. Augmentierungen unter Zuhilfenahme gleicher Zusatzdaten würden eher zu einem Bias führen. Eine stärkere Generalisierung ist nur durch zusätzliche, vollständige, echte Datensätze möglich.

	CD LC
Sensitivity	74.4%
Specificity	83.4%
Pos Pred Value	81.6%
Neg Pred Value	76.7%
Precision	81.6%
Recall	74.4%
F1	77.8%
Prevalence	49.8%
Detection Rate	37.0%
Detection Prevalence	45.4%
Balanced Accuracy	78.9%
Samples Change	13188
Samples NoChange	13312

Abbildung 4.25. Change Detection (Land Cover) mittels direktem Machine-Learning-Ansatz; Dieses Random-Forest-Modell wurde auf die dichotome Entscheidung Change / No Change trainiert; im Training: CNN (RGB) jeweils AREA3 und 4, sowie Satellitenzeitreihen jeweils AREA3 und 4.

Mit 77,8% Gesamtgenauigkeit (F1) wird derzeit jedes vierte bis fünfte „Change“ oder „No Change“-Element falsch detektiert. Dabei sollte jedoch beachtet werden, dass die statistische Baseline dieser dichotomen Entscheidung bei 50% liegt. Die Gesamtgenauigkeit ist somit höher als mit dem tabellarischen Ansatz, jedoch u.U. niedriger als die Genauigkeit einzelner grosser Klassen. Aufgrund der grundsätzlich niedrigeren Genauigkeiten beider CD-Modelle empfiehlt sich beim derzeitigen Stand jedoch die Anwendung eines reinen Klassifikationsmodells (AREA4 LC mit Zusatzdaten) in Kombination mit einem Datenbankabgleich zur Ground-Truth-Klasse der vorherigen Erhebungsperiode. Die erreichbare Change-Detection-Präzision dieses Ansatzes würde der Gesamtpräzision des Modells entsprechen und somit bei 88,51% liegen (vgl. Kapitel 4.3.3).

Bei der Variablengewicht-Verteilung zeigt sich eine Schwachstelle des dichotomen Change-Detection-trainierten Ansatzes: Die Klassifikationsresultate des AREA3-CNN-Modells überwiegen die Inputgewichte des AREA4-CNN-Modells stark. Dies bedeutet, dass häufiger rein von der AREA3-Einstufung auf ein Change / No Change-Event geschlossen wird (Abbildung 4.26).

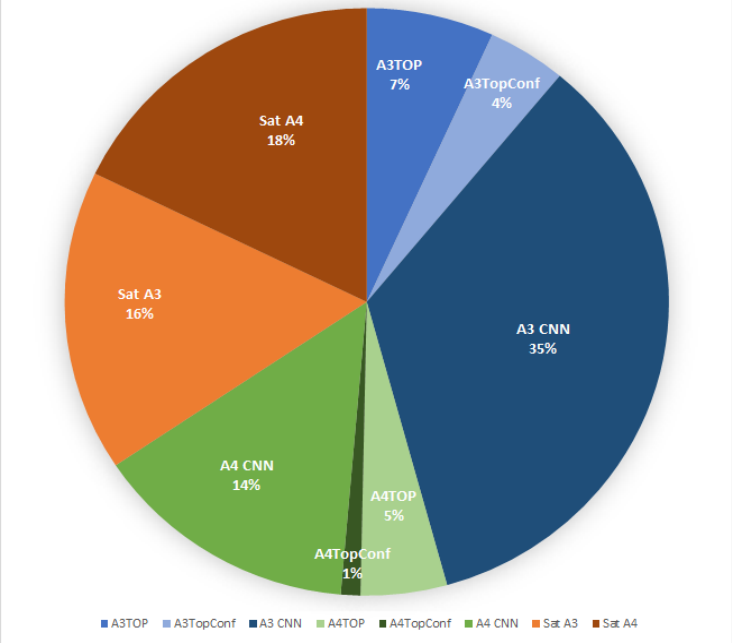
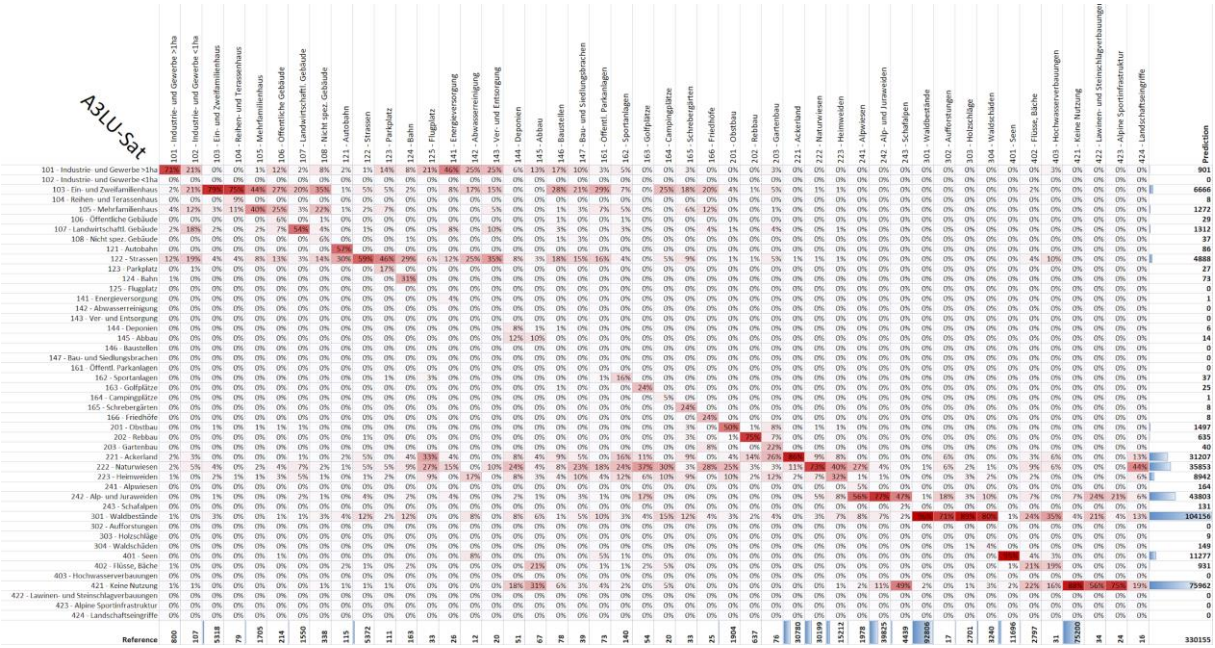


Abbildung 4.26. Verteilung der Variablengewichte im dichotomen Change Detection Random Forest Modell (Land Cover).

4.3.6 Land Use AREA3

Ähnlich dem LC-Modell für AREA3 ist auch für das LU-Modell als Inputdatensatz lediglich eine Kombination aus Deep-Learning-Resultaten und den Satellitenzeitreihen möglich, da andere Daten aufgrund ihrer Zeitstände nicht kompatibel sind. Im Training befanden sich 165‘058 Samples, in der Prädiktion 330‘155 Samples. Die Overall Accuracy lag mit 77,28% ähnlich hoch wie im AREA3-LC-Modell (78,22%, vgl. Kapitel 4.3.1).



A3LU-Sat	101 - Industrie- und Gewerbe >1ha	102 - Industrie- und Gewerbe <1ha	103 - Ein- und Zweifamilienhaus	104 - Reihen- und Terassenhaus	105 - Mehrfamilienhaus	106 - Öffentliche Gebäude	107 - Landwirtschaftl. Gebäude	108 - Nicht spez. Gebäude	121 - Autobahn	122 - Strassen	123 - Parkplatz	124 - Bahn	125 - Flugplatz	141 - Energieversorgung	142 - Abwasserreinigung	143 - Ver- und Entsorgung	144 - Deponien	145 - Abbau	146 - Baustellen	147 - Bau- und Siedlungsbrachen	161 - Öffentl. Parkanlagen	162 - Sportanlagen	163 - Golfplätze	164 - Campingplätze	165 - Schrebergärten	166 - Friedhöfe	201 - Obstbau	202 - Rebbau	203 - Gartenbau	221 - Ackerland	222 - Naturwiesen	223 - Heimweiden	241 - Alpweiden	242 - Alp- und Juraweiden	243 - Schafalpen	301 - Waldbestände	302 - Aufforstungen	303 - Holzschläge	304 - Waldschäden	401 - Seen	402 - Flüsse, Bäche	403 - Hochwasserverbauungen	421 - Keine Nutzung	422 - Lawinen- und Steinschlagverbauungen	423 - Alpine Sportinfrastruktur	424 - Landschaftseingriffe	Prediction
101 - Industrie- und Gewerbe >1ha	71%	21%	0%	0%	1%	12%	2%	8%	2%	1%	14%	8%	21%	46%	25%	25%	6%	13%	17%	10%	3%	5%	0%	0%	3%	0%	0%	0%	0%	3%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3%	0%	0%	0%	0%	0%	901
102 - Industrie- und Gewerbe <1ha	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0
103 - Ein- und Zweifamilienhaus	2%	21%	71%	7%	44%	27%	20%	35%	1%	5%	5%	2%	0%	8%	17%	15%	0%	0%	28%	21%	29%	7%	0%	25%	18%	20%	4%	1%	5%	0%	1%	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%	0%	0%	0%	6666
104 - Reihen- und Terassenhaus	0%	0%	0%	9%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	8
105 - Mehrfamilienhaus	4%	12%	3%	11%	40%	25%	3%	22%	1%	2%	7%	0%	0%	0%	0%	5%	0%	0%	1%	3%	7%	5%	0%	0%	6%	12%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1272
106 - Öffentliche Gebäude	0%	0%	0%	0%	0%	6%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	29
107 - Landwirtschaftl. Gebäude	2%	18%	2%	0%	2%	2%	7%	54%	4%	0%	1%	0%	0%	8%	0%	10%	0%	0%	3%	0%	0%	3%	0%	0%	0%	4%	1%	0%	4%	0%	0%	3%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1312
108 - Nicht spez. Gebäude	0%	0%	0%	0%	0%	0%	0%	6%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	1%	3%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	37
121 - Autobahn	0%	0%	0%	0%	0%	0%	0%	0%	57%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	86
122 - Strassen	12%	19%	4%	4%	8%	13%	3%	14%	30%	39%	46%	29%	6%	12%	25%	35%	8%	3%	18%	15%	16%	4%	0%	5%	9%	0%	1%	1%	5%	1%	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	4%	10%	0%	0%	0%	0%	4888
123 - Parkplatz	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	17%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	27
124 - Bahn	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	31%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	73
125 - Flugplatz	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0
141 - Energieversorgung	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	4%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1
142 - Abwasserreinigung	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0
143 - Ver- und Entsorgung	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0
144 - Deponien	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	8%	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	6
145 - Abbau	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	12%	10%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	14
146 - Baustellen	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0
147 - Bau- und Siedlungsbrachen	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0
161 - Öffentl. Parkanlagen	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0
162 - Sportanlagen	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	3%	0%	0%	0%	0%	0%	0%	0%	0%	1%	16%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	37
163 - Golfplätze	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	24%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	25
164 - Campingplätze	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	5%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1
165 - Schrebergärten	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	24%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	8
166 - Friedhöfe	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	24%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	8
201 - Obstbau	0%	0%	1%	0%	1%	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3%	0%	32%	1%	8%	0%	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1497
202 - Rebbau	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3%	0%	1%	95%	7%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	635
203 - Gartenbau	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	8%	0%	0%	22%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	40
221 - Ackerland	2%	3%	0%	0%	0%	0%	1%	0%	2%	5%	0%	4%	33%	4%	0%	0%	8%	4%	9%	5%	0%	16%	11%	0%	9%	0%	4%	14%	76%	98%	9%	8%	0%	0%	0%	0%	6%	0%	0%	0%	0%	3%	6%	0%	0%	0%	13%	31207
222 - Naturwiesen	2%	5%	4%	0%	2%	4%	7%	2%	1%	5%	5%	9%	27%	15%	0%	10%	24%	4%	8%	23%	18%	24%	37%	30%	3%	28%	25%	3%	3%	1%	23%	40%	27%	4%	0%	1%	6%	2%	1%	0%	9%	6%	0%	0%	0%	0%	48%	35853
223 - Heimweiden	1%	0%	2%	1%	1%	3%	5%	1%	0%	1%	2%	0%	9%	0%	17%	0%	8%	3%	4%	10%	4%	12%	6%	10%	9%	0%	10%	2%	12%	2%	7%	12%	1%	1%	0%	0%	0%	3%	2%	0%	2%	0%	0%	0%	0%	0%	6%	8942
241 - Alpweiden	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	5%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	164
242 - Alp- und Juraweiden	0%	0%	1%	0%	0%	0%	2%	1%	0%	4%	0%	2%	0%	4%	0%	0%	2%	1%	1%	3%	3%	1%	0%	17%	0%	0%	0%	0%	0%	0%	0%	5%	8%	56%	27%	47%	1%	18%	3%	10%	0%	7%	0%	7%	24%	21%	6%	43803
243 - Schafalpen	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	131
301 - Waldbestände	1%	0%	3%	0%	0%	1%	1%	3%	4%	12%	2%	12%	0%	0%	8%	0%	8%	6%	1%	5%	10%	3%	4%	15%	12%	4%	3%	2%	4%	0%	3%	7%	8%	7%	2%	100%	71%	100%	90%	1%	24%	35%	4%	21%	4%	13%	104156
302 - Aufforstungen	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0
303 - Holzschläge	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	9
304 - Waldschäden	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	4%	0%	0%	0%	0%	0%	0%	0%	149
401 - Seen	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	8%	0%	0%	0%	0%	0%	5%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%	4%	3%	0%	0%	0%	0%	11277
402 - Flüsse, Bäche	1%	0%	0%	0%	0%	0%	0%	0%	2%	1%	0%	2%	0%	0%	0%	0%	21%	0%	0%	1%	1%	1%	2%	5%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	21%	19%	0%	0%	0%	0%	931
403 - Hochwasserverbauungen	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0
421 - Keine Nutzung	1%	1%	0%	0%	0%	0%	0%	1%	1%	1%	1%	0%	0%	0%	0%	18%	11%	6%	3%	4%	2%	0%	5%	0%	0%	0%	0%	0%	0%	0%	0%	1%	2%	11%	49%	2%	0%	1%	3%	2%	22%	16%	100%	56%	72%	19%	75962
422 - Lawinen- und Steinschlagverbauungen	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0
423 - Alpine Sportinfrastruktur	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0
424 - Landschaftseingriffe	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0
Reference	800	107	5318	79	1705	214	1550	338	115	5372	111	163	33	26	12	20	51	67	78	39	73	140	54	20	33	25	1904	637	76	22702	30139	15214	1978	39825	4439	92800	17	2701	3240	11698	2797	31	72320	34	24	16	330155

Abbildung 4.27. Confusion Matrix (normalisiert) für Land Use AREA3 - Im Training: CNN-Klassifikationsergebnisse (RGB) und Satelliten-Zeitreihen.

Genauigkeit A3LU-Sat (F1)	Sensitivity	Specificity	Pos. Pred.	Neg. Pred.	Precision	Recall	F1	Prevalence	Det. Rate	Det. Preval.	Bal. Accuracy	Samples	Correct Samples
101 - Industrie- und Gewerbe >1ha	70.8%	99.9%	62.8%	99.9%	62.8%	70.8%	66.5%	0.2%	0.2%	0.3%	70.6%	9'073	6'038
102 - Industrie- und Gewerbe <1ha	0.0%	100.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	1'214	-
103 - Ein- und Zweifamilienhaus	79.3%	99.2%	63.2%	99.7%	63.2%	79.3%	70.4%	1.6%	1.3%	2.0%	78.5%	60'315	42'438
104 - Reihen- und Terassenhaus	8.9%	100.0%	87.5%	100.0%	87.5%	8.9%	16.1%	0.0%	0.0%	0.0%	8.9%	896	144
105 - Mehrfamilienhaus	39.9%	99.8%	53.5%	99.7%	53.5%	39.9%	45.8%	0.5%	0.2%	0.4%	39.8%	19'337	8'847
106 - Öffentliche Gebäude	6.1%	100.0%	44.8%	99.9%	44.8%	6.1%	10.7%	0.1%	0.0%	0.0%	6.1%	2'427	260
107 - Landwirtschaftl. Gebäude	53.7%	99.9%	63.4%	99.8%	63.4%	53.7%	58.1%	0.5%	0.3%	0.4%	53.5%	17'579	10'221
108 - Nicht spez. Gebäude	5.6%	100.0%	51.4%	99.9%	51.4%	5.6%	10.1%	0.1%	0.0%	0.0%	5.6%	3'833	388
121 - Autobahn	56.5%	100.0%	75.6%	100.0%	75.6%	56.5%	64.7%	0.0%	0.0%	0.0%	56.5%	1'304	844
122 - Strassen	59.3%	99.5%	65.2%	99.3%	65.2%	59.3%	62.1%	1.6%	1.0%	1.5%	58.8%	60'927	37'851
123 - Parkplatz	17.1%	100.0%	70.4%	100.0%	70.4%	17.1%	27.5%	0.0%	0.0%	0.0%	17.1%	1'259	347
124 - Bahn	30.7%	100.0%	68.5%	100.0%	68.5%	30.7%	42.4%	0.0%	0.0%	0.0%	30.7%	1'849	783
125 - Flugplatz	0.0%	100.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	374	-
141 - Energieversorgung	3.8%	100.0%	100.0%	100.0%	100.0%	3.8%	7.4%	0.0%	0.0%	0.0%	3.8%	295	22
142 - Abwasserreinigung	0.0%	100.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	136	-
143 - Ver- und Entsorgung	0.0%	100.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	227	-
144 - Deponien	7.8%	100.0%	66.7%	100.0%	66.7%	7.8%	14.0%	0.0%	0.0%	0.0%	7.8%	578	81
145 - Abbau	10.4%	100.0%	50.0%	100.0%	50.0%	10.4%	17.3%	0.0%	0.0%	0.0%	10.4%	760	131
146 - Baustellen	0.0%	100.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	885	-
147 - Bau- und Siedlungsbrachen	0.0%	100.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	442	-
161 - Öffentl. Parkanlagen	0.0%	100.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	828	-
162 - Sportanlagen	15.7%	100.0%	59.5%	100.0%	59.5%	15.7%	24.9%	0.0%	0.0%	0.0%	15.7%	1'588	395
163 - Golfplätze	24.1%	100.0%	52.0%	100.0%	52.0%	24.1%	32.9%	0.0%	0.0%	0.0%	24.1%	612	202
164 - Campingplätze	5.0%	100.0%	100.0%	100.0%	100.0%	5.0%	9.5%	0.0%	0.0%	0.0%	5.0%	227	22
165 - Schrebergärten	24.2%	100.0%	100.0%	100.0%	100.0%	24.2%	39.0%	0.0%	0.0%	0.0%	24.2%	374	146
166 - Friedhöfe	24.0%	100.0%	75.0%	100.0%	75.0%	24.0%	36.4%	0.0%	0.0%	0.0%	24.0%	284	103
201 - Obstbau	50.5%	99.8%	64.2%	99.7%	64.2%	50.5%	56.5%	0.6%	0.3%	0.5%	50.3%	21'594	12'204
202 - Rebbau	74.6%	100.0%	74.8%	100.0%	74.8%	74.6%	74.7%	0.2%	0.1%	0.2%	74.5%	7'225	5'396
203 - Gartenbau	22.4%	100.0%	42.5%	100.0%	42.5%	22.4%	29.3%	0.0%	0.0%	0.0%	22.4%	862	253
221 - Ackerland	86.0%	98.4%	84.8%	98.6%	84.8%	86.0%	85.4%	9.3%	8.0%	9.5%	84.4%	349'094	298'177
222 - Naturwiesen	73.0%	95.4%	61.5%	97.2%	61.5%	73.0%	66.7%	9.1%	6.7%	10.9%	68.4%	342'504	228'582
223 - Heimweiden	32.0%	98.7%	54.4%	96.8%	54.4%	32.0%	40.3%	4.6%	1.5%	2.7%	30.7%	172'528	69'457
241 - Alpwiesen	5.0%	100.0%	60.4%	99.4%	60.4%	5.0%	9.2%	0.6%	0.0%	0.0%	5.0%	22'434	2'074
242 - Alp- und Juraweiden	76.9%	95.5%	69.9%	96.8%	69.9%	76.9%	73.2%	12.1%	9.3%	13.3%	72.3%	451'679	330'771
243 - Schafalpen	1.8%	100.0%	61.8%	98.7%	61.8%	1.8%	3.5%	1.3%	0.0%	0.0%	1.8%	50'345	1'785
301 - Waldbestände	95.7%	93.5%	85.3%	98.2%	85.3%	95.7%	90.2%	28.1%	26.9%	31.5%	89.3%	1'052'567	949'566
302 - Aufforstungen	0.0%	100.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	193	-
303 - Holzschläge	0.3%	100.0%	77.8%	99.2%	77.8%	0.3%	0.5%	0.8%	0.0%	0.0%	0.3%	30'634	158
304 - Waldschäden	3.7%	100.0%	81.2%	99.1%	81.2%	3.7%	7.1%	1.0%	0.0%	0.0%	3.7%	36'747	2'624
401 - Seen	95.0%	99.9%	98.5%	99.8%	98.5%	95.0%	96.7%	3.5%	3.4%	3.4%	95.0%	132'651	128'326
402 - Flüsse, Bäche	21.4%	99.9%	64.3%	99.3%	64.3%	21.4%	32.1%	0.8%	0.2%	0.3%	21.3%	31'722	10'194
403 - Hochwasserverbauungen	0.0%	100.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	352	-
421 - Keine Nutzung	87.8%	96.1%	86.9%	96.4%	86.9%	87.8%	87.3%	22.8%	20.0%	23.0%	83.9%	852'887	744'782
422 - Lawinen- und Steinschlagvert	0.0%	100.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	386	-
423 - Alpine Sportinfrastruktur	0.0%	100.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	272	-
424 - Landschaftseingriffe	0.0%	100.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	181	-
												3'744'480	2'893'610

Abbildung 4.28. Genauigkeitsmetriken für Land Use AREA3 - Im Training: CNN-Klassifikationsergebnisse (RGB) und Satelliten-Zeitreihen; rechts dargestellt die bei derzeitiger Datenlage maximal vorhersagbare Anzahl Samples mit diesem Modell. Overall Accuracy: 77,28%

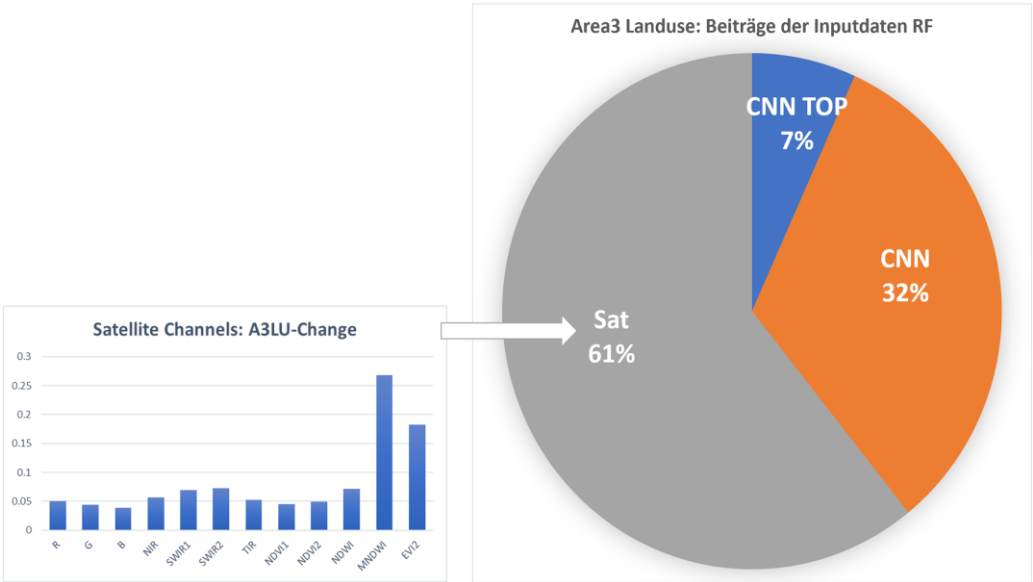


Abbildung 4.29. Variablengewichte für Land Use AREA3 - Im Training: CNN-Klassifikationsergebnisse (RGB) und Satelliten-Zeitreihen.

Aufgrund der grösseren Anzahl Features in der LU-Klassifikation ist auch die Grösse des CNN-Input-Vektors pro Sample gewachsen: 46x Klassenkonfidenzen + 46x TOP1 One-Hot-Encoded Klassenfaktoren + 1x TOP1 Konfidenz = 93 Variablen. Die Satelliten-Zeitreihen bestanden weiterhin aus 120 Variablen. Erstaunlich ist, dass sich die Variablengewichte des CNN trotz des gegenüber LC gewachsenen Inputvektors im Verhältnis zu den Satellitendaten reduzierten. Im Umkehrschluss ist also die Aussagekraft der CNN-Klassifikation bei AREA3 hinsichtlich Land Use niedriger als bei Land Cover. Eine extrem starke Zunahme der Gewichtung findet sich vor allem bei den Satellitenindizes MNDWI und EVI2.

Das Bild, dass stark besetzte, homogene Klassen besser abschneiden als Klassen mit niedriger Prävalenz setzt sich teilweise auch für Land Use fort. Werte über 85% Genauigkeit (F1) werden von den Klassen 221 – Ackerland, 301 – Waldbestände, 401 – Seen und 421 – Keine Nutzung erreicht. Alle Klassen mit einer Prävalenz < 0,15% (27 der 46 LU-Klassen) konnten keine Genauigkeiten (F1) oberhalb von 50% erreichen.

4.3.7 Land Use AREA4

Im Vergleich zum AREA3-LU-Modell (77,28%) erreicht das AREA4-LU-Modell eine leicht höhere Gesamtgenauigkeit von 78,18%, jedoch schneiden vor allem kleinere Klassen nun deutlich besser ab. Im Vergleich mit dem korresponierenden LC-Modell (83,90%) liegt die Gesamtgenauigkeit jedoch niedriger. Im Training befanden sich hier 210‘333 Samples, in der Prädiktion 841‘331 Samples. Der bereits bei den LC-Modellen gewonnene Eindruck, dass die höher auflösenden Inputdaten des CNN auch zu einer höheren Gewichtung dieser Klassifikationsergebnisse führen, verfestigt sich auch bei den LU-Modellen: Der CNN-Variabelgewicht-Anteil nahm um 11% zu. Innerhalb der Satellitenzeitreihen diversifiziert sich der Input über mehrere Indizes. Der Anteil ist mit 50% jedoch immer noch wesentlich höher als beim AREA4-LC-Modell. Zusätzlich erreicht nun die Klasse 202 – Rebbau eine Klassifikationsgenauigkeit von >90%, aber auch andere Landnutzungsklassen können erheblich besser klassifiziert werden.

Im Vergleich zum AREA3-LU-Modell erreicht das AREA4-LU-Modell eine leicht höhere Gesamtgenauigkeit von 78,18%, jedoch schneiden vor allem kleinere Klassen nun deutlich besser ab. Im Training befanden sich hier 210‘333 Samples, in der Prädiktion 841‘331 Samples.

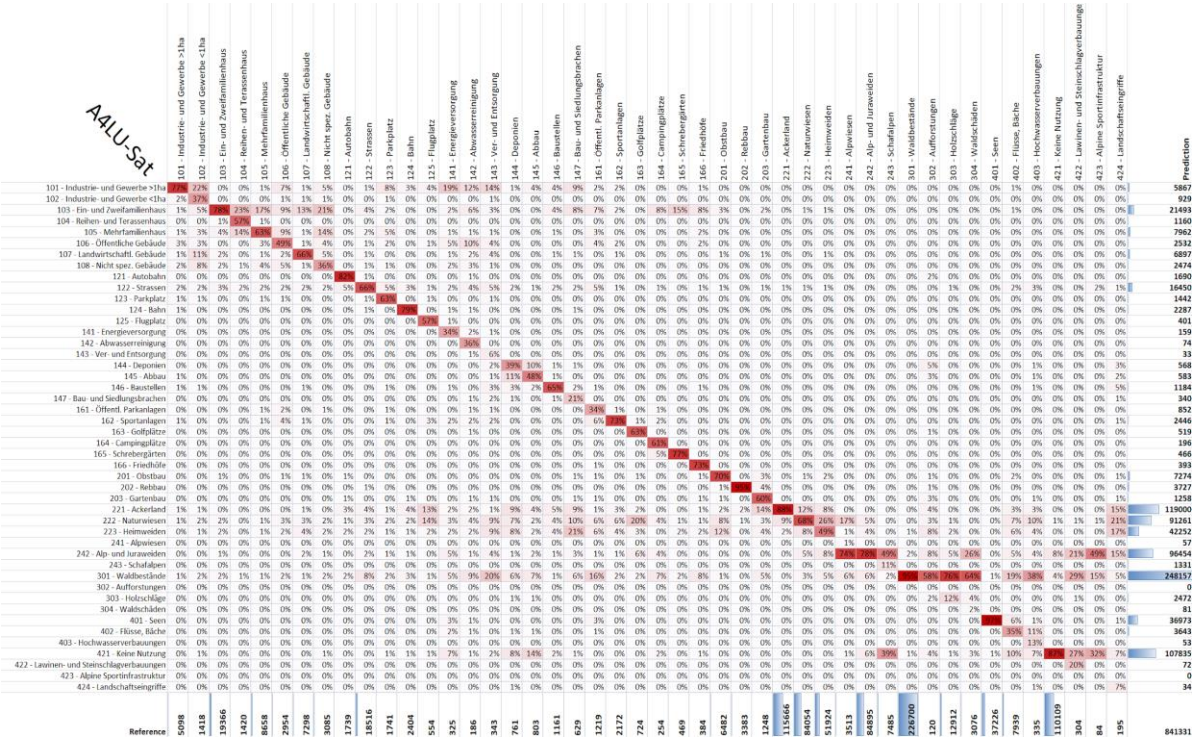


Abbildung 4.30. Confusion Matrix (normalisiert) für Land Use AREA4 - Im Training: CNN-Klassifikationsergebnisse (RGB) und Satelliten-Zeitreihen.

Genauigkeit A4LU-Sat (F1)	Sensitivity	Specificity	Pos. Pred.	Neg. Pred.	Precision	Recall	F1	Prevalence	Det. Rate	Det. Preval	Bal. Accuracy	Samples	Correct Samples
101 - Industrie- und Gewerbe >1h	77,1%	99,8%	67,0%	99,9%	67,0%	77,1%	71,7%	0,6%	0,5%	0,7%	76,9%	13.008	9.329
102 - Industrie- und Gewerbe <1h	36,8%	100,0%	56,2%	99,9%	56,2%	36,8%	44,5%	0,2%	0,1%	0,1%	36,8%	3.618	1.609
103 - Ein- und Zweifamilienhaus	78,2%	99,2%	70,4%	99,5%	70,4%	78,2%	74,1%	2,3%	1,8%	2,6%	77,4%	49.414	36.623
104 - Reihen- und Terassenhaus	57,1%	100,0%	69,9%	99,9%	69,9%	57,1%	62,9%	0,2%	0,1%	0,1%	57,1%	3.623	2.278
105 - Mehrfamilienhaus	63,4%	99,7%	69,0%	99,6%	69,0%	63,4%	66,1%	1,0%	0,7%	0,9%	63,1%	22.092	14.600
106 - Öffentliche Gebäude	49,0%	99,9%	57,1%	99,8%	57,1%	49,0%	52,8%	0,4%	0,2%	0,3%	48,9%	7.537	3.976
107 - Landwirtschaftl. Gebäude	65,9%	99,7%	69,8%	99,7%	69,8%	65,9%	67,8%	0,9%	0,6%	0,8%	65,7%	18.622	12.623
108 - Nicht spez. Gebäude	36,3%	99,8%	45,2%	99,8%	45,2%	36,3%	40,3%	0,4%	0,1%	0,3%	36,1%	7.872	3.169
121 - Autobahn	82,4%	100,0%	84,8%	100,0%	84,8%	82,4%	83,6%	0,2%	0,2%	0,2%	82,4%	4.437	3.709
122 - Strassen	66,0%	99,5%	74,3%	99,2%	74,3%	66,0%	69,9%	2,2%	1,5%	2,0%	65,5%	47.245	33.031
123 - Parkplatz	62,6%	100,0%	75,6%	99,9%	75,6%	62,6%	68,5%	0,2%	0,1%	0,2%	62,6%	4.442	3.043
124 - Bahn	78,7%	100,0%	82,7%	99,9%	82,7%	78,7%	80,6%	0,3%	0,2%	0,3%	78,6%	6.134	4.945
125 - Flugplatz	57,4%	100,0%	79,3%	100,0%	79,3%	57,4%	66,6%	0,1%	0,0%	0,0%	57,4%	1.414	941
141 - Energieversorgung	34,5%	100,0%	70,4%	100,0%	70,4%	34,5%	46,3%	0,0%	0,0%	0,0%	34,5%	829	384
142 - Abwasserreinigung	36,0%	100,0%	90,5%	100,0%	90,5%	36,0%	51,5%	0,0%	0,0%	0,0%	36,0%	475	245
143 - Ver- und Entsorgung	5,8%	100,0%	60,6%	100,0%	60,6%	5,8%	10,6%	0,0%	0,0%	0,0%	5,8%	875	93
144 - Deponien	38,9%	100,0%	52,1%	99,9%	52,1%	38,9%	44,5%	0,1%	0,0%	0,1%	38,9%	1.942	865
145 - Abbau	48,4%	100,0%	66,7%	100,0%	66,7%	48,4%	56,1%	0,1%	0,0%	0,1%	48,4%	2.049	1.150
146 - Baustellen	65,0%	99,9%	63,8%	100,0%	63,8%	65,0%	64,4%	0,1%	0,1%	0,1%	65,0%	2.962	1.908
147 - Bau- und Siedlungsbrachen	20,8%	100,0%	38,5%	99,9%	38,5%	20,8%	27,0%	0,1%	0,0%	0,0%	20,8%	1.605	434
161 - Öffentl. Parkanlagen	33,7%	99,9%	48,2%	99,9%	48,2%	33,7%	39,7%	0,1%	0,0%	0,1%	33,7%	3.110	1.235
162 - Sportanlagen	73,2%	99,9%	65,0%	99,9%	65,0%	73,2%	68,8%	0,3%	0,2%	0,3%	73,1%	5.542	3.814
163 - Golfplätze	62,6%	100,0%	87,3%	100,0%	87,3%	62,6%	72,9%	0,1%	0,1%	0,1%	62,6%	1.847	1.347
164 - Campingplätze	61,4%	100,0%	79,6%	100,0%	79,6%	61,4%	69,3%	0,0%	0,0%	0,0%	61,4%	648	449
165 - Schrebergärten	76,5%	100,0%	77,0%	100,0%	77,0%	76,5%	76,8%	0,1%	0,0%	0,1%	76,5%	1.197	919
166 - Friedhöfe	72,9%	100,0%	71,2%	100,0%	71,2%	72,9%	72,1%	0,0%	0,0%	0,0%	72,9%	980	706
201 - Obstbau	70,0%	99,7%	62,4%	99,8%	62,4%	70,0%	66,0%	0,8%	0,5%	0,9%	69,7%	16.539	10.917
202 - Rebbau	95,4%	99,9%	86,6%	100,0%	86,6%	95,4%	90,8%	0,4%	0,4%	0,4%	95,3%	8.632	7.836
203 - Gartenbau	60,4%	99,9%	59,9%	99,9%	59,9%	60,4%	60,2%	0,1%	0,1%	0,1%	60,4%	3.184	1.916
221 - Ackerland	88,1%	97,6%	85,7%	98,1%	85,7%	88,1%	86,9%	13,7%	12,1%	14,1%	85,8%	295.134	256.395
222 - Naturwiesen	68,4%	95,5%	63,0%	96,5%	63,0%	68,4%	65,5%	10,0%	6,8%	10,8%	63,9%	214.472	140.581
223 - Heimweiden	48,7%	97,9%	59,9%	96,7%	59,9%	48,7%	53,7%	6,2%	3,0%	5,0%	46,6%	132.489	71.157
241 - Alpwiesen	1,0%	100,0%	59,6%	99,6%	59,6%	1,0%	1,9%	0,4%	0,0%	0,0%	1,0%	8.964	171
242 - Alp- und Juraweiden	78,2%	96,0%	68,8%	97,5%	68,8%	78,2%	73,2%	10,1%	7,9%	11,5%	74,2%	216.618	158.510
243 - Schafalpen	10,6%	99,9%	59,8%	99,2%	59,8%	10,6%	18,1%	0,9%	0,1%	0,2%	10,6%	19.099	3.449
301 - Waldbestände	95,5%	94,8%	87,2%	98,3%	87,2%	95,5%	91,2%	26,9%	25,7%	29,5%	90,3%	578.448	527.302
302 - Aufforstungen	0,0%	100,0%	0,0%	100,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	306	-
303 - Holzschläge	12,5%	99,9%	65,3%	98,7%	65,3%	12,5%	21,0%	1,5%	0,2%	0,3%	12,4%	32.946	6.909
304 - Waldschäden	1,9%	100,0%	71,6%	99,6%	71,6%	1,9%	3,7%	0,4%	0,0%	0,0%	1,9%	7.849	288
401 - Seen	97,3%	99,9%	98,0%	99,9%	98,0%	97,3%	97,7%	4,4%	4,3%	4,4%	97,2%	94.986	92.773
402 - Flüsse, Bäche	35,4%	99,9%	77,1%	99,4%	77,1%	35,4%	48,5%	0,9%	0,3%	0,4%	35,3%	20.257	9.823
403 - Hochwasserverbauungen	12,8%	100,0%	81,1%	100,0%	81,1%	12,8%	22,2%	0,0%	0,0%	0,0%	12,8%	855	189
421 - Keine Nutzung	86,8%	98,3%	88,6%	98,0%	88,6%	86,8%	87,7%	13,1%	11,4%	12,8%	85,1%	280.954	246.388
422 - Lawinen- und Steinschlagve	20,1%	100,0%	84,7%	100,0%	84,7%	20,1%	32,4%	0,0%	0,0%	0,0%	20,1%	776	252
423 - Alpine Sportinfrastruktur	0,0%	100,0%	0,0%	100,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	214	-
424 - Landschaftseingriffe	7,2%	100,0%	41,2%	100,0%	41,2%	7,2%	12,2%	0,0%	0,0%	0,0%	7,2%	498	61
												2.146.743	**1.678.340**

Abbildung 4.31. Genauigkeitsmetriken für Land Use AREA4 - Im Training: CNN-Klassifikationsergebnisse (RGB) und Satelliten-Zeitreihen; rechts dargestellt die bei derzeitiger Datenlage maximal vorhersagbare Anzahl Samples mit diesem Modell. Overall Accuracy: 78,18%.

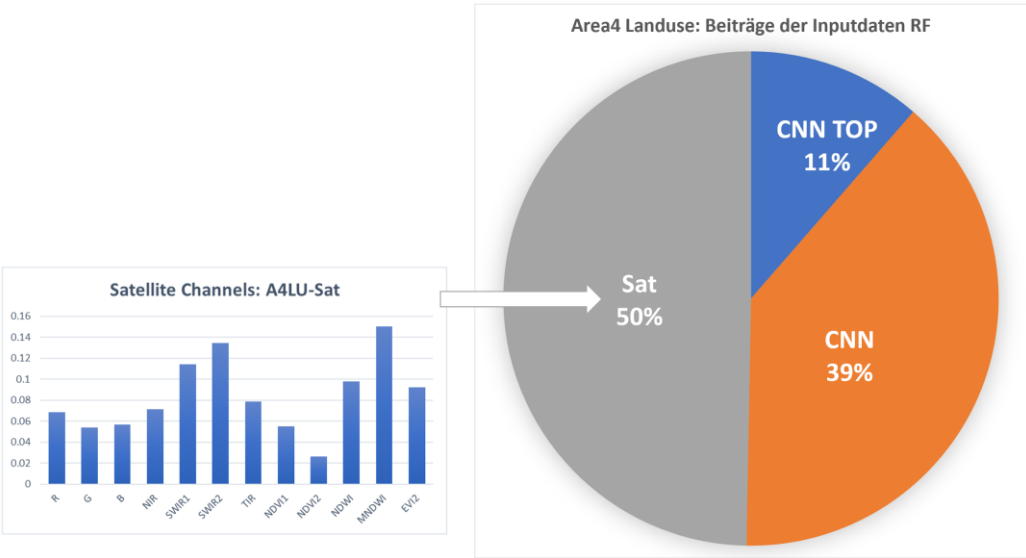


Abbildung 4.32. Variablengewichte für Land Use AREA4 - Im Training: CNN-Klassifikationsergebnisse (RGB) und Satelliten-Zeitreihen.

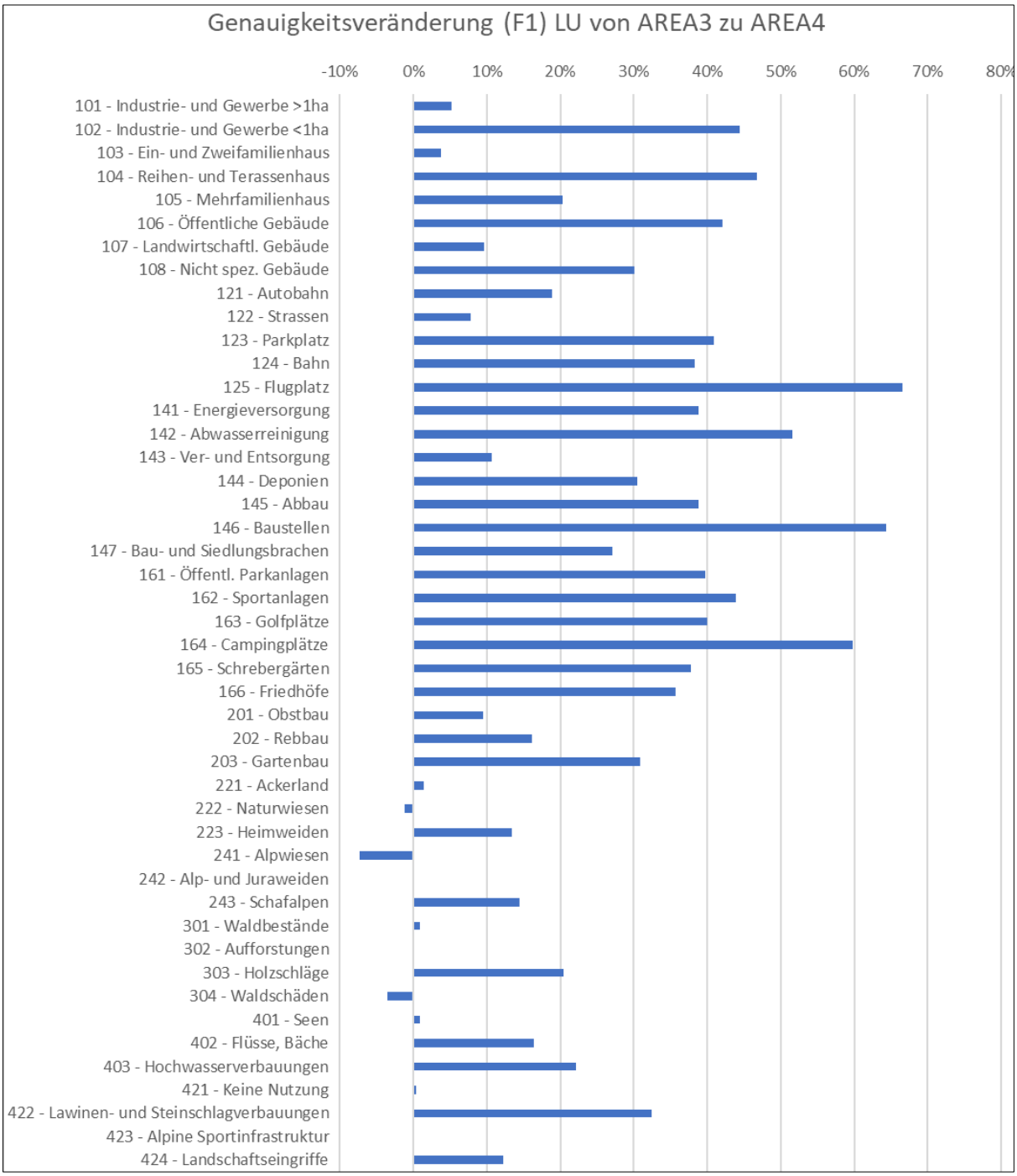


Abbildung 4.33. Veränderungen des F1-Scores nach Klasse; LU-Modelle für AREA3 und AREA4 im Vergleich.

Sehr starke Verbesserungen zeigen sich vor allem bei den Klassen 102 – Industrie- und Gewerbeareal <1ha (+44%), 104 - Reihen- und Terassenhaus (+47%), 106 – Öffentliche Gebäude (+42%), 123 Parkplatz (+41%), 124 – Bahn (+38%), 125 – Flugplatz (+67%), 141 – Energieversorgung (+39%), 142 – Abwasserreinigung (+52%), 145 – Abbau (+39%), 146 – Baustellen (+64%), 161 - Öffentl. Parkanlagen (40%), 162 – Sportanlagen (+44%), 163 – Golfplätze (+40%), 164 – Campingplätze (+60%), 165 – Schrebergärten (+38%), 166 – Friedhöfe (+36%). Verschlechterungen traten kaum auf.

4.3.8 Land Use AREA4 mit Zusatzdaten

Unter Zuhilfenahme der Zusatzdaten ist im AREA4-LU-Modell noch einmal eine deutliche Steigerung der Gesamtgenauigkeit von 78,1% (ohne Zusatzdaten) auf 84,20% zu beobachten. Der Abstand zur immer noch leicht besseren Gesamtgenauigkeit des LC-Modells mit Zusatzdaten (88,51%) verringert sich hier. Weiterhin sind es vor allem die grösseren, homogenen Klassen die sehr gute Klassifikationsresultate erreichen. In diesem Modell können jedoch auch viele sehr kleine Klassen mit einer Prävalenz <0,05% mit einer Genauigkeit von >70% erkannt werden (z.B. Parkplätze, Golfplätze, Schrebergärten oder Friedhöfe). Im Training befanden 665‘401 Samples und ebenso viele in der Prädiktion. Der Anteil des Höhenmodells, sowie der Anteil FCIR-CNN-Resultate an den Variablengewichten ist im LU-Modell wesentlich höher als im LC-Modell. Die RGB-CNN-Resultate nehmen im Vergleich an Wichtigkeit ab.

A4LU-Sat	101 - Industrie- und Gewerbe >1ha	102 - Industrie- und Gewerbe <1ha	103 - Ein- und Zweifamilienhaus	104 - Reihen- und Terassenhaus	105 - Mehrfamilienhaus	106 - Öffentliche Gebäude	107 - Landwirtschaftl. Gebäude	108 - Nicht spez. Gebäude	121 - Autobahn	122 - Strassen	123 - Parkplatz	124 - Bahn	125 - Flugplatz	141 - Energieversorgung	142 - Abwasserreinigung	143 - Ver- und Entsorgung	144 - Deponien	145 - Abbau	146 - Baustellen	147 - Bau- und Siedlungsbrachen	161 - Öffentl. Parkanlagen	162 - Sportanlagen	163 - Golfplätze	164 - Campingplätze	165 - Schrebergärten	166 - Friedhöfe	201 - Obstbau	202 - Rebbau	203 - Gartenbau	221 - Ackerland	222 - Naturwiesen	223 - Heimweiden	241 - Alpwiesen	242 - Alp- und Juraweiden	243 - Schafalpen	301 - Waldbestände	302 - Aufforstungen	303 - Holzschläge	304 - Waldschäden	401 - Seen	402 - Flüsse, Bäche	403 - Hochwasserverbauungen	421 - Keine Nutzung	422 - Lawinen- und Steinschlagverbauungen	423 - Alpine Sportinfrastruktur	424 - Landschaftseingriffe	Prediction
101 - Industrie- und Gewerbe >1ha	74%	13%	0%	0%	0%	5%	1%	2%	0%	0%	5%	2%	1%	17%	10%	13%	2%	2%	3%	7%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	927
102 - Industrie- und Gewerbe <1ha	3%	38%	0%	0%	0%	1%	0%	3%	0%	0%	1%	0%	0%	0%	3%	2%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	199
103 - Ein- und Zweifamilienhaus	2%	13%	88%	40%	28%	24%	15%	47%	1%	3%	5%	0%	0%	8%	32%	12%	0%	0%	9%	16%	41%	6%	0%	35%	22%	18%	6%	1%	3%	0%	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	1%	4%	0%	2%	0%	3%	1	16706
104 - Reihen- und Terassenhaus	0%	0%	0%	42%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	138
105 - Mehrfamilienhaus	1%	1%	2%	10%	61%	5%	1%	11%	0%	0%	4%	0%	0%	2%	0%	0%	0%	0%	0%	2%	3%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2203
106 - Öffentliche Gebäude	3%	1%	0%	0%	1%	45%	0%	2%	0%	0%	2%	0%	0%	2%	3%	2%	0%	0%	0%	0%	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	387
107 - Landwirtschaftl. Gebäude	2%	14%	2%	0%	1%	3%	66%	5%	0%	0%	0%	0%	1%	4%	13%	3%	1%	0%	0%	2%	0%	1%	0%	0%	0%	0%	0%	0%	3%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2806
108 - Nicht spez. Gebäude	1%	3%	0%	0%	1%	2%	0%	21%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	245
121 - Autobahn	0%	0%	0%	0%	0%	0%	0%	0%	78%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	273
122 - Strassen	4%	6%	3%	6%	3%	4%	4%	3%	18%	78%	9%	7%	5%	10%	6%	7%	4%	4%	8%	5%	7%	1%	0%	4%	4%	10%	1%	1%	0%	0%	1%	1%	0%	0%	0%	0%	12%	1%	0%	0%	2%	6%	0%	0%	0%	0%	0%	12955
123 - Parkplatz	1%	0%	0%	0%	1%	0%	0%	0%	0%	66%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	274
124 - Bahn	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	73%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	412
125 - Flugplatz	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	38%	0%	0%	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	35
141 - Energieversorgung	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	17%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	10
142 - Abwasserreinigung	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	16%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	6
143 - Ver- und Entsorgung	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0
144 - Deponien	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	28%	12%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	65
145 - Abbau	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	9%	48%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3%	111
146 - Baustellen	1%	1%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	2%	1%	2%	54%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3%	171
147 - Bau- und Siedlungsbrachen	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	5%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	8
161 - Öffentl. Parkanlagen	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	15%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	58
162 - Sportanlagen	1%	0%	0%	0%	0%	2%	0%	0%	0%	0%	1%	0%	1%	0%	0%	0%	0%	1%	0%	1%	4%	69%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	375
163 - Golfplätze	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	63%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	96
164 - Campingplätze	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	31%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	18
165 - Schrebergärten	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	2%	59%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	48
166 - Friedhöfe	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	63%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	54
201 - Obstbau	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	56%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1830
202 - Rebbau	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	820
203 - Gartenbau	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	42%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	131
221 - Ackerland	1%	0%	0%	0%	0%	1%	1%	0%	2%	4%	1%	3%	38%	2%	3%	5%	4%	6%	3%	7%	1%	6%	0%	0%	3%	0%	4%	2%	32%	81%	12%	7%	0%	0%	0%	0%	0%	0%	0%	0%	3%	10%	0%	0%	0%	40%	1	121936
222 - Naturwiesen	1%	2%	2%	0%	1%	1%	3%	2%	1%	3%	2%	4%	14%	8%	3%	12%	6%	2%	7%	13%	7%	3%	18%	8%	5%	0%	9%	0%	3%	7%	70%	22%	16%	5%	0%	0%	0%	0%	0%	0%	6%	12%	0%	0%	0%	20%	1	73633
223 - Heimweiden	1%	3%	2%	0%	1%	3%	6%	2%	2%	2%	1%	1%	1%	8%	0%	17%	6%	4%	6%	21%	7%	8%	3%	0%	5%	1%	23%	0%	9%	2%	11%	59%	1%	4%	0%	0%	8%	1%	0%	0%	5%	0%	0%	0%	0%	14%	1	44672
241 - Alpwiesen	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	6
242 - Alp- und Juraweiden	0%	0%	0%	0%	0%	1%	2%	1%	0%	2%	1%	1%	0%	4%	3%	5%	0%	0%	1%	3%	0%	0%	9%	0%	0%	0%	0%	0%	0%	0%	4%	7%	25%	81%	55%	1%	4%	2%	9%	0%	3%	4%	7%	19%	27%	9%	1	78586
243 - Schafalpen	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	26
301 - Waldbestände	0%	1%	1%	0%	1%	2%	1%	2%	3%	7%	1%	6%	0%	4%	6%	10%	11%	3%	1%	6%	8%	1%	3%	17%	0%	6%	0%	1%	5%	0%	2%	3%	4%	5%	1%	37%	54%	73%	79%	0%	19%	33%	4%	37%	8%	0%	1	224076
302 - Aufforstungen	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0
303 - Holzschläge	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	1%	0%	0%	0%	2%	4%	1%	0%	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	12%	23%	7%	0%	0%	2%	0%	2%	0%	0%	1	3078
304 - Waldschäden	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	7
401 - Seen	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%	1%	0%	0%	0%	0%	0%	1	20977
402 - Flüsse, Bäche	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	6%	0%	2%	1%	1%	0%	0%	2%	0%	2%	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	51%	18%	0%	0%	0%	0%	0%	1950
403 - Hochwasserverbauungen	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0
421 - Keine Nutzung	1%	0%	0%	0%	0%	0%	0%	0%	1%	0%	1%	1%	0%	4%	0%	5%	23%	15%	2%	5%	0%	0%	1%	2%	1%	0%	0%	1%	0%	0%	0%	0%	0%	3%	43%	1%	12%	0%	3%	1%	8%	10%	48%	37%	15%	6%	1	55091
422 - Lawinen- und Steinschlagverbauungen	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%	1	1
423 - Alpine Sportinfrastruktur	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0
424 - Landschaftseingriffe	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3%	1	1
Reference	936	275	14870	218	2833	535	3187	537	337	13299	330	455	74	48	31	60	125	168	213	121	231	402	130	52	73	71	2270	823	207	120611	68909	50108	712	76937	1652	111161	26	9276	674	21199	3070	49	57015	43	13	35	665401

Abbildung 4.34. Confusion Matrix (normalisiert) für Land Use AREA4 - Im Training: CNN-Klassifikationsergebnisse (RGB und FCIR getrennt), Satelliten-Zeitreihen und alle verfügbaren Zusatzdaten (Höhenmodell, Vegetationshöhen, Amtl. Vermessung).

Problematisch bleibt der bereits auch bei allen anderen Modellen beobachtete Trend, dass kleine Klassen mit hoher Wechselbarkeit einer grossen Klasse besonders häufig fehlklassifiziert werden. Möglichkeiten, diese Fälle besser zu klassifizieren, sind unter Kapitel 5.3 beschrieben.

Genauigkeit A4LU-Full (F1)	Sensitivity	Specificity	Pos. Pred.	Neg. Pred.	Precision	Recall	F1	Prevalence	Det. Rate	Det. Preval.	Bal. Accuracy	Samples	Correct Samples
101 - Industrie- und Gewerbe >1ha	73.6%	100.0%	74.3%	100.0%	74.3%	73.6%	74.0%	0.1%	0.1%	0.1%	73.6%	1'878	1'389
102 - Industrie- und Gewerbe <1ha	37.8%	100.0%	52.3%	100.0%	52.3%	37.8%	43.9%	0.0%	0.0%	0.0%	37.8%	552	242
103 - Ein- und Zweifamilienhaus	87.9%	99.4%	78.2%	99.7%	78.2%	87.9%	82.7%	2.2%	2.0%	2.5%	87.3%	29'828	24'682
104 - Reihen- und Terassenhaus	42.2%	100.0%	66.7%	100.0%	66.7%	42.2%	51.7%	0.0%	0.0%	0.0%	42.2%	437	226
105 - Mehrfamilienhaus	60.5%	99.9%	77.8%	99.8%	77.8%	60.5%	68.1%	0.4%	0.3%	0.3%	60.4%	5'683	3'868
106 - Öffentliche Gebäude	45.2%	100.0%	62.5%	100.0%	62.5%	45.2%	52.5%	0.1%	0.0%	0.1%	45.2%	1'073	563
107 - Landwirtschaftl. Gebäude	65.8%	99.9%	74.7%	99.8%	74.7%	65.8%	69.9%	0.5%	0.3%	0.4%	65.7%	6'393	4'472
108 - Nicht spez. Gebäude	20.7%	100.0%	45.3%	99.9%	45.3%	20.7%	28.4%	0.1%	0.0%	0.0%	20.7%	1'077	306
121 - Autobahn	70.0%	100.0%	86.4%	100.0%	86.4%	70.0%	77.4%	0.1%	0.0%	0.0%	70.0%	676	523
122 - Strassen	77.6%	99.6%	79.6%	99.5%	79.6%	77.6%	78.6%	2.0%	1.6%	1.9%	77.2%	26'677	20'964
123 - Parkplatz	66.1%	100.0%	79.6%	100.0%	79.6%	66.1%	72.2%	0.0%	0.0%	0.0%	66.1%	662	478
124 - Bahn	72.7%	100.0%	80.3%	100.0%	80.3%	72.7%	76.4%	0.1%	0.0%	0.1%	72.7%	913	697
125 - Flugplatz	37.8%	100.0%	80.0%	100.0%	80.0%	37.8%	51.4%	0.0%	0.0%	0.0%	37.8%	148	76
141 - Energieversorgung	16.7%	100.0%	80.0%	100.0%	80.0%	16.7%	27.6%	0.0%	0.0%	0.0%	16.7%	96	27
142 - Abwasserreinigung	16.1%	100.0%	83.3%	100.0%	83.3%	16.1%	27.0%	0.0%	0.0%	0.0%	16.1%	62	17
143 - Ver- und Entsorgung	0.0%	100.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	120	-
144 - Deponien	28.0%	100.0%	53.8%	100.0%	53.8%	28.0%	36.8%	0.0%	0.0%	0.0%	28.0%	251	92
145 - Abbau	47.6%	100.0%	72.1%	100.0%	72.1%	47.6%	57.3%	0.0%	0.0%	0.0%	47.6%	337	193
146 - Baustellen	54.5%	100.0%	67.8%	100.0%	67.8%	54.5%	60.4%	0.0%	0.0%	0.0%	54.5%	427	258
147 - Bau- und Siedlungsbrachen	5.0%	100.0%	75.0%	100.0%	75.0%	5.0%	9.3%	0.0%	0.0%	0.0%	5.0%	243	23
161 - Öffentl. Parkanlagen	14.7%	100.0%	58.6%	100.0%	58.6%	14.7%	23.5%	0.0%	0.0%	0.0%	14.7%	463	109
162 - Sportanlagen	69.4%	100.0%	74.4%	100.0%	74.4%	69.4%	71.8%	0.1%	0.0%	0.1%	69.4%	806	579
163 - Golfplätze	63.1%	100.0%	85.4%	100.0%	85.4%	63.1%	72.6%	0.0%	0.0%	0.0%	63.1%	261	189
164 - Campingplätze	30.8%	100.0%	88.9%	100.0%	88.9%	30.8%	45.7%	0.0%	0.0%	0.0%	30.8%	104	48
165 - Schrebergärten	58.9%	100.0%	89.6%	100.0%	89.6%	58.9%	71.1%	0.0%	0.0%	0.0%	58.9%	146	104
166 - Friedhöfe	63.4%	100.0%	83.3%	100.0%	83.3%	63.4%	72.0%	0.0%	0.0%	0.0%	63.4%	142	103
201 - Obstbau	56.3%	99.9%	69.8%	99.9%	69.8%	56.3%	62.3%	0.3%	0.2%	0.3%	56.2%	4'553	2'839
202 - Rebbau	94.2%	100.0%	94.5%	100.0%	94.5%	94.2%	94.3%	0.1%	0.1%	0.1%	94.2%	1'651	1'557
203 - Gartenbau	41.5%	100.0%	65.6%	100.0%	65.6%	41.5%	50.9%	0.0%	0.0%	0.0%	41.5%	415	211
221 - Ackerland	90.6%	97.7%	89.6%	97.9%	89.6%	90.6%	90.1%	18.1%	16.4%	18.3%	88.3%	241'939	218'002
222 - Naturwiesen	69.9%	95.7%	65.4%	96.5%	65.4%	69.9%	67.6%	10.4%	7.2%	11.1%	65.6%	138'227	93'393
223 - Heimweiden	59.3%	97.6%	66.6%	96.7%	66.6%	59.3%	62.7%	7.5%	4.5%	6.7%	56.9%	100'514	63'070
241 - Alpwiesen	0.8%	100.0%	100.0%	99.9%	100.0%	0.8%	1.7%	0.1%	0.0%	0.0%	0.8%	1'428	24
242 - Alp- und Juraweiden	83.2%	97.5%	81.5%	97.8%	81.5%	83.2%	82.3%	11.6%	9.6%	11.8%	80.7%	154'331	127'059
243 - Schafalpen	1.2%	100.0%	76.9%	99.8%	76.9%	1.2%	2.4%	0.2%	0.0%	0.0%	1.2%	3'314	79
301 - Waldbestände	97.3%	96.1%	92.1%	98.7%	92.1%	97.3%	94.7%	31.9%	31.0%	33.7%	93.4%	425'583	402'829
302 - Aufforstungen	0.0%	100.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	52	-
303 - Holzschläge	23.0%	99.9%	69.3%	98.9%	69.3%	23.0%	34.5%	1.4%	0.3%	0.5%	22.9%	18'607	6'425
304 - Waldschäden	1.0%	100.0%	100.0%	99.9%	100.0%	1.0%	2.1%	0.1%	0.0%	0.0%	1.0%	1'352	28
401 - Seen	98.5%	100.0%	99.5%	99.9%	99.5%	98.5%	99.0%	3.2%	3.1%	3.2%	98.5%	42'524	42'096
402 - Flüsse, Bäche	51.2%	99.9%	80.6%	99.8%	80.6%	51.2%	62.6%	0.5%	0.2%	0.3%	51.1%	6'158	3'854
403 - Hochwasserverbauungen	0.0%	100.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	98	-
421 - Keine Nutzung	87.9%	99.2%	90.9%	98.9%	90.9%	87.9%	89.4%	8.6%	7.5%	8.3%	87.0%	114'369	102'206
422 - Lawinen- und Steinschlagvert	2.3%	100.0%	100.0%	100.0%	100.0%	2.3%	4.5%	0.0%	0.0%	0.0%	2.3%	86	4
423 - Alpine Sportinfrastruktur	0.0%	100.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	26	-
424 - Landschaftseingriffe	2.9%	100.0%	100.0%	100.0%	100.0%	2.9%	5.6%	0.0%	0.0%	0.0%	2.9%	70	4
												1'334'756	1'123'909

Abbildung 4.35. Genauigkeitsmetriken für Land Cover AREA4 - Im Training: CNN-Klassifikationsergebnisse (RGB und FCIR getrennt), Satelliten-Zeitreihen und alle verfügbaren Zusatzdaten (Höhenmodell, Vegetationshöhen, Amtl. Vermessung); rechts dargestellt die bei derzeitiger Datenlage maximal vorhersagbare Anzahl Samples mit diesem Modell. Overall Accuracy: 84,20%.

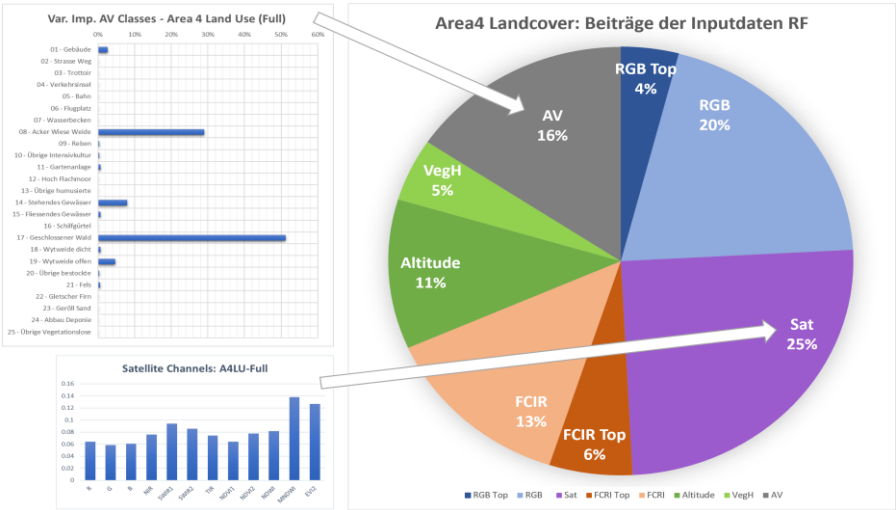


Abbildung 4.36. Variablengewichte für Land Use AREA4 - Im Training: CNN-Klassifikationsergebnisse (RGB und FCIR getrennt), Satelliten-Zeitreihen und alle verfügbaren Zusatzdaten (Höhenmodell, Vegetationshöhen, Amtl. Vermessung).

4.3.9 Change Detection Land Use

Auch hier gilt, dass beide zusätzlichen Change Detection-Ansätze sich von einer Anwendung des AREA4-LU-Modells mit Zusatzdaten in der Grundannahme unterscheiden, dass für sie die Ground-Truth-Werte für AREA3 (ausser natürlich für den Trainigsvorgang und die Evaluation) nicht notwendigerweise bekannt sein müssen.

CD LU: Tabellarischer Ansatz

Die Overall-Accuracy des tabellarischen LU-Ansatzes ist vor allem im Vergleich mit den sehr guten Resultaten der AREA4-LU-Klassifikation mit Zusatzdaten (84,20%) mit einem Wert von 60,41% relativ niedrig ausgefallen. Dies ist vor allem auf die durchwachsenen Resultate der AREA3-Klassifikation im Hinblick auf kleine, heterogene, verwechselbare Klassen zurückzuführen. Lediglich Waldbestände und Seen erreichen mit diesem Ansatz Genauigkeiten >80%.

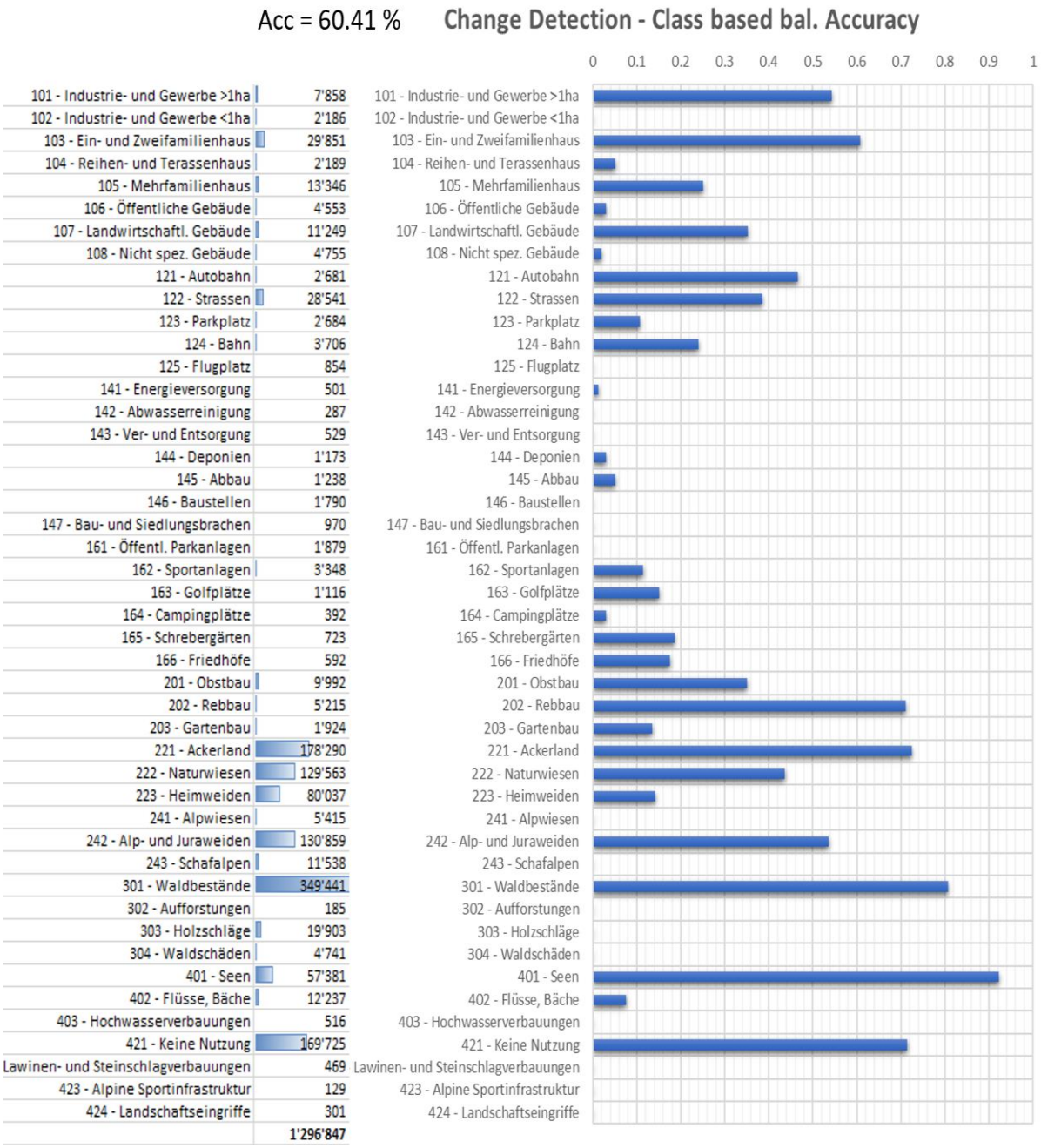


Abbildung 4.37. Change Detection mittels tabellarischem Ansatz; im Vergleich die Land Use Klassifikationsresultate AREA3 vs. AREA4 aus den Modellen CNN (RGB) + Satellitenzeitreihen.

CD LU: Dichotomer RF-Ansatz

Der Change-Detection-Ansatz mit eigenem RF-Modell schneidet wie bei CD-LC mit 73.5% (F1) etwas besser ab als der tabellarische Ansatz, unterliegt jedoch weiterhin deutlich dem direkten Ansatz unter Verwendung des LU-AREA4-Modells mit Zusatzdaten (84,20%). Es fällt jedoch auf, dass die Variablengewichte zwischen AREA3 und AREA4 deutlich ausgeglichener als im CD-LC-Modell sind. Die Satellitenzeitreihen überwiegen dabei die RGB-CNN-Resultate leicht.

	CD LU RF
Sensitivity	68.2%
Specificity	82.8%
Pos Pred Value	79.6%
Neg Pred Value	72.6%
Precision	79.6%
Recall	68.2%
F1	73.5%
Prevalence	49.6%
Detection Rate	33.8%
Detection Prevalence	42.5%
Balanced Accuracy	75.5%
Samples Change	10606
Samples NoChange	10794

Abbildung 4.38. Change Detection (Land Use) mittels direktem Machine-Learning-Ansatz; Dieses Random-Forest-Modell wurde auf die dichotome Entscheidung Change / No Change trainiert; im Training: CNN (RGB) jeweils AREA3 und 4, sowie Satellitenzeitreihen jeweils AREA3 und 4.

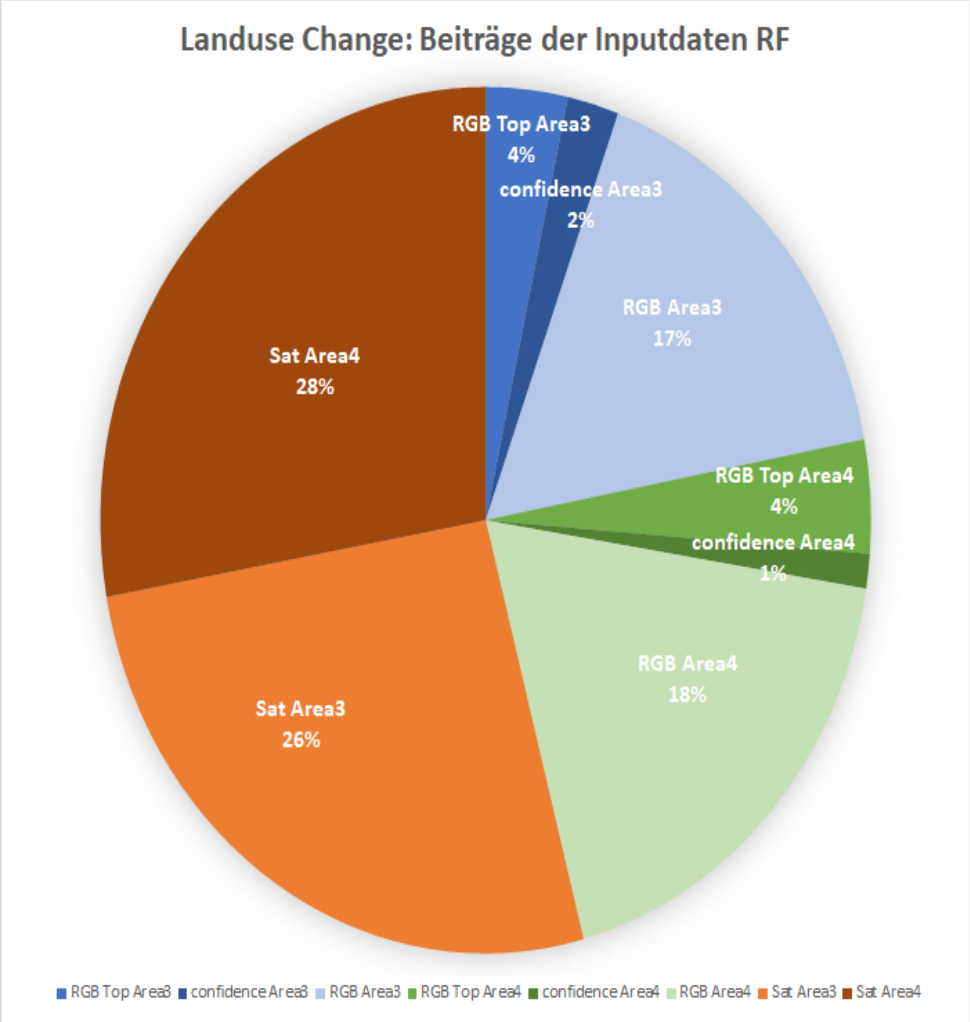


Abbildung 4.39. Verteilung der Variablengewichte im dichotomen Change Detection Random Forest Modell (Land Use).

4.3.10 Kombinierte LU/LC-Klassifikation

Für die kombinierte Random-Forest-Klassifikation von Land Cover und Land Use wurden für die Input-Daten AREA4-CNN-Klassifikationsresultate für RGB und FCIR jeweils für Land Use und Land Cover mit allen verfügbaren Zusatzdaten verkettet. Die ursprüngliche Anzahl Klassen in der Ground Truth stieg auf über 300 an,

wobei jedoch nur die 58 grössten Klassen verwertet werden konnten. Als Ausschlusskriterium wurde eine Prävalenz von <0.1% der Samples des Gesamtdatensatzes verwendet. Dies entspricht einer Anzahl trainierbarer Samples von mindestens n=60. Im Training befanden sich 61‘805 Samples, in der Prädiktion 60‘000 Samples. Dies entsprach jeweils ca. 5% des zur Verfügung stehenden Datensatzes von max. 1‘218‘060 Samples.

Die Gesamtgenauigkeit des kombinierten Modells erreichte mit 77,20% einen Wert der unterhalb der Einzelwerte für LC von 88,4% und LU von 84,2% in der jeweiligen Klassifikation mit Zusatzdaten liegt. Typische Probleme wie kleine heterogene Klassen von grossen homogenen Klassen zu unterscheiden bleiben jedoch trotzdem bestehen. Viele Klassenkombinationen können aufgrund ihrer extrem niedrigen Prävalenz von diesem Modell darüber hinaus systematisch nicht erreicht werden. Für die kombinierte LU/LC-Klassifikation wurde kein gesonderter dichotomer Change-Detection-Durchlauf ausgeführt, da die Klassenbesetzung innerhalb der „Change“-Stichproben prohibitiv niedrig ausgefallen wäre.

	c11c101c	c11c103c	c11c105c	c11c107c	c11c122c	c12c103c	c12c105c	c12c107c	c15c103c	c15c105c	c21c103c	c21c122c	c21c221c	c21c222c	c21c223c	c21c241c	c21c242c	c21c243c	c21c402c	c21c421c	c31c303c	c31c421c	c32c242c	c32c303c	c32c421c	c33c201c	c34c202c	c41c301c	c41c302c	c42c301c	c43c301c	c44c223c	c44c242c	c44c301c	c44c303c	c45c301c	c46c221c	c46c222c	c46c223c	c46c242c	c46c402c	c46c421c	c47c103c	c47c201c	c47c222c	c47c223c	c47c242c	c47c303c	c47c421c	c51c421c	c52c402c	c52c421c	c53c242c	c53c421c	c61c401c	c61c402c	c62c421c	c63c242c
c11c101c	76%	0%	2%	1%	0%	0%	0%	0%	0%	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c11c103c	0%	57%	20%	12%	2%	24%	4%	3%	24%	8%	10%	3%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	1%	0%	11%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%
c11c105c	0%	1%	42%	0%	0%	2%	4%	0%	1%	3%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c11c107c	4%	2%	2%	59%	0%	2%	0%	29%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c11c122c	12%	15%	13%	8%	83%	1%	8%	0%	1%	2%	0%	36%	0%	1%	1%	2%	0%	0%	7%	0%	0%	0%	0%	0%	3%	1%	0%	3%	0%	0%	0%	0%	0%	1%	0%	0%	0%	1%	1%	0%	6%	0%	7%	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3%	0%	0%
c12c103c	4%	7%	2%	1%	0%	67%	35%	17%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c12c105c	0%	0%	0%	0%	0%	1%	38%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c12c107c	0%	0%	0%	4%	0%	2%	4%	34%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c15c103c	4%	9%	0%	0%	0%	0%	0%	3%	57%	17%	30%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	21%	1%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c15c105c	0%	1%	16%	0%	0%	0%	4%	0%	4%	58%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c21c103c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c21c122c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c21c221c	0%	0%	0%	1%	2%	0%	0%	0%	0%	0%	20%	81%	14%	8%	0%	0%	0%	30%	0%	0%	0%	0%	0%	0%	0%	0%	10%	1%	0%	4%	0%	0%	0%	0%	0%	0%	32%	7%	1%	0%	3%	0%	0%	1%	3%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%
c21c222c	0%	2%	0%	5%	2%	0%	4%	0%	7%	6%	30%	17%	7%	68%	25%	12%	7%	0%	26%	0%	0%	1%	0%	0%	1%	1%	10%	0%	0%	8%	2%	0%	0%	0%	0%	0%	32%	30%	15%	0%	9%	0%	4%	16%	34%	12%	1%	0%	1%	0%	4%	0%	0%	0%	0%	3%	0%	0%
c21c223c	0%	2%	2%	9%	1%	1%	0%	11%	2%	3%	20%	4%	2%	11%	58%	0%	5%	0%	11%	1%	2%	1%	1%	0%	1%	0%	0%	0%	12%	2%	9%	0%	0%	0%	0%	0%	14%	24%	44%	6%	9%	8%	18%	30%	35%	51%	2%	2%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%
c21c241c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c21c242c	0%	1%	0%	1%	2%	0%	0%	0%	1%	0%	0%	4%	0%	5%	6%	80%	87%	78%	0%	42%	0%	43%	82%	3%	43%	0%	0%	0%	4%	1%	0%	11%	2%	0%	16%	0%	7%	2%	35%	0%	4%	0%	0%	20%	14%	49%	2%	17%	0%	7%	1%	47%	10%	0%	5%	0%	100%
c21c243c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c21c402c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	4%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c21c421c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	10%	0%	20%	0%	3%	1%	0%	3%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	1%	0%	0%	2%	5%	0%	0%	0%	0%	0%
c31c303c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	28%	1%	1%	45%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	1%	0%	0%	0%	0%	0%	0%	4%	0%	0%	0%	1%	0%	13%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%
c31c421c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	2%	0%	15%	4%	0%	7%	0%	0%	0%	0%	0%	0%	0%	1%	1%	0%	12%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	4%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%
c32c242c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	3%	5%	0%	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c32c303c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c32c421c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c33c201c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	4%	0%	0%	0%	0%	0%	0%	81%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c34c202c	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	95%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c41c301c	0%	1%	0%	0%	6%	0%	0%	0%	0%	2%	0%	11%	0%	1%	1%	0%	1%	0%	15%	2%	69%	16%	2%	38%	7%	0%	0%	99%	69%	90%	82%	42%	74%	94%	47%	9%	10%	21%	38%	24%	25%	14%	1%	1%	8%	12%	72%	26%	1%	15%	1%	1%	2%	0%	11%	0%	0%
c42c301c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c43c301c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	1%	1%	0%	3%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%
c44c223c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c44c242c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	1%	0%	0%	0%	0%	0%	4%	0%	5%	27%	2%	1%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	1%	7%	0%	4%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c44c301c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	1%	1%	0%	7%	6%	0%	0%	0%	0%	0%	0%	0%	3%	11%	2%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	8%	3%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%
c44c303c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c45c301c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	16%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c46c221c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c46c222c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	5%	8%	4%	0%	1%	13%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c46c223c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	5%	8%	5%	0%	6%	8%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%
c46c242c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	4%	0%	4%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c46c402c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	5%	2%	1%	0%	25%	13%	4%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	4%	0%
c46c421c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	1%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c47c103c	0%	1%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	14%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c47c201c	0%	0%	0%	0%	0%	0%	0%	7%	0%	10%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	5%	0%	0%	0%	0%	0%	0%	1%	0%	1%	0%	7%	50%	1%	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c47c222c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c47c223c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c47c242c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	6%	1%	1%	0%	0%	1%	3%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	15%	2%	0%	0%	0%	1%	0%	14%	0%	13%	0%	0%	3%	7%	26%	0%	14%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%
c47c303c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c47c421c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	4%	0%	0%	0%	0%	0%	0%	5%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c51c421c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	7%	0%	6%	0%	0%	5%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	12%	83%	0%	9%	1%	18%	0%	2%	3%	0%
c52c402c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c52c421c	0%	0%	0%	0%	0%	0%	3%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	4%	0%	3%	0%	0%	5%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	4%	0%	0%	0%	0%	0%	0%	4%	9%	41%	81%	8%	13%	0%	3%	3%	0%
c53c242c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	11%	2%	0%	0%	0%	0%	0%
c53c421c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	7%	0%	21%	0%	6%	1%	0%	15%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3%	0%	0%	0%	0%	0%	0%	0%	0%	0%	8%	4%	7%	6%	31%	47%	0%	0%	0%	0%	0%	0%	0%	0%	0%
c61c401c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	99%	0%	0%	0%	0%
c61c402c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	4%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	7%	0%	0%	0%	0%	0%	0%	0%	0%	0%	26%	0%	0%	0%	0%	82%	0%	0%	0%
c62c421c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	2%	0%	0%	0%	0%	93%	0%
c63c242c	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Abbildung 4.40. Normalisierte Confusion Matrix für ein Modell mit 60’000 Samples in der Prädiktion für kombinierte Klassen von Land Cover und Land Use. Overall Accuracy = 77,20%

	Sensitivity	Specificity	Pos Pred Value	Neg Pred Value	Precision	Recall	F1	Prevalence	Detection Rate	Detection Prevalen	Balanced Accuracy	Samples	Correct	
c11c101c	76.00%	99.99%	76.00%	99.99%	76.00%	76.00%	76.00%	0.04%	0.03%	0.04%	75.99%	25	19	
c11c103c	57.32%	99.74%	53.64%	99.77%	53.64%	57.32%	55.42%	0.53%	0.30%	0.56%	57.06%	316	180	
c11c105c	42.22%	99.98%	55.88%	99.96%	55.88%	42.22%	48.10%	0.07%	0.03%	0.06%	42.20%	44	19	
c11c107c	58.56%	99.94%	64.36%	99.92%	64.36%	58.56%	61.32%	0.18%	0.11%	0.17%	58.50%	109	64	
c11c122c	82.69%	99.54%	77.51%	99.67%	77.51%	82.69%	80.02%	1.87%	1.55%	1.99%	82.23%	1121	922	
c12c103c	66.67%	99.93%	70.34%	99.92%	70.34%	66.67%	68.46%	0.25%	0.17%	0.24%	66.60%	151	100	
c12c105c	38.46%	99.99%	66.67%	99.97%	66.67%	38.46%	48.78%	0.04%	0.02%	0.02%	38.45%	26	10	
c12c107c	34.29%	99.98%	54.55%	99.96%	54.55%	34.29%	42.11%	0.06%	0.02%	0.04%	34.27%	34	12	
c15c103c	56.84%	99.86%	61.86%	99.83%	61.86%	56.84%	59.24%	0.38%	0.22%	0.35%	56.70%	231	131	
c15c105c	58.46%	99.96%	58.46%	99.96%	58.46%	58.46%	58.46%	0.11%	0.06%	0.11%	58.42%	64	37	
c21c103c	0.00%	100.00%	0.00%	99.97%	0.00%	0.00%	0.00%	0.03%	0.00%	0.00%	0.00%	20	0	
c21c122c	1.04%	100.00%	100.00%	99.84%	100.00%	1.04%	2.06%	0.16%	0.00%	0.00%	1.04%	95	1	
c21c221c	90.96%	97.14%	88.63%	97.77%	88.63%	90.96%	89.78%	19.71%	17.93%	20.23%	88.09%	11825	10417	
c21c222c	67.94%	95.11%	62.47%	96.12%	62.47%	67.94%	65.09%	10.71%	7.28%	11.65%	63.05%	6425	4051	
c21c223c	58.17%	97.28%	63.32%	96.64%	63.32%	58.17%	60.63%	7.48%	4.35%	6.87%	55.44%	4490	2489	
c21c241c	0.00%	100.00%	0.00%	99.92%	0.00%	0.00%	0.00%	0.08%	0.00%	0.00%	0.00%	48	0	
c21c242c	84.71%	96.66%	72.64%	98.37%	72.64%	84.71%	78.21%	9.48%	8.03%	11.05%	81.37%	5686	4627	
c21c243c	0.62%	100.00%	50.00%	99.74%	50.00%	0.62%	1.23%	0.26%	0.00%	0.00%	0.62%	159	1	
c21c402c	3.70%	100.00%	50.00%	99.96%	50.00%	3.70%	6.90%	0.04%	0.00%	0.00%	3.70%	27	1	
c21c421c	19.86%	99.86%	50.59%	99.43%	50.59%	19.86%	28.52%	0.71%	0.14%	0.28%	19.72%	427	84	
c31c303c	27.80%	99.90%	53.17%	99.71%	53.17%	27.80%	36.51%	0.40%	0.11%	0.21%	27.70%	237	66	
c31c421c	14.61%	99.90%	30.59%	99.75%	30.59%	14.61%	19.77%	0.29%	0.04%	0.14%	14.51%	175	25	
c32c242c	4.71%	99.98%	40.91%	99.70%	40.91%	4.71%	8.45%	0.31%	0.01%	0.04%	4.69%	188	9	
c32c303c	0.00%	100.00%	0.00%	99.95%	0.00%	0.00%	0.00%	0.05%	0.00%	0.00%	0.00%	29	0	
c32c421c	1.01%	100.00%	33.33%	99.84%	33.33%	1.01%	1.96%	0.16%	0.00%	0.00%	1.01%	98	1	
c33c201c	80.95%	100.00%	89.47%	99.99%	89.47%	80.95%	85.00%	0.03%	0.03%	0.03%	80.95%	21	17	
c34c202c	94.94%	99.99%	94.94%	99.99%	94.94%	94.94%	94.94%	0.13%	0.12%	0.13%	94.93%	78	74	
c41c301c	98.88%	95.00%	90.16%	99.46%	90.16%	98.88%	94.32%	31.65%	31.29%	34.71%	93.89%	18988	17827	
c42c301c	0.00%	100.00%	0.00%	99.96%	0.00%	0.00%	0.00%	0.04%	0.00%	0.00%	0.00%	26	0	
c43c301c	1.10%	99.99%	36.36%	99.41%	36.36%	1.10%	2.14%	0.59%	0.01%	0.02%	1.09%	357	4	
c44c223c	0.00%	100.00%	0.00%	99.96%	0.00%	0.00%	0.00%	0.04%	0.00%	0.00%	0.00%	22	0	
c44c242c	27.22%	99.84%	48.63%	99.61%	48.63%	27.22%	34.90%	0.54%	0.15%	0.30%	27.06%	322	87	
c44c301c	11.47%	99.88%	43.85%	99.28%	43.85%	11.47%	18.18%	0.82%	0.09%	0.21%	11.35%	490	56	
c44c303c	1.62%	99.99%	58.33%	99.30%	58.33%	1.62%	3.16%	0.71%	0.01%	0.02%	1.62%	425	7	
c45c301c	16.18%	99.99%	73.33%	99.91%	73.33%	16.18%	26.51%	0.11%	0.02%	0.02%	16.17%	67	11	
c46c221c	0.00%	100.00%	0.00%	99.96%	0.00%	0.00%	0.00%	0.04%	0.00%	0.00%	0.00%	22	0	
c46c222c	7.77%	99.98%	34.78%	99.84%	34.78%	7.77%	12.70%	0.17%	0.01%	0.04%	7.74%	101	8	
c46c223c	5.49%	99.97%	23.81%	99.86%	23.81%	5.49%	8.93%	0.15%	0.01%	0.03%	5.47%	90	5	
c46c242c	4.23%	100.00%	60.00%	99.89%	60.00%	4.23%	7.89%	0.12%	0.00%	0.01%	4.22%	70	3	
c46c402c	25.00%	99.97%	44.74%	99.92%	44.74%	25.00%	32.08%	0.11%	0.03%	0.06%	24.97%	67	17	
c46c421c	0.00%	100.00%	0.00%	99.96%	0.00%	0.00%	0.00%	0.04%	0.00%	0.00%	0.00%	24	0	
c47c103c	14.29%	99.99%	33.33%	99.96%	33.33%	14.29%	20.00%	0.05%	0.01%	0.02%	14.27%	28	4	
c47c201c	49.59%	99.93%	59.41%	99.90%	59.41%	49.59%	54.05%	0.20%	0.10%	0.17%	49.52%	119	59	
c47c222c	0.00%	100.00%	0.00%	99.88%	0.00%	0.00%	0.00%	0.12%	0.00%	0.00%	0.00%	70	0	
c47c223c	0.72%	100.00%	33.33%	99.77%	33.33%	0.72%	1.41%	0.23%	0.00%	0.00%	0.72%	137	1	
c47c242c	26.30%	99.70%	36.93%	99.51%	36.93%	26.30%	30.72%	0.66%	0.17%	0.47%	26.00%	397	103	
c47c303c	1.03%	100.00%	33.33%	99.84%	33.33%	1.03%	2.00%	0.16%	0.00%	0.00%	1.03%	96	1	
c47c421c	4.65%	99.98%	37.50%	99.80%	37.50%	4.65%	8.28%	0.21%	0.01%	0.03%	4.63%	127	6	
c51c421c	82.78%	99.50%	78.27%	99.62%	78.27%	82.78%	80.46%	2.14%	1.77%	2.26%	82.28%	1282	1055	
c52c402c	0.00%	100.00%	0.00%	99.96%	0.00%	0.00%	0.00%	0.04%	0.00%	0.00%	0.00%	27	0	
c52c421c	80.73%	99.53%	77.20%	99.62%	77.20%	80.73%	78.93%	1.94%	1.57%	2.03%	80.25%	1166	936	
c53c242c	10.76%	99.96%	43.59%	99.77%	43.59%	10.76%	17.26%	0.26%	0.03%	0.06%	10.72%	156	17	
c53c421c	47.01%	99.45%	43.37%	99.53%	43.37%	47.01%	45.12%	0.88%	0.41%	0.95%	46.47%	528	245	
c61c401c	99.33%	100.00%	99.90%	99.98%	99.90%	99.33%	99.61%	3.42%	3.40%	3.40%	99.32%	2053	2039	
c61c402c	61.90%	99.96%	73.03%	99.93%	73.03%	61.90%	67.01%	0.17%	0.11%	0.15%	61.87%	103	64	
c62c421c	93.39%	99.92%	89.57%	99.95%	89.57%	93.39%	91.44%	0.77%	0.72%	0.80%	93.31%	462	431	
c63c242c	0.00%	100.00%	0.00%	99.94%	0.00%	0.00%	0.00%	0.06%	0.00%	0.00%	0.00%	34	0	
												60000	**46342**	77.2%

Abbildung 4.41. Präzisionsmetriken für das kombinierte LU/LC-Modell.

4.3.11 Variablengewichte der Satellitenkanäle im Vergleich

Die Angaben über die internen Variablengewichte der Satellitenkanäle sind auf den jeweiligen Anteil der Satellitenzeitreihen an der Gesamtklassifikation normiert. Grössere Abweichungen von einem rechnerischen Durchschnittsanteil von 0,083 zeigen sich eher in den Modellen ohne Zusatzvariablen A3LC, A4LC, sowie A3LU und A4LU. Bei den Modellen mit Zusatzvariablen (A4LC-Full und A4LU-Full) ist die Verteilung hingegen homogener. Wichtige Beiträge liefern insbesondere die SWIR-Kanäle, sowie die Indizes, die sensitiv in Bezug auf Vegetationseigenschaften und Wasser sind (z.B. EVI oder MNDWI).

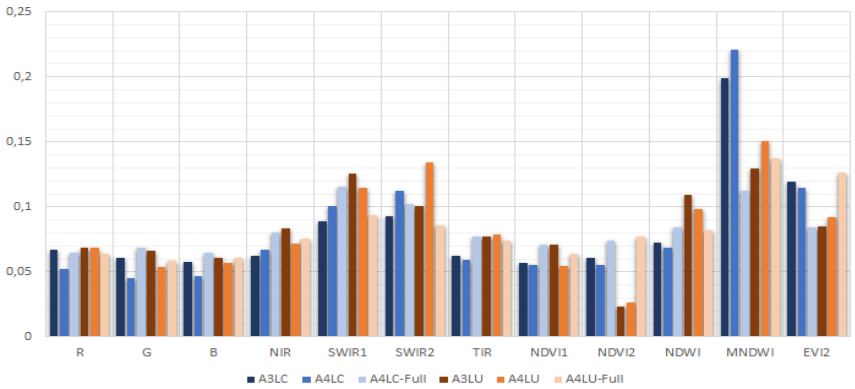


Abbildung 4.42. Variablengewichte über alle RF-Klassifikationsmodelle.

4.3.12 Variablengewichte der AV-Daten im Vergleich

Besonders deutlich fallen bei beiden Modellen die Beträge der Klassen „8 – Acker/Wiese/Weide“ und „17 – Geschlossener Wald“ auf, welche weit über dem theoretischen rechnerischen Durchschnittsbeitrag von 4% liegen. Ursächlich ist dafür vermutlich der Anteil Ground-Truth-Prävalenzen dieses Typus im Trainings-Datensatz. Eine starke Veränderung dieser Werte ist nur zu erwarten, wenn der RF mit einem symmetrischeren Datensatz aufgebaut würde, da die AV-Werte mit den starken Klassen korrellieren.. Grundsätzlich sind Random-Forest-Modelle aber in einer Umgebung mit vielen Klassen unempfindlich gegenüber assymetrischen Datensätzen. Die dargestellten Ungleichgewichte repräsentieren somit auch einen Ausschnitt der reellen Prävelanz.

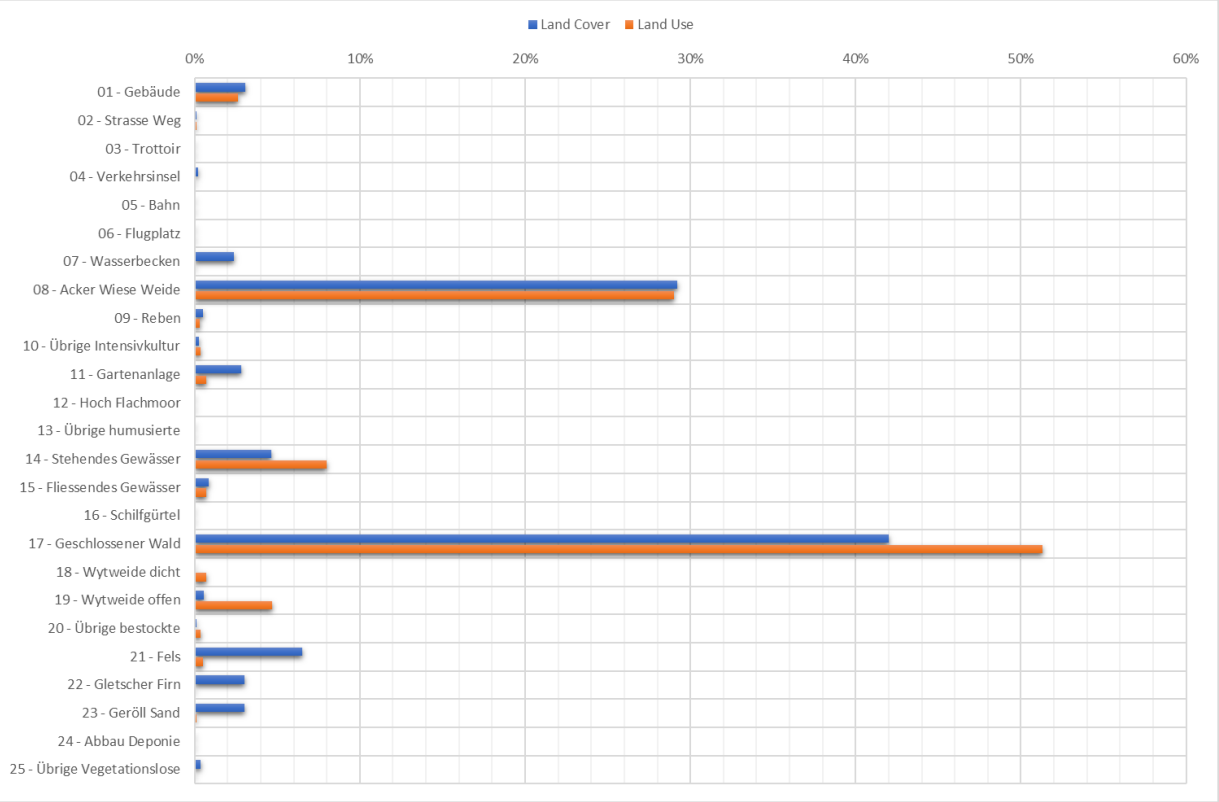


Abbildung 4.43. Variablengewichte der faktorisierten Bodenbedeckung aus der amtlichen Vermessung die Modelle für Land Use und Land Cover mit allen Zusatzdaten.

4.3.13 Trainingsdauer

Tabelle 4.15. Trainingsdauer verschiedener LC-Durchläufe in Abhängigkeit von der Anzahl trainierter Samples und dem Umfang der Inputvektoren. Hardware: 2017 Intel Xeon Octacore Desktop, 32GB RAM, SSD.

Number of Samples	LC Model	Training Time	Prediction Accuracy
400‘000	DL RGB + SAT	3-4 h	79 %
300‘000	DL RGB + FCIR + SAT + AV +DEM + CHM	3-4 h	88 %
13‘000	Change Detection RGB (x2) + SAT (x2)	<1 h	79 %

5 Empfehlungen und Optimierungspotential

Der im Rahmen des Pilot-Projektes entwickelte Prototyp AI ist vollumfänglich funktionsfähig, dokumentiert und erlaubt die eigenständige Umsetzung aller Prozessierungs- und Klassifikationsschritte von Seiten des BFS. In Hinblick auf eine effizientere Anwendung, wäre nachfolgend die Integration aller Teilmodule in einer ready-to-use Software mit intuitiver GUI empfehlenswert, da dies die Handhabung des Prototypen AI in einem operationellen Setup deutlich optimieren würde. Insbesondere bei iterativen Vorgehensweisen würde solch ein Software-Paket die Parametrisierung und Automatisierung stark vereinfachen und unterstützen. Nachfolgend wird darüberhinausgehend noch auf einzelne Komponenten eingegangen, für die ein Optimierungspotenzial gesehen wird.

5.1 Datengrundlage/Datenverwendung

Grundsätzlich kann festgehalten werden, dass die Basisdaten (= die verschiedenen SWISSIMAGE - Datensätze) sehr gut für Klassifikationsansätze basierend auf künstlicher Intelligenz geeignet sind. Dies ist insofern interessant, da mit den Orthobildern noch nicht mal die optimale Datengrundlage in der auf ImageNet vortrainierte Xception CNN-Architektur Verwendung fand. Mit der Implementierung eines Modules zur Prozessierung der Oblique-Luftbilder könnten dementsprechend für bestimmte Klassen bessere Ergebnisse erzielt werden. Klassen, bei denen ausschliesslich die exakte Punktinformation für die Klassenzuweisung herangezogen wird, könnten hingegen schlechter abgebildet werden, da durch die unterschiedlichen Blickwinkel - insbesondere in vertikal stark strukturierten Bereichen wie in urbanen Gebieten - zum Teil grosse Unterschiede in den Merkmalsräumen auftreten. In Hinblick auf die spektralen Eigenschaften der Basisdaten wäre eine Verbesserung des Informationsgehaltes in Schattenbereichen empfehlenswert. Für Stichprobenpunkte, deren Interpretation durch Schatten beeinflusst wird (z.B. bedingt durch Topografie oder Gebäudelage), könnte durch eine radiometrische Korrektur sowohl eine deutliche Verbesserung der visuellen Interpretation als auch der entsprechenden automatisierten Klassifikation im Rahmen des Prototypen AI erreicht werden. Um die komplexe und zeitintensive Prozessierung möglichst optimal anzuwenden, empfiehlt es sich, diese nur auf diejenigen Stichprobenpunkte anzuwenden, welche in einem vorgängigen Schritt im Rahmen der Schattendetektion bereits selektiert wurden. Die Schattendetektion könnte hierbei auf dem in Kapitel 2.3.1 beschriebenen Vorgehen basieren oder unter Verwendung einer räumlich hochgenauen Schattenmodellierung unter Verwendung von TLM3D/swissALTI3D umgesetzt werden.

Das Optimierungspotential in Hinblick auf die Hilfs- und Zusatzdaten betrifft vor allem die Verwendung von Satellitendaten im Klassifikationsprozess sowie die Aufbereitung des thematischen Informationsgehaltes der Geodaten des Bundes. Mit den Sentinel-Missionen der ESA stehen für AREA 4 (teilweise) und AREA5 (vollständig) eine Datenquelle zur Verfügung, die sowohl zeitlich als auch räumlich deutliche Verbesserungen im Vergleich zu den im Rahmen dieser Studie verwendeten Landsat-Daten aufweisen kann (die Ergebnisse in Kapitel 4.3.6 deuten bereits das Potential dieser Zusatzinformationen im Rahmen der Klassifikation an). So bedeuten die 10 x 10 m Pixelgrösse eine 9-fache bessere räumliche Auflösung bei einer theoretisch möglichen wöchentlichen Abdeckung der ganzen Schweiz. Zudem zeigen eine Vielzahl an Studien, dass durch eine hohe spektrale Auflösung bzw. durch die Kombination aus passiv-optischen Sensoren und aktiven Sensoren im Bereich der Mikrowellen Klassifikationen der Landnutzung/-bedeckung mit deutlich höheren Genauigkeiten durchgeführt werden können. Dementsprechend wäre eine Integration von Daten der Sentinel-1 und Sentinel-2 Missionen für die zukünftigen Erhebungen der Arealstatistik sehr empfehlenswert.

Die aus unserer Sicht wichtigste Optimierung betrifft die Verwendung der thematischen Daten des Bundes (z.B. AV oder swissTLM3D). Die Klassentiefen und Klassenbeschreibungen sind aktuell nur bedingt kompatibel mit der Klassifikationsstruktur, die in der Arealstatistik Verwendung findet. Damit kommt es zu grossen Überschneidungen zwischen verschiedenen Klassen, was wiederum die Zuweisungen im Rahmen des Klassifikationsprozesses erschwert und somit auch die mittleren klassenspezifischen Genauigkeiten verringert. Dies könnte man dahingehend verbessern, dass man für die RF-Klassifikation die Klassen auf einer übergeordneten Ebene harmonisiert und anschliessend über einen Plausibilisierungsansatz die finale individuelle Klassenzuweisung vornimmt. In diesem Kontext ist auch die (automatisierte) Berücksichtigung der zeitlichen

Komponente notwendig, um den Einfluss abweichender Aufnahme-/Erstellungszeitpunkte auf die Klassifikation so gering wie möglich zu halten. Dies ist insofern eine Herausforderung, da nicht alle thematischen Datensätze einen entsprechenden Zeitstempel verfügen. Eine weitere Herausforderung bei der Verwendung der thematischen Daten ist die (quantifizierbare) Güte der Daten - also wie zuverlässig sind die räumlichen und thematischen Informationen. Da mögliche Ungenauigkeiten oder Fehler in den Daten auch im Rahmen der Klassifikation zu Unsicherheiten führen, wird eine Interpretation der Ergebnisse erschwert und die statistischen Metriken bezüglich der Genauigkeiten beziehen sich nicht nur auf den Klassifikationsprozess, sondern beinhalten auch die Datenunsicherheit. Um dies zu trennen bzw. den jeweiligen Anteil quantifizieren zu können, müsste eine entsprechende Validierung für jeden thematischen Datensatz, der im Prototyp AI Verwendung findet, durchgeführt werden. Im Zusammenhang mit der Verwendung der thematischen Daten und für eine datentyp-unabhängigen Verschneidung dieser Daten wäre der Aufbau einer geo-räumlichen Datenbank sehr hilfreich. Diese könnte auch in den Prototypen AI problemlos integriert werden und die Anwendung bezüglich der Auswahl des Dateninputs (z.B. im Rahmen eines iterativen Samplings) und der entsprechenden Vorbereitung der Daten vereinfachen.

5.2 CNN

Auf der Ebene des Deep Learning CNN sind mehrere mögliche Optimierungen vorstellbar, die Potential für eine Steigerung der Klassifikationsgenauigkeit zulassen. So wurde beispielsweise im Rahmen dieses Prototyps lediglich eine Kachelgrösse getestet und trainiert: 50x50m / bzw. 200x200 Px (sowohl für AREA3, als auch AREA4-RGB und AREA4-FCIR). Diese Kachelgrösse würde aufgrund der Erkenntnisse der Bachelor Thesis von Markus Schär verwendet, da er mit der Wahl dieser Kacheln die besten Ergebnisse erzielte. Die Untersuchung abweichender Kachelgrössen könnte sich trotzdem lohnen, da Schär eine andere Netzwerkarchitektur verwendete und statt des Transfer-Learnings ein von Grund auf neu trainiertes Netzwerk einsetzte. Die derzeitige Implementierung des Preprocessing-Workflows gestattet den Export anderer Kachelgrössen, sodass ein Durchlauf mit einer anderen Kachelgrösse verhältnismässig einfach zu bewerkstelligen sein sollte. Insbesondere wäre eine Vergrösserung der Kacheln auf 299x299 Px – die native Resizing-Auflösung des Xception-Netzwerkes – untersuchenswert. Dies entspräche ca. 75x75m bei 25cm AREA4 GSD. Auch im Zusammenhang mit den neuen 10cm-GSD-SwissImage-Bildern, welche in AREA5 zum Einsatz kommen werden, wäre die Untersuchung einer insgesamt höheren Bodenauflösung in Zukunft möglich (50x50m entsprächen hier 500x500 Px).

Eine weitere Optimierungsmöglichkeit existiert beim Klassenausgleich der Inputdaten. Derzeit sind einige Klassen noch stark unterrepräsentiert, da in ihnen nur eingeschränkt Ground-Truth-Daten existieren. Dies gilt insbesondere für die grösseren Trainingsdurchläufe, bei welchen – wo es möglich war – 10‘000 oder 18‘000 Samples pro Klasse trainiert wurden. Eine oft angewendete Methode zur Vergrösserung dieser Datensätze ist die Augmentierung von Ground-Truth-Daten zur Verbreiterung der Datengrundlage. Durchgeführte Experimente zeigten jedoch bisher keine signifikanten Verbesserungen in der Trainingsleistung bei der Verwendung von Augmentierungen für dieses Netzwerk. Eine andere Möglichkeit wäre das manuelle Sampling neuer Tiles für die schwach besetzten Klassen aus Orthophotos.

Eine bisher nicht durchgeführte Optimierung des Klassifikationsmodells, welche auch das Problem unterrepräsentierter Klassen teilweise lösen könnte, wäre das gemeinsame Training von AREA3- und AREA4-Inputdaten zu einem Kombinationsmodell. Durch die unterschiedlichen Zeitstände (Frühjahr/Herbst) und unterschiedliche Bodenauflösungen wäre es möglich, einen höheren Grad der Generalisierung im CNN-Klassifikator zu erreichen. Dies würde Anpassungen in Preprocessing-Modul nach sich ziehen, da eine neue Auftrennung in die Subsets Train / Validation / Test notwendig wäre. Derzeit wird die Resultat-Klassifikation des High-Level-Deep-Feature-Vektors (ein Vektor mit 2048 numerischen Variablen) durch eine lineare Regression durchgeführt (Xception Default). Der Resultatvektor verfügt nur noch über so viele Variablen, wie sich Klassen im Training befinden. Eine Optimierungsmöglichkeit wäre, die Hilfs- und Zusatzdaten bei diesem Schritt bereits einfliessen zu lassen, um eine weniger komplexe, einschrittige Post-Klassifikation zu ermöglichen. Eine genaue Bestimmung der Variablengewichte wäre darüber hinaus ausserdem nur noch schwer möglich. Eine Alternative für die Optimierung des Post-Klassifikators wäre daher, keine lineare Regression zum Schluss des CNN durchzuführen, sondern den High-Level-Deep-Feature-Vektor „as is“ in einem komplexen Random Forest mit Zusatzvariablen zu integrieren.

5.3 RF

Eine starke Limitation des RF-Modells zeigt sich bei kleinen Klassen, die mutmasslich eine starke Ähnlichkeit mit einer grossen Klasse aufweisen. So wurden beispielsweise bei „AREA4-LU mit Zusatzdaten“ viele Elemente kleinerer Klassen wie Reihenhäuser, Mehrfamilienhäuser und nicht-spezifizierte Gebäude in die grosse Klasse „Ein- und Zweifamilienhaus“ klassifiziert. Auch viele Aufforstungen, Holzschläge und Waldschäden wurden der grossen Klasse „Waldbestände“ zugerechnet, ebenso wie viele Wiese-/Weidetypen in ihre korrespondierenden grösseren Klassen fielen. Diese auf visuelle Durchsicht beruhenden Vermutung, müsste umfassend untersucht werden, indem fehlkassifizierte Patches systematisch hinsichtlich ihrer Vertauschbarkeit zu analysieren wären

Eine vielversprechende Möglichkeit diese Fehlklassifikationen zu minimieren, wäre die Klassen mit grosser Ähnlichkeit zunächst zusammenzufassen (Pooling) und den kompletten CNN-RF-Durchlauf mit dieser Überklasse durchzuführen. In einem zweiten Schritt würde ein nur für die Unterscheidung innerhalb dieser Klasse trainiertes CNN in Verwendung mit einem weiteren nur für diese Unterscheidung aufgebauten RF die Feinklassifizierung vornehmen. Der Vorteil wäre, dass es einfacher ist, balancierte Ausgangsdatensätze zu erzeugen und darüberhinaus ein detaillierterer Fokus des CNN möglicherweise in besseren RF-Inputdaten resultiert.

Unsymmetrische Klassen sind grundsätzlich nur dann ein Problem, wenn eine Fehleinschätzung höhere Kosten bedeutet. Wenn es kleine Minderheitenklassen gibt und es nicht teurer ist, sie als Mehrheitsklasse einzustufen, als umgekehrt, dann ist es sinnvoll, eine Fehlklassifizierung von Minderheitenklassen zuzulassen. Im Falle des RF ist eine Augmentierung von unterrepräsentierten Datensätzen noch einmal wesentlich komplexer als beim CNN, weshalb die einzig praktische Möglichkeit der Erzeugung eines symmetrischen Datensatzes im Oversampling von Minderheitsklassen besteht, während gleichzeitig sehr grosse Klassen einem Undersampling unterzogen werden müssten. Dies würde jedoch unter Umständen die Gesamtgenauigkeit des Modells erheblich verschlechtern. Der Ansatz ist aber trotzdem untersuchenswert. Auch ein zweistufiges Verfahren könnte noch einen erheblichen Genauigkeitsgewinn erbringen, besonders beispielsweise im Hinblick auf eine Feindifferenzierung von Wald- oder Siedlungsflächen.

Auf experimenteller Ebene ist insbesondere die Durchführung umfangreicherer Change-Detection-Modelle (z.B. mit dem Maximum von ~100‘000 Change und ~100‘000 No-Change Samples im 80% Training / 20% Validation – Split) jeweils für Land Cover und Land Use mit moderatem Zusatzaufwand möglich, und die Frage nach der Leistungsfähigkeit eines solchen Modells liesse sich damit adressieren. Dazu müsste ein Datensatz mit identischer Variablenselektion über zwei Erhebungszeitpunkten erstellt werden. Da jedoch grundsätzlich die Genauigkeiten der Klassifikationsmodelle jeweils für LC und LU über den erreichten Werten der gesonderten CD-Modellen lagen, ist ein tabellarischer Ansatz unter der Grundannahme, dass die Ground-Truth-Werte der letzten Arealstatistikerhebung bekannt sind, vermutlich weiterhin überlegen. Optimierungspotential zeigt sich insbesondere noch durch Einbezug von Ground-Truth-Daten früherer Erhebungen (in diesem Fall AREA3), sowie beim Training eines maximal grossen Stichprobendatensatzes.

Auch eine Ausweitung der gemeinsamen LULC-Klassifikation im RF auf einen grösseren Anteil der bis zu 1,2 Mio möglichen Samples im 80% Training / 20% Validation – Split wäre denkbar, vor allem um zusätzliche Klassenkombinationen mit sehr niedriger Prävalenz in das Training aufnehmen zu können. Die Grösse des Inputvektors mit sehr vielen Variablen-Spalten und die hohe Anzahl möglicher Outputklassen erhöhte jedoch bereits im vorliegenden Training (in welchem nur 5% des Datensatzes genutzt wurden) die Berechnungsdauer massiv (ca. Faktor x3). Dies scheint auch dem Versuch entgegenzuwirken, den Deep Feature Vektor mit 2048 Einzelvariablen statt des klassifizierten CNN-Outputs zu verwenden. Die tatsächliche Berechnungsdauer müsste jedoch experimentell bestimmt werden.

Eine deutliche Verbesserung der Usability ist im Hinblick auf das Preprocessing der RF-Eingabedaten möglich. Derzeit müssen diese Datensätze „von Hand“ verkettet werden, indem je nach Einsatzszenario eine Reihe Pandas-Kommandos in einer interaktiven Pythonkonsole ausgeführt werden. Hier würde die Programmierung von Zusatzroutinen und eine weitere Standardisierung der Inputdaten eine Stratifizierung des Workflows ermöglichen. Das Ziel für eine produktive Umgebung über das Stadium eines Prototyps hinaus wäre die Einbettung des Random Forest in das existierende Command-Line-Tool, sodass Preprocessing, Deep Learning, Random Forest und Evaluierung aus einem kohärenten System heraus aufgerufen werden können. Die Einbindung des Random Forest

könnte beispielsweise über einen R/Python-Adapter erfolgen. Rpy2 ist eine R-Schnittstelle, die innerhalb eines Python-Prozesses aufgerufen wird. Das Paket ist ausgereift, stabil, gut dokumentiert, einfach einzubinden und wird vielfach produktiv genutzt. Alternativ könnte der Random Forest auch mit dem Python-Machine-Learning-Framework SciKit-Learn umgesetzt werden, das eigene RF-Implementierungen mitbringt. In den letzten Jahren wurde auch die SciKit-Learn RF-Implementierung auf ein Cython-Backend umgestellt, sodass die Berechnungszeiten auch bei diesem Ansatz dramatisch gesunken sind. Prozedurale Python-Calls stellen jedoch immer noch aufgrund eines starken Overheads einen „Flaschenhals“ dar und erfordern so eine stringente Datentypoptimierung. Andererseits ist das native Speichermanagement von pandas/numpy demjenigen von R überlegen, sodass selbst sehr grosse Datensätze ohne splitting/batch processing direkt weiterverarbeitet werden können. Hier könnten aber auch Speicheroptimierungen im R-Script (z.B. mit den Modulen bigmemory, ff oder Hadoop) noch erhebliche Verbesserungen bringen. Eine Entscheidung für den einen oder anderen Entwicklungsweg sollte sich am aktuellen Stand der Forschung und den zur Verfügung stehenden Ressourcen orientieren.

Eine massive Verbesserung für den Workflow, sowie die Geschwindigkeit von Dataset-Input/Output-Prozessen wäre darüber hinaus mit der Implementierung einer Datenbank-Struktur umsetzbar (z.B. PostgreSQL). Der derzeitige Ansatz, Zwischenergebnisse und Zusatzdaten als CSV-Tabellen zu speichern erhöht zwar die funktionale Kompatibilität und Flexibilität des Prototyps und war für den Entwicklungsprozess auch zwingend notwendig. Gleichzeitig bedingt der Ansatz aber auch eine zeitintensive, komplexe Zusammenstellung und Verschaltung der Workflow-Elemente, die für eine produktive Umgebung im Hinblick auf die Usability noch reduziert werden könnte.

5.4 Sensitivitätsanalysen und Validierung

Bedingt durch die Vielzahl an Möglichkeiten der Klassifikator-Parametrisierung und Datenverwendung, ist aus unserer Sicht ein strukturiertes Vorgehen bei der Validierung unbedingt zu empfehlen. Erst wenn der jeweilige Einfluss eines Parameters/ eines Datensatzes verstanden und quantifizierbar ist, können die Ergebnisse der Klassifikation richtig interpretiert und aussagekräftig validiert werden. Dementsprechend sind vor einer thematischen Validierung verschiedene Sensitivitätsanalysen notwendig. Diese können in system-spezifische und daten-spezifische Sensitivitäten unterschieden werden. System-spezifische Sensitivitäten beschreiben die Performance und den Output der Klassifikatoren (z.B. in Bezug auf Stabilität und Genauigkeit) in Abhängigkeit der Parametrisierung und des Daten-Inputs. Die daten-spezifischen Sensitivitäten beschreiben den Einfluss verwendeter Datensätze auf die Performance und den Output der Klassifikatoren. Damit wird auch deutlich, dass sich daten- und system-spezifische Sensitivitäten wechselseitig beeinflussen und zwar getrennt quantifiziert aber jeweils in Abhängigkeit zueinander analysiert werden müssen. Basierend auf den aktuell in der Arealstatistik verwendeten Basis-, Hilfs- und Zusatzdaten, den Parametrisierungsmöglichkeiten der Klassifikatoren und ein auf das Minimum reduzierte Samplingverfahren würde dies mind. 500 Klassifikationsdurchläufe bedeuten. Da sowohl der zeitliche Aufwand der Klassifikationen als auch der Interpretationsaufwand bezüglich der Ergebnisse dementsprechend hoch ist, sind eine vorgängige Priorisierung und ein klar strukturiertes Vorgehen unabdingbar.

Bei der Validierung empfiehlt sich ebenfalls ein strukturiertes Vorgehen. Grundsätzlich können basierend auf den Gesamtgenauigkeiten keine aussagekräftigen Rückschlüsse über die Qualität und dem Mehrwert des Prototypen AI gewonnen werden. Der Vorteil des umgesetzten Klassifikationsverfahrens liegt darin, dass man die punktspezifischen Genauigkeiten auswerten und somit unabhängig von den klassen-spezifischen Genauigkeiten Aussagen über die Güte der Klassenzuweisung für jeden einzelnen Stichprobenpunkt treffen kann. Dabei ist zu berücksichtigen, welche thematische Ebene man auswertet - also ob man die Change/No-Change - Klassifikation beurteilen möchte, oder die Landnutzung, Landbedeckung oder eine Kombination aus Landbedeckung/Landnutzung analysiert (siehe Abbildung 4.9). Die voraussichtlich besten Ergebnisse bezüglich der Klassifikationsgüte sind hierbei mit einem iterativen Verfahren zu erwarten, in dem man jeweils die Stichprobenpunkte mit den am zuverlässigsten ausgewiesenen Klassen von dem nachfolgenden Klassifikationsprozessen ausschliesst. Mit diesem Ansatz könnte auch umgesetzt werden, dass man für bestimmte Klassen angepasste Parametrisierungen vornimmt - also bestimmte Eingangsdaten und Klassifikatoreneinstellungen spezielle für ausgewählte Zielklassen optimiert (z.B. optimiert auf die Ausweisung von Waldklassen oder landwirtschaftliche Flächen). Wenn solch ein strukturiertes Vorgehen auf- und umgesetzt

ist, kann man dann auch den tatsächlichen Mehrwert des Prototypen AI im Rahmen des Erhebungsprozesses der Arealstatistik quantifizieren (z.B. Anzahl Stichprobenpunkte, die nicht mehr interpretiert werden müssen oder die für die Kontrolle Verwendung finden).

6 Schlussfolgerungen und Ausblick

Das von der FHNW und ExoLabs im Auftrag des BFS bearbeitete Projekt „Prototyp AI“ dient der Potenzialabschätzung automatisierter, datengetriebener Methoden, die zukünftig unterstützend bei der Durchführung der Arealstatistik des BFS eingesetzt werden sollen. Der hierfür entwickelte und in vorliegendem Bericht dargestellte Software Prototyp weist als gesamtheitlicher Ansatz darauf hin, dass die State-of-the-Art Methoden des maschinellen Lernens in Kombination mit einer räumlich-temporalen Datenfusion unterschiedlichster Formate dazu geeignet sind, eine substanzielle Arbeitsunterstützung herbeizuführen. Der Prototyp AI ist modular aufgebaut und setzt sich aus Modulen für die Datenaufbereitung, die Klassifikation der präsentierten Multisensordaten durch Deep Learning/ Random Forest - Algorithmen und für die Datenauswertung zusammen (Abbildung 6.1). Zudem wurde der Prototyp dahingehend entwickelt und dokumentiert, dass nach Abschluss des Projektes eine eigenständige Anwendung, Optimierung und Weiterentwicklung möglich ist.

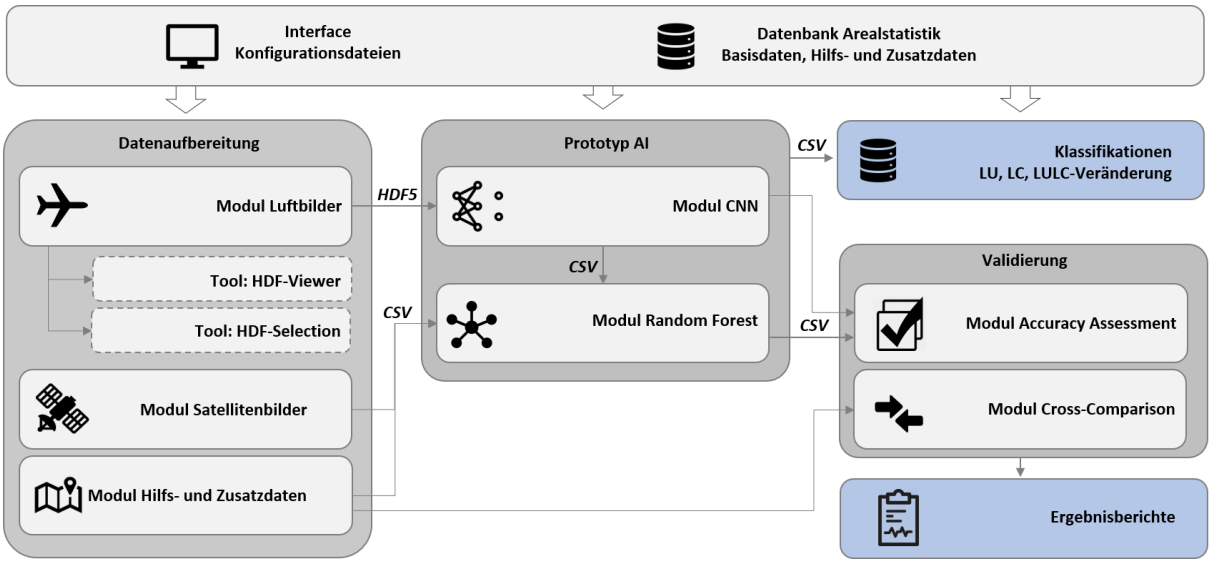


Abbildung 6.1: Aufbau und Funktionsweise des Prototyps AI.

Das Innovationspotenzial der bereitgestellten Methodik kann bereits in diesem prototypischen Stadium genutzt werden, indem die maschinell generierten Klassenzugehörigkeiten sowohl für Landnutzung und Landbedeckung als auch für deren Veränderung über die Zeit dem manuellen Klassifikationsprozess unterstützend zugefügt werden können. Die Klassifikation von Veränderungen ist dahingehend eingeschränkt, dass einige Datenquellen nicht simultan für beide Erhebungsperioden verfügbar sind. Es wurden dementsprechend zusätzlich Ansätze entwickelt, die davon unabhängig angewandt werden können, indem die Veränderungsdetektion anhand beider Klassifikationsergebnisse durchgeführt wird (vgl. Kapitel 2.5.7). Im Rahmen des Projektes wurden diesbezüglich erste Klassifikationsdurchläufe testweise durchgeführt, deren Ergebnisse nachfolgend zusammengefasst sind:

Bei den Klassifikationen der Landbedeckung wurden je nach Parametrisierung und verwendetem Dateninput Gesamtgenauigkeiten (Overall Accuracy) von über 88.5% erreicht (vgl. Kapitel 4.3.3). Wenn die beiden Klassen mit den höchsten Anteilen betrachtet werden (entspricht einer Anzahl von über 2 Millionen Stichprobenpunkten), so können die Klassen “21 Gras-, Krautvegetation” und “41 Geschlossene Baumbestände“ mit Genauigkeiten (F-Mass) von 95.6% (Klasse 21) bzw. 94.4% (Klasse 41) ausgewiesen werden. Im Hinblick auf die Gewichtung der Variablen tragen neben den Informationen aus den Luftbildern (52%) vor allem die Satellitendaten (22%) und die Amtliche Vermessung (17%) zur Klassifikationsgüte bei. Die Validierung der Veränderungsdetektionen für Landbedeckung resultierte mit einer vorläufigen Parametrisierung und einem reduzierten Datensatz in einer Gesamtgenauigkeit von 88.5% (vgl. Kapitel 4.3.5) basierend auf einem direkten Ansatz und von 77.8% unter

Verwendung des dichotomen Ansatzes. Die Gewichtung der Variablen für die Veränderungsklassifikation verteilt sich auf die Luftbilddaten bzw. die darauf basierenden CNN-Ergebnisse (66%) und die Satellitendaten (34%).

Bei den Klassifikationen der Landnutzung wurden Gesamtgenauigkeiten von über 84.2% erreicht (vgl. Kapitel 4.3.8). Wenn wieder die beiden Klassen mit den höchsten Anteilen betrachtet werden (= über 2 Millionen Stichprobenpunkte), so können die Klassen “301 Waldbestände” und “421 Keine Nutzung“ mit Genauigkeiten von 94.7% (Klasse 301) bzw. 89.4% (Klasse 421) ausgewiesen werden. Vergleichbar zur Klassifikation der Landbedeckung tragen neben den Informationen aus den Luftbildern (43%) wieder die Satellitendaten (25%) und die Amtliche Vermessung (16%) primär zur Klassifikationsgüte bei. Die Validierung der Veränderungsdetektionen resultierte in Gesamtgenauigkeiten von bis zu 84.2% (73.5% dichotomer Ansatz) (vgl. Kapitel 4.3.9). Neben den Luftbilddaten sind auch bei der Veränderungsdetektion vor allem die Satellitendaten (54% Variablengewichtung) von besonderer Bedeutung.

Die direkte Klassifikation der kombinierten Landbedeckung/Landnutzung ist einigen Limitationen unterworfen, da für viele Klassenkombinationen keine ausreichend grosse Stichprobenanzahl für ein Training/eine Klassifikation zur Verfügung stehen. Dementsprechend konnten nur 58 Klassen (von über 300) ausgewiesen werden, die jeweils mindestens 0.1% Anteil an den Samples hatten. Die Gesamtgenauigkeit für diese Klassen lag bei 77.2% (vgl. Kapitel 4.3.10). Die Klassenkombinationen mit den höchsten Anteilen (= über 2 Millionen Stichprobenpunkte) wurden mit einem F-Mass von 94.3% (41/301 – Geschlossene Baumbestände/ Waldbestände) und 89.7% (21/221 – Gras-, Krautvegetation/ Ackerland) ausgewiesen.

Diese vielversprechenden Voruntersuchungen stützen auf zuverlässige Referenzdaten des BFS, welche im Rahmen des maschinellen Lernprozesses genutzt wurden. Auf dieser Grundlage wurde in einem äusserst anspruchsvollen Setting - bestehend aus komplexen Eingangsdaten, Identifizierung geeigneter Klassifikationsmethoden, Beherrschung grosser Datenmengen und zeitlicher Limitierungen - ein auf die Bedürfnisse des BFS zugeschnittenes, ready-to-use Softwarepaket erfolgreich umgesetzt und in die Infrastruktur des BFS implementiert. Wir sind überzeugt, dass der vorgestellte Ansatz aus moderner Machine Learning - Methodik und Multisensor - Datenfusion zukunftsweisend ist.

In Hinblick auf eine effizientere Anwendung, wäre nachfolgend die Integration aller Teilmodule in einer ready-to-use Software mit intuitiver GUI empfehlenswert, da dies die Handhabung des Prototypen AI in einem operationellen Setup deutlich optimieren würde (vgl. Kapitel 5). Insbesondere bei iterativen Vorgehensweisen würde solch ein Software-Paket die Parametrisierung und Automatisierung stark vereinfachen und unterstützen. Zudem besteht Potential zur Optimierung der Klassifikatoren und des Klassifikationsprozesses sowie zur Erweiterung des Prototyps AI in Hinblick auf die Integretation zusätzlicher Datenquellen. Diese sind in der „Liste Optionen – Prototyp AI“ zusammengefasst.

7 Literatur

Bundesamt für Statistik (2004): Arealstatistik nach Nomenklatur 2004 - Standard. GEOSTAT-Datenbeschreibung Arealstatistik NOAS04-Standard. Unter Mitarbeit von Rainer Humbel, A. Beyeler, J. Burkhalter, R. Pfister, J. Sager und H.-U. Zaugg. Bundesamt für Statistik. Online verfügbar unter https://www.bfs.admin.ch/bfsstatic/dam/as-sets/264654/master; accessed on: 2018-06-12.

Bundesamt für Statistik (2018a): Arealstatistik 2013/18 - Stand der Datenerhebung am 01.06.2018. Online verfügbar unter https://www.bfs.admin.ch/bfs/de/home/statisti-ken/raum-umwelt/erhebungen/area/2013-18/datenerhebung.assetde-tail.5346492.html; accessed on: 2018-06-12.

Bundesamt für Statistik (2018b): Arealstatistik Schweiz. Online verfügbar unter https://www.bfs.admin.ch/bfs/de/home/statistiken/raum-umwelt/erhebun-gen/area.html; accessed on: 2018-07-04.

Chollet, F. (2017): Xception: Deep Learning with Depthwise Separable Convolutions. In: 30th IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, S. 1800–1807; accessed on: 2018-07-20.

Hochuli, S. (2018): Bachelor-Thesis "Deep Learning für die Arealstatistik der Schweiz", Institut Geomatik, Fachhochschule Nordwestschweiz; Abgabe: 2018-07-27.

Morris, B. (2018): Mastering the Learning Rate to Speed Up Deep Learning, Arizon State University, https://www.kdnuggets.com/2018/11/mastering-learning-rate-speed-up-deep-learning.html; accessed on: 2018-10-20.

Picterra (2017): Feasibility Study - Arealstatistik 2020, v2.0; Catalag ID: do-t-02.02.01-Study; accessed on: 2019-04-05.

Wright, M. N. & Ziegler, A. (2017): ranger - A fast implementation of random forests for high dimensional data in C++ and R. Journal of Statistical Software 77:1-17

8 Anhang

Nachfolgend findet sich ein Inventar der im Rahmen der vorliegenden Studie erbrachten Leistungen.

1) Projektdokumentation – allgemeine Projektbeschreibung inklusive Evaluation der Funktionsweise des Prototyps (= vorliegendes Dokument)
2) Technische Dokumentation Prototyp AI – Quellcode-Beschreibung der jeweiligen Softwaremodule inklusive Erläuterungen zur Implementierung (basierend auf SPHINX = Anhang_B_PrototypeAI)
3) Skripte – mit Dokumentation im Quellcode, zum Teil untereinander verknüpft und mit eingebetteten Funktionen
    a. Datenvorverarbeitung/ Datenaufbereitung (Raster/Vector und CSV-Datenformate)
    b. Daten-Sampling (für einen direkten Input in die Klassifikatoren)
    c. HDF5-Visualisierung (Visualisierung der Luftbild-Patches)
    d. CNN (Klassifikator)
    e. RF (Klassifikator)
    f. Validierung & Cross-Comparison (Berechnung und Ausgabe gängiger Genauigkeitsmetriken)
4) Ordner: Ergebnisstabellen- und Abbildungen von Zwischenschritten und Testdurchläufen
5) Ordner: Testdatensätze
6) Liste Optionen – Prototyp AI