

Mit KI auf der Spur von Wildschweinen am Streckennetz der SNCF

Einleitung

Kollisionen von Zügen mit Wildschweinen führen zu:

- Sachschäden
- Verspätungen
- Zugausfällen
- wirtschaftliche Einbußen.

Ziel dieser Bachelor-Thesis:

- spezifische Bodenbedeckungen entlang von SNCF-Eisenbahnstrecken als Indikator für die Habitatpräferenz von Wildschweinen zu klassifizieren.

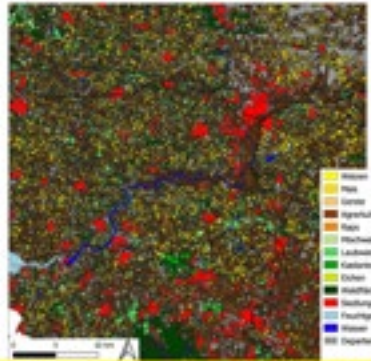


10. Juni 2024 Autor: Théo Reibel, Examinator: Prof. Dr. Denis Jordan, Betreuer: Adrian Meyer, Experten: SNCF Réseau © IGEO FHNW

Referenzdaten

Verwendete Daten:

- Agrikulturreferenzen: Régistre Parcellaire Graphique (RPG)
- Walddaten: BD Forêt
- Gewässer: OpenStreetMap (OSM)
- Siedlungen: CORINE Land Cover (CLC)



10. Juni 2024 Autor: Théo Reibel, Examinator: Prof. Dr. Denis Jordan, Betreuer: Adrian Meyer, Experten: SNCF Réseau © IGEO FHNW

Künstliche Intelligenz

Modell:

- Pflanz-100m-multi-temporal-crop-classification
- Anforderungen des Modells:
 - Satellitenbilder (HLS) mit 18 Bänder (drei Zeitstände)
 - 224x224 Pixel grosse Masken/Bilder
 - 30 Meter geometrische Auflösung

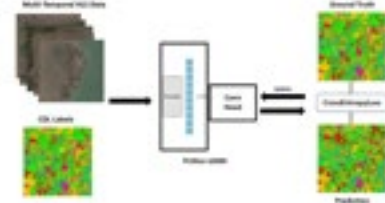


Abb. 1: Schéma der Pflanz-100m-multi-temporal-crop-classification Modell (deep-learning-basiertes neuronales Netzwerk/Pflanz-100m-multi-temporal-crop-classification)

10. Juni 2024 Autor: Théo Reibel, Examinator: Prof. Dr. Denis Jordan, Betreuer: Adrian Meyer, Experten: SNCF Réseau © IGEO FHNW

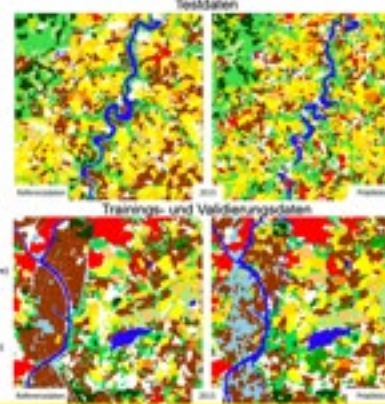
KI-Training Resultat

Training:

- 80 % Trainingsdaten, 20 % Validierungsdaten
- 160 Trainingsepochen
- Neues Datensatz in der Größenordnung der Validierungsdatsatz für die Qualitätsbeurteilung

Testdaten

Trainings- und Validierungsdaten



10. Juni 2024 Autor: Théo Reibel, Examinator: Prof. Dr. Denis Jordan, Betreuer: Adrian Meyer, Experten: SNCF Réseau © IGEO FHNW

Satellitenbildern

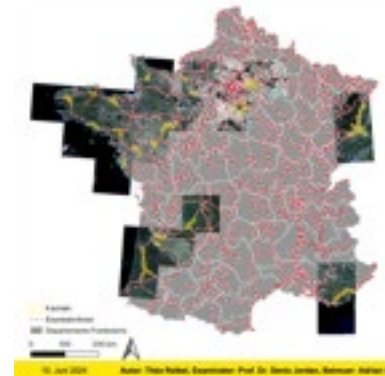
Harmonized Landsat- und Sentinel-2 (HLS)

Vorteile:

- Historisierung
- Zeitliche Auflösung (2 bis 3 Tage)
- Daten von verschiedenen Satelliten (Landsat und Sentinel-2)
- Spektrale Auflösung

Nachteile:

- Geometrische Auflösung (GSD = 30 Meter)
- Komplexes Herunterladen mit MGRS-Tiles
- Wolkenbedeckung und fehlende Abdeckung am Boden



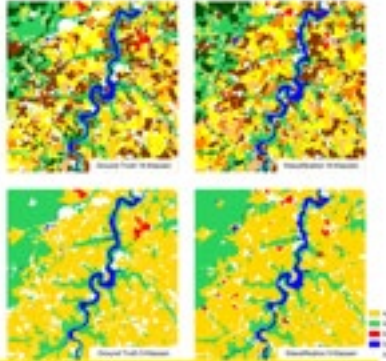
10. Juni 2024 Autor: Théo Reibel, Examinator: Prof. Dr. Denis Jordan, Betreuer: Adrian Meyer, Experten: SNCF Réseau © IGEO FHNW

Diskussion

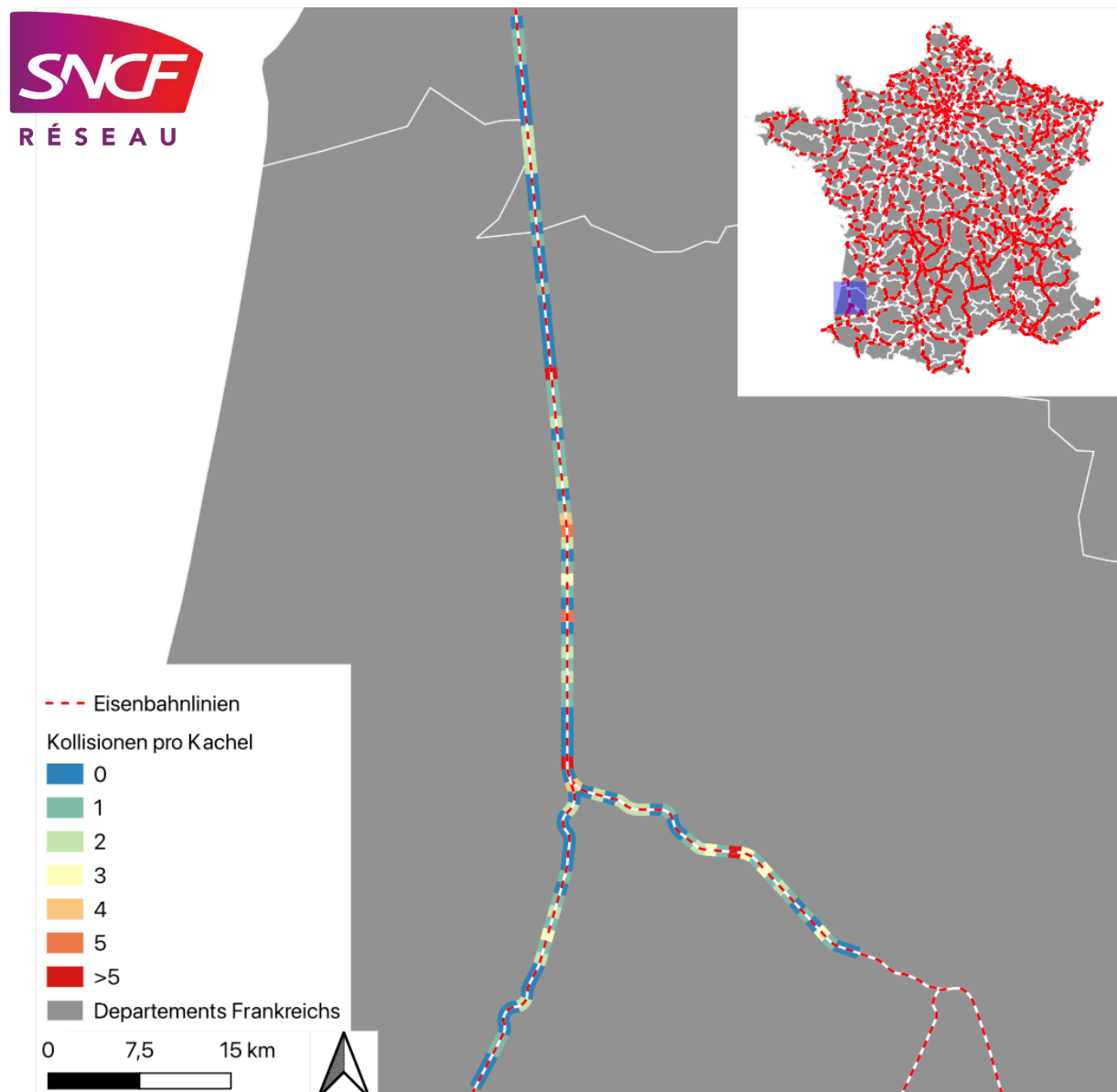
Ein Vergleich von Klassifikationen zeigt:

- Bei 5 Klassen erreicht die Klassifikation eine Genauigkeit von 78,94 % auf den Testdaten.
- Bei 14 Klassen ist die Genauigkeit auf 51,79 %.

Die Klassifikation mit weniger Klassen (5) ist also präziser als die mit mehr Klassen (14).



10. Juni 2024 Autor: Théo Reibel, Examinator: Prof. Dr. Denis Jordan, Betreuer: Adrian Meyer, Experten: SNCF Réseau © IGEO FHNW



Einleitung

Kollisionen von Zügen mit Wildschweinen führen zu:

- Sachschäden
- Verspätungen
- Zugausfällen
- wirtschaftliche Einbussen.

Ziel diese Bachelor-Thesis:

- spezifische Bodenbedeckungen entlang von SNCF-Eisenbahnstrecken als Indikator für die Habitatpräferenz von Wildschweinen zu klassifizieren.

Künstliche Intelligenz

Modell:

- Prithvi-100m-multi-temporal-crop-classification
- Anforderungen des Modells:
 - Satellitenbilder (HLS) mit 18 Bänder (drei Zeitstände)
 - 224x224 Pixel grosse Masken/Bilder
 - 30 Meter geometrische Auflösung

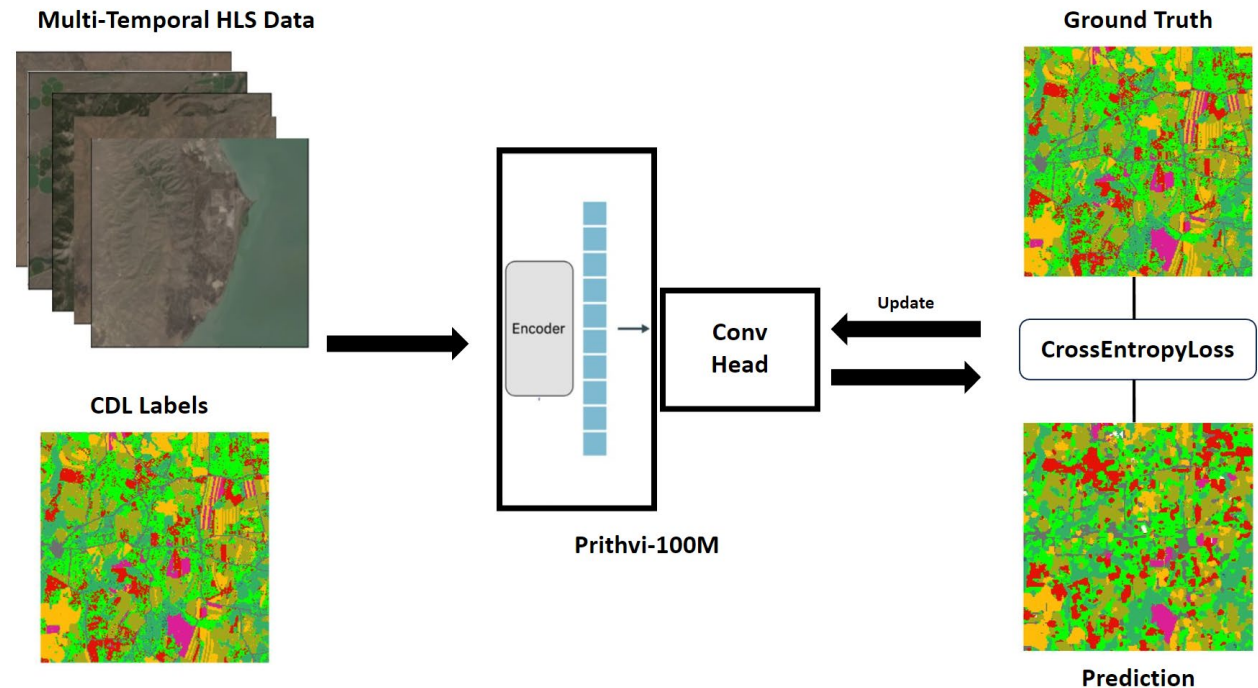


Abb. 1: Schema des Prithvi-100m-multi-temporal-crop-classification Modell (<https://huggingface.co/ibm-nasa-geospatial/Prithvi-100M-multi-temporal-crop-classification>).

Satellitenbildern

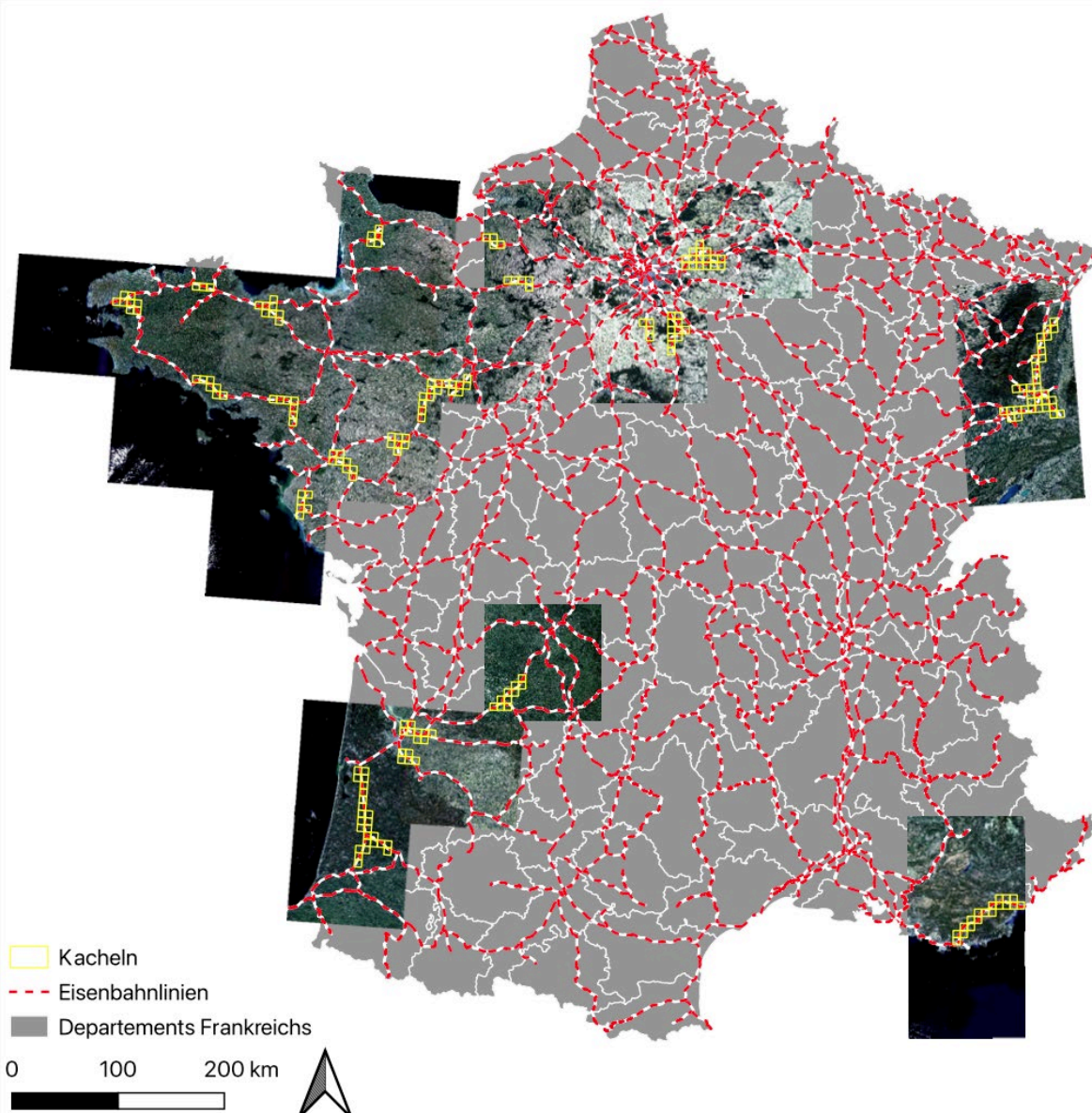
Harmonized Landsat- und Sentinel-2 (HLS)

Vorteile:

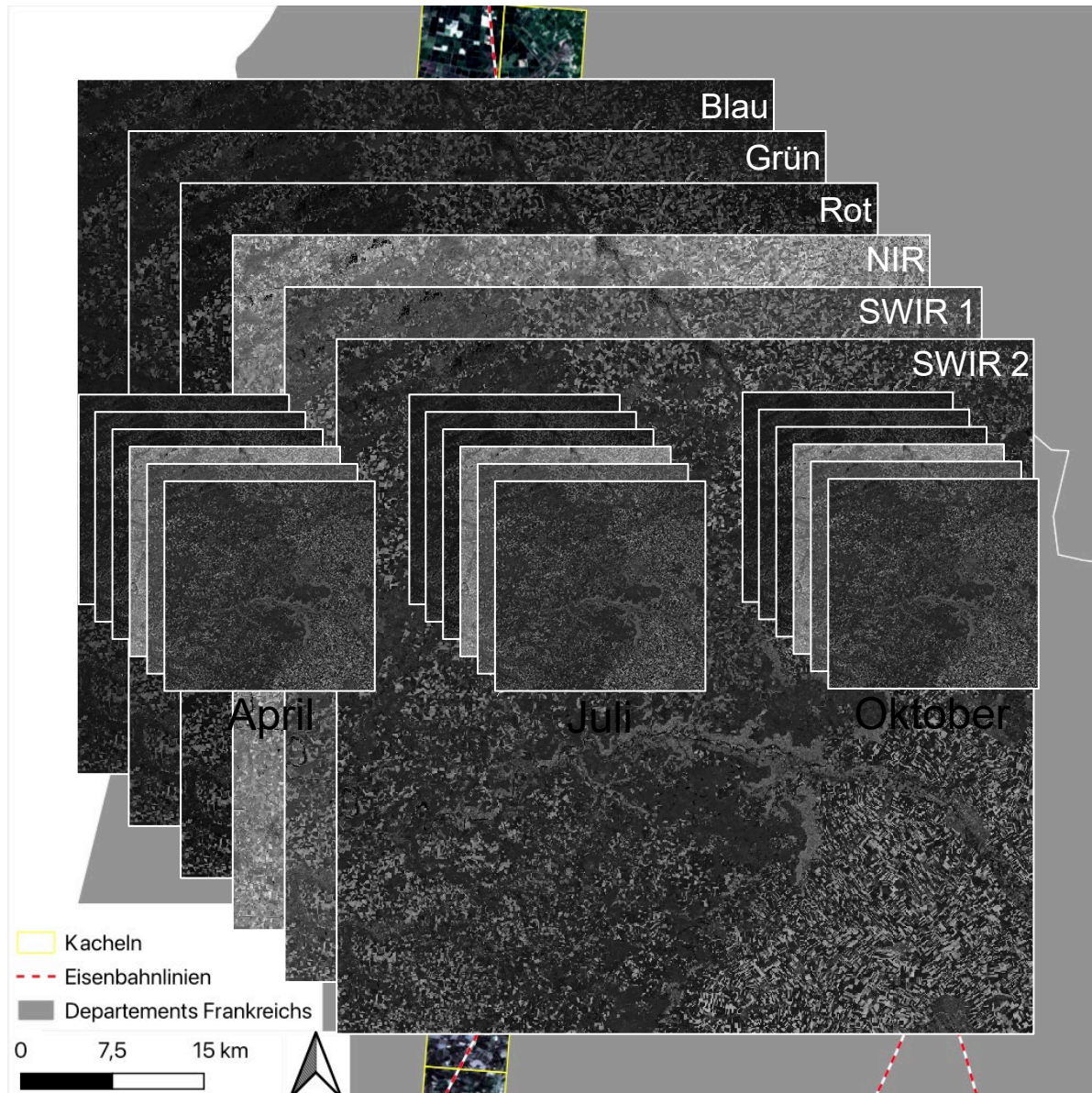
- Historisierung
- Zeitliche Auflösung (2 bis 3 Tage)
- Daten von verschiedenen Satelliten (Landsat und Sentinel-2)
- Spektrale Auflösung

Nachteile:

- Geometrische Auflösung (GSD = 30 Meter)
- Kompliziertes Herunterladen mit MGRS-Tiles
- Wolkenbedeckung und fehlende Abdeckung am Boden



MGRS = Military Grid Reference System



Datenverarbeitung

Herstellung von Satellitenzeitreihen:

- Layerstack pro Zeitstand (April, Juli, Oktober): Blau, Grün, Rot, NIR, SWIR 1, SWIR 2
- Layerstack von drei Zeitstände
- Ergebnis: Satellitenzeitreihen mit 18 Bänder (pro Jahr)

Segmentierung der Satellitenzeitreihen:

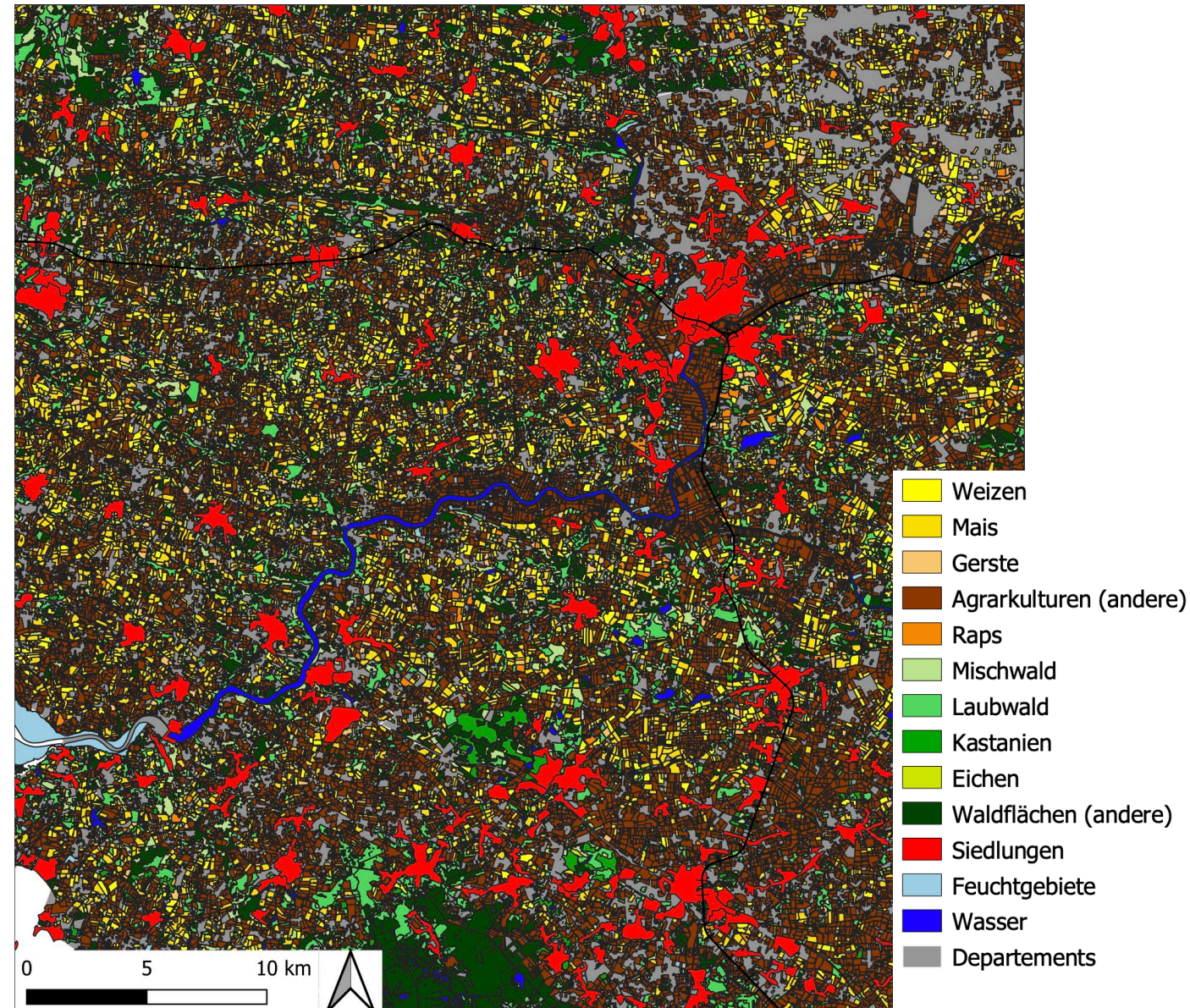
- Zuschneiden der Satellitenbilder auf 6'720x6'720 Meter grosse Kacheln
- Dies entspricht der benötigten Grösse für das KI-Training: 224x224 Pixel, wobei 1 Pixel = 30 Meter

NIR = Near Infrared
SWIR = Short-Wave Infrared

Referenzdaten

Verwendete Daten:

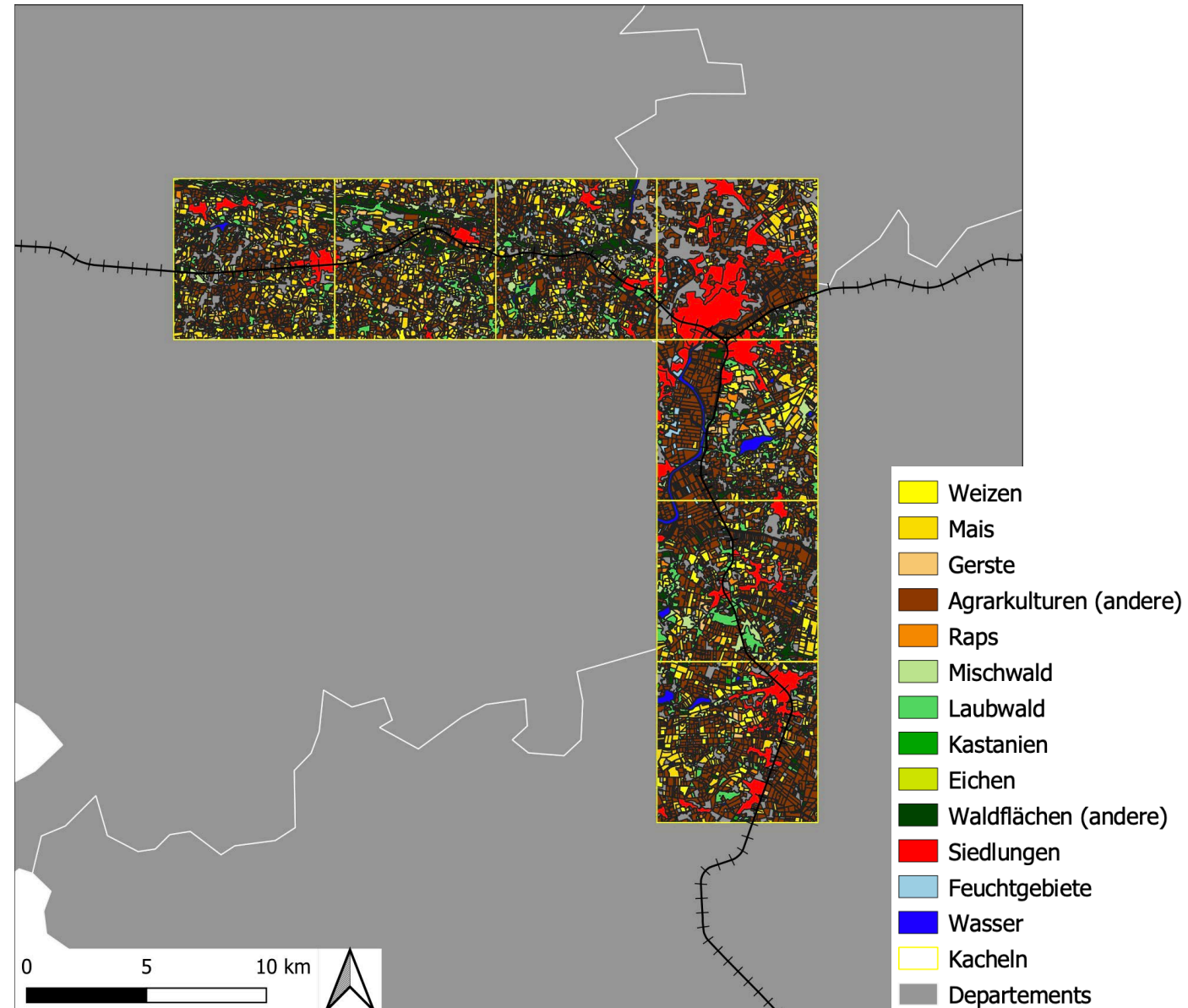
- Agrarkulturrendaten: Registre Parcellaire Graphique (RPG)
- Walddaten: BD Forêt
- Gewässer: OpenStreetMap (OSM)
- Siedlungen: CORINE Land Cover (CLC)



Segmentierung

Segmentierung der Referenzdaten:

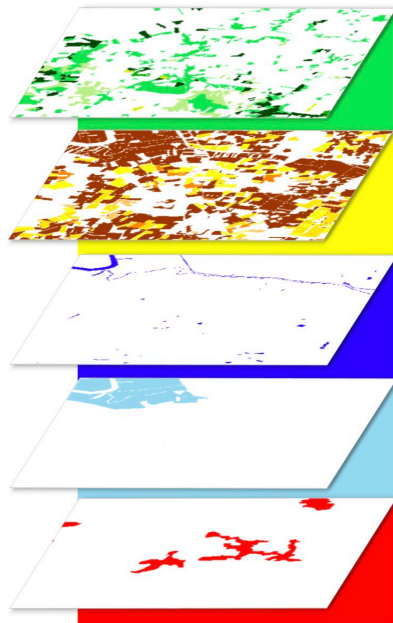
- Zuschneiden der Referenzdaten auf 6'720x6'720 Meter grosse Kacheln
- Dies entspricht der benötigten Grösse für das KI-Training: 224x224 Pixel, wobei 1 Pixel = 30 Meter



Rasterisierung

Eigenschaften der Masken:

- 14 Klassen
- Pixelgrösse: 30x30 Meter
- Klassenreihenfolge:



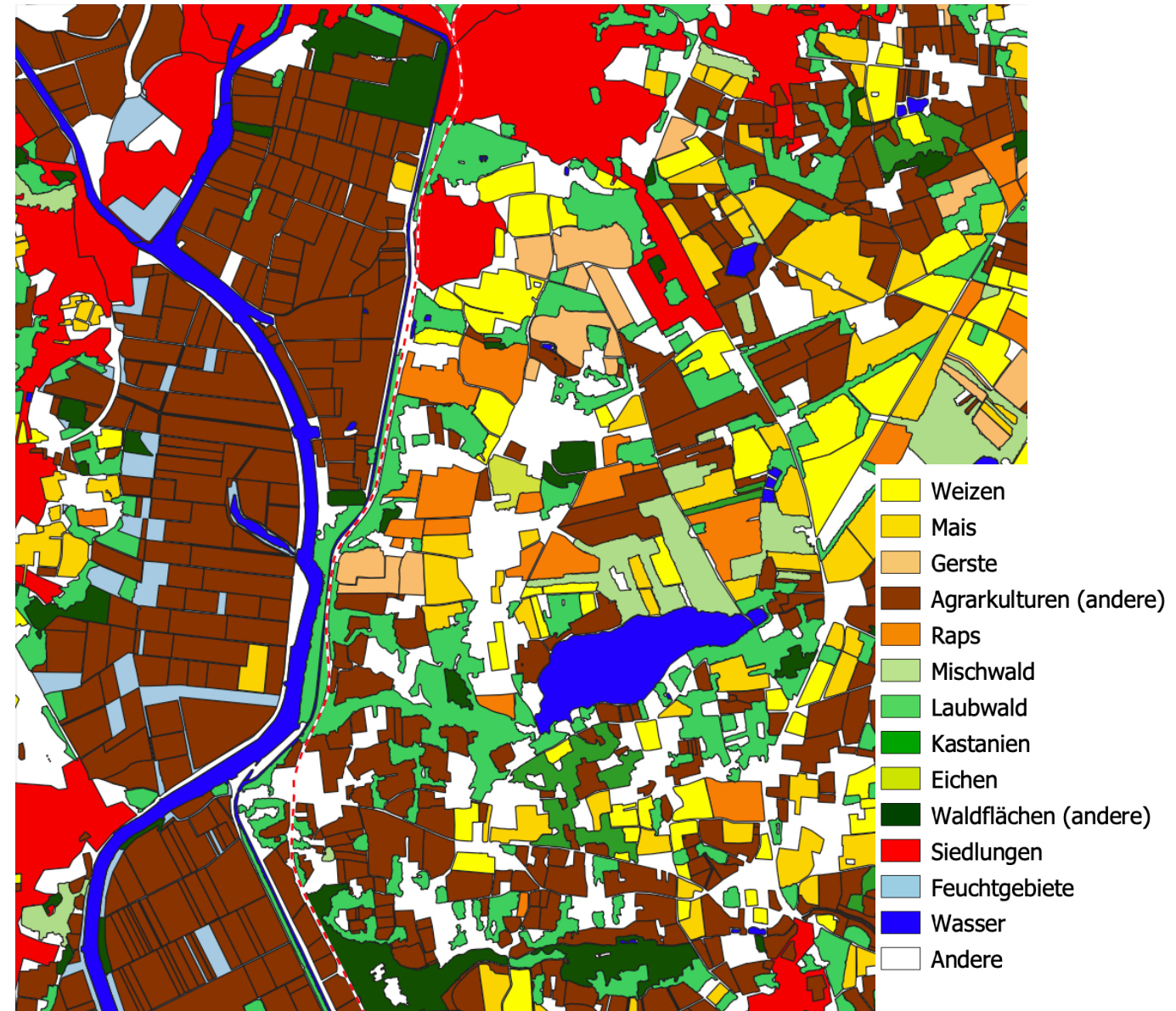
BD Forêt

RPG

Wasser

Feuchtgebiete

Siedlungen

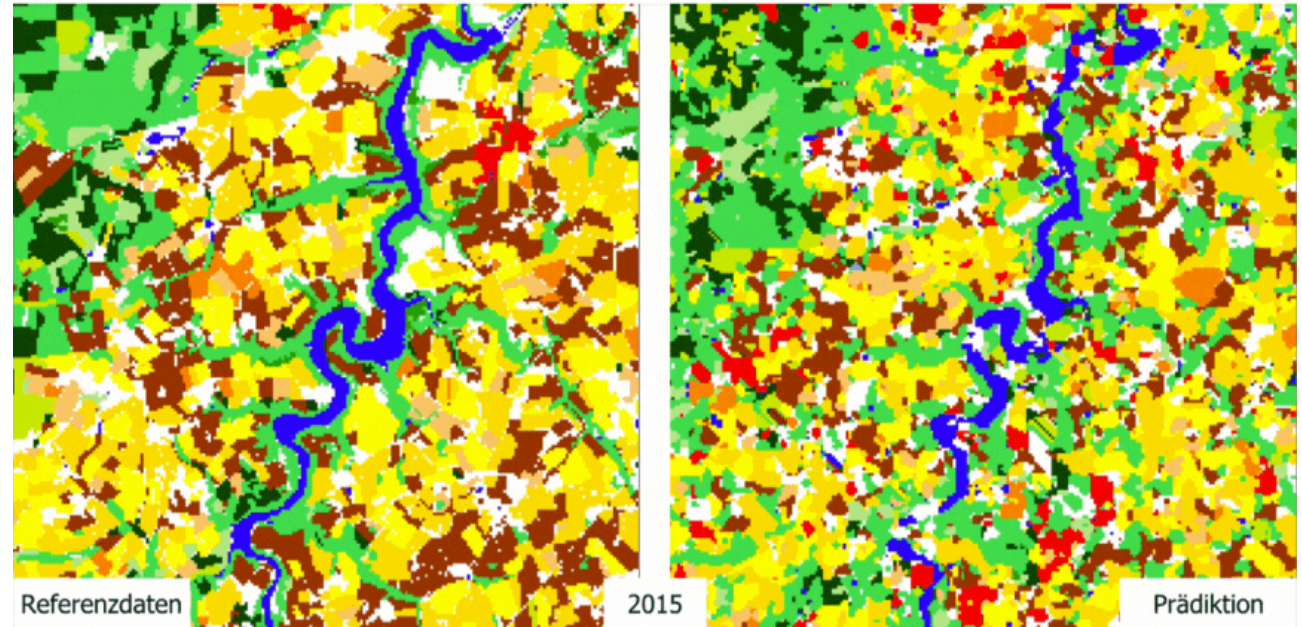


KI-Training Resultat

Training:

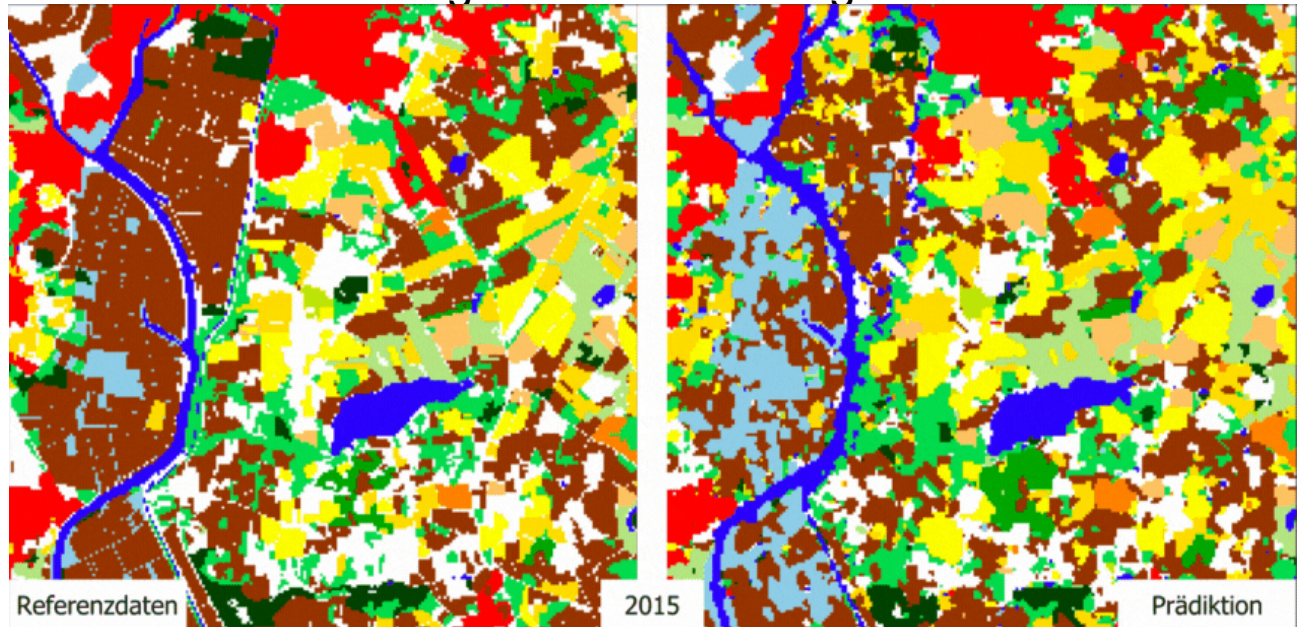
- 80 % Trainingsdaten, 20 % Validierungsdaten
- 160 Trainingsepochen
- Neues Datensatz in der Grössenordnung der Validierungsdatsatz für die Qualitätsbeurteilung

Testdaten

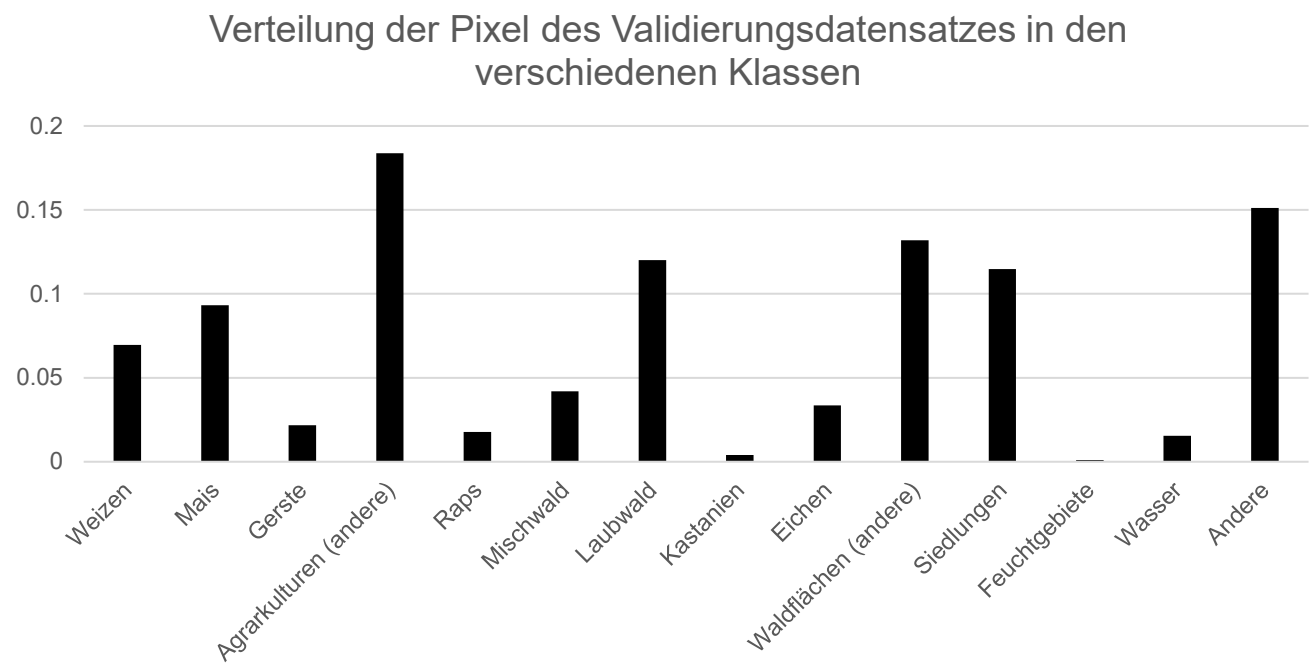


Trainings- und Validierungsdaten

- Weizen
- Mais
- Gerste
- Agrarkulturen (andere)
- Raps
- Mischwald
- Laubwald
- Kastanien
- Eichen
- Waldflächen (andere)
- Siedlungen
- Feuchtgebiete
- Wasser
- Andere



Genauigkeit (Validierungsdaten)



Klassen	Klassengenauigkeit	IoU
Weizen	83%	71%
Mais	83%	71%
Gerste	81%	67%
Agrarkulturen (andere)	74%	60%
Raps	83%	71%
Mischwald	84%	72%
Laubwald	82%	70%
Kastanien	85%	71%
Eichen	86%	76%
Waldflächen (andere)	89%	81%
Siedlungen	92%	86%
Feuchtgebiete	71%	43%
Wasser	86%	74%
Andere	64%	47%

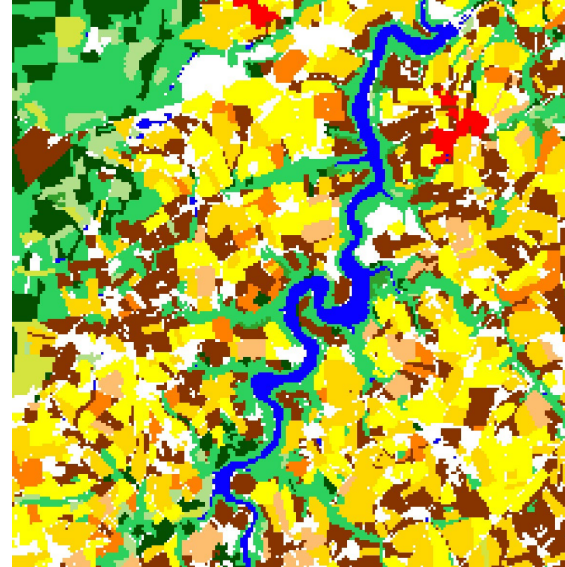
Gesamtgenauigkeit: 80.25 %

Diskussion

Ein Vergleich von Klassifikationen zeigt:

- Bei 5 Klassen erreicht die Klassifikation eine Genauigkeit von 78,94 % auf den Testdaten.
- Bei 14 Klassen ist die Genauigkeit auf 51,79 %.

Die Klassifikation mit weniger Klassen (5) ist also präziser als die mit mehr Klassen (14).



- Agrarkulturen
- Waldflächen
- Siedlungen
- Gewässer
- Andere

Fazit

Die Ergebnisse sind erfolgreich, die Klassifikation hat gut funktioniert.

In den Testdaten:

- Kastanien und Feuchtgebiete nicht gut erkannt
- Verschiedene Wälder allgemein gut als Wälder erkannt, aber nicht immer die richtige Art von Wald
- Analoges gilt für Agrarkulturen

