

# KIHBA – Mit künstlicher Intelligenz Baumarten erkennen

**Bettina Felten, Steffen Günster, Peter Schauer, Nicole Schmeitzner, Rudolf Seitz, Adelheid Wallner**

Als Grundlage für die Umsetzung von Maßnahmen der Klimaanpassung, des Monitorings, der Beratung von Waldbesitzern sowie weiterer forstlicher Aufgaben ist eine flächendeckende, räumlich hochaufgelöste Baumartenkarte ein lang gehegter Wunsch der Forstpraxis. Kann dieser Wunsch mit Hilfe künstlicher Intelligenz erfüllt werden? Dieses Ziel verfolgt die LWF im Fernerkundungsprojekt KIHBA – Künstliche Intelligenz für hochaufgelöste Baumartenerkennung, in Kooperation mit der Firma IABG mbH.

Eine hochaufgelöste, räumliche Erfassung der Hauptbaumarten für ganz Bayern existiert in diesem Sinne nicht. Informationen über bayernweite Baumartenanteile und ihre Verteilung beruhen auf der Hochrechnung und räumlichen Interpolation von Stichprobeninformationen der Bundeswaldinventur. Diese Ergebnisse erlauben aber keine Ableitung einer flächenscharfen, räumlichen Verteilung der Baumarten.

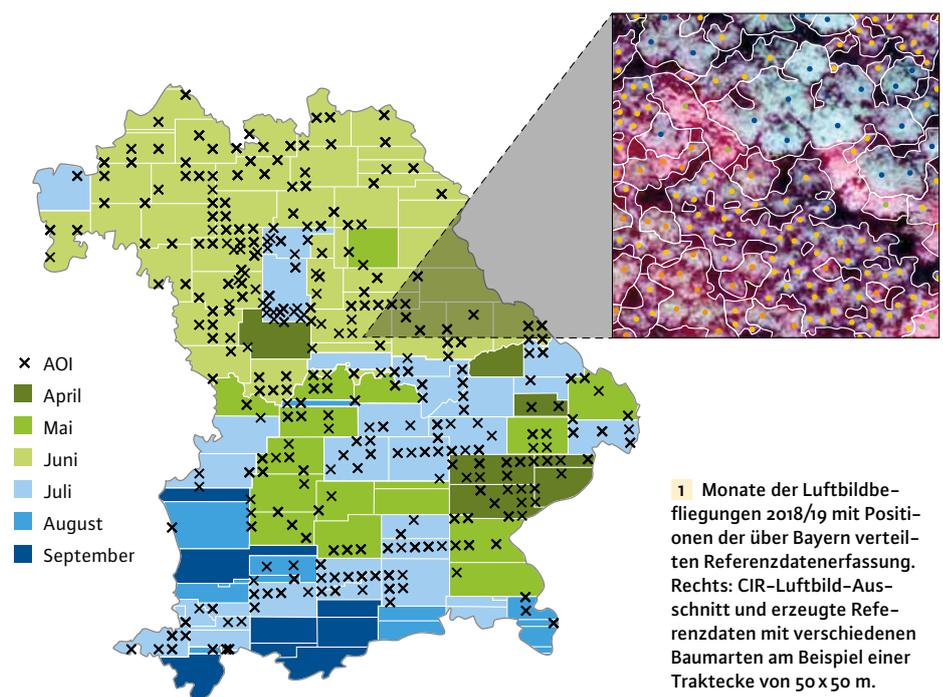
Seit Ende 2022 gibt es erstmals eine deutschlandweite Baumartenkarte mit einer räumlichen Auflösung von 10 m x 10 m, die durch das Thünen-Institut erstellt und veröffentlicht wurde (Blickensdörfer et al. 2023). Diese Karte beruht auf klassischen Verfahren des maschinellen Lernens (ML). Als Datengrundlage dienten Satellitenaufnahmen (2017/2018) des Sentinel-2 Satelliten aus dem europäischen Erdbeobachtungsprogramm Copernicus. Aufgrund der geringen räumlichen Auflösung ist eine Baumartenerkennung zwar auf Bestandesebene möglich, jedoch nicht auf Einzelbaumebene.

Im Projekt KIHBA erhoffen wir uns mit Hilfe von Deep-Learning (DL)-Algorithmen und einer höheren räumlichen Auflösung der verwendeten Fernerkundungsdaten ein Modell zur Ableitung einer bayernweiten Baumartenkarte auf Einzelbaumebene erzeugen zu können. Die für die Modellierung gewählten Klassen entsprechen dabei den in Bayern vorkommenden Hauptbaumarten: Buche, Eiche, Fichte und Kiefer sowie sonstiges Nadel- und Laubholz. Zusätzlich wurde eine weitere Klasse Totholz definiert.

Ergebnisse aus den bisher an der LWF in Kooperation mit der IABG durchgeführten Fernerkundungsstudien BeechSAT und IpsSAT (Straub et al. 2021, Straub et al. 2023) zeigten höhere Klassifizierungsgenauigkeiten von DL-Verfahren gegenüber klassischen ML-Verfahren. Da DL-Methoden zur Bildauswertung räumliche Strukturen und Zusammenhänge erfassen können, verspricht DL eine rein auf Farbinformation basierende Auswertung durch die Nutzung von Strukturinformation zu ergänzen. Gegenüber klassischem ML benötigen Methoden des DL jedoch bedeutend größere Mengen an Referenzdaten, die in der Regel arbeitsaufwändig erhoben werden müssen.

## Datengrundlage und Referenzdatenerzeugung

Zur Modellierung auf Einzelbaumebene und Nutzung von Strukturinformationen werden im Projekt möglichst hochauflösende Fernerkundungsdaten verwendet. Dabei handelt es sich um digitale True-Orthophotos (tDOP, verzerrungs- und verkipprungs-korrigierte Luftbildaufnahmen). Diese sehr lagegenauen, amtlichen Befliegungsdaten der bayerischen Vermessungsverwaltung werden im 2-Jahres-Turnus aktualisiert und liegen für ganz Bayern mit 0,2 m räumlicher Auflösung in 4-Kanälen (Rot, Grün, Blau (RGB) + nahes Infrarot) vor (LDBV Bayern 2023). Die Abdeckung Bayerns beruht dabei auf einer Vielzahl von Bildflügen zu verschiedenen Zeitpunkten, wodurch Inhomogenitäten im Bildmaterial aufgrund verschiedener Beleuchtungsbedingungen entstehen. Alternative Daten von satellitengestützten Sensoren mit Auflösungen von < 1 m (z. B. WorldView 2/3, SkySAT) wurden ebenfalls in Betracht gezogen, schied jedoch wegen starker Verkipprung der Bäume durch Schrägaufnahme, hoher Lagegenauigkeit, mangelnder Verfügbarkeit sowie hoher Kosten aus.



Mögliche klimatische und geographische Einflussfaktoren, die Phänologie und sonstige Heterogenität der Eingangsdaten müssen im Rahmen der Baumarterfassung bedacht und in den zur Modellierung genutzten Referenzdaten abgebildet sein. Dies bedingt, dass die Referenzdaten bestmöglich über Bayern verteilt erfasst werden müssen. Für deren Erfassung wurden terrestrische Inventurdaten aus der Bundeswaldinventur (BWI, 2012) und der jährlich durchgeführten Waldzustandserhebung sowie hochaufgelöste Drohnenaufnahmen aus weiteren Projekten an der LWF als Hilfestellung verwendet.

Insgesamt wurden 809 über Bayern verteilte Traktecken von BWI-Inventurpunkten zur Ableitung von Referenzdaten gewählt. Um jede Traktecke wurde eine Quadratfläche von 50 m x 50 m definiert. Entsprechend dieser Auswahl fanden tDOP aus den Jahren 2017/18/19 sowie 2021/22 Anwendung, die vorwiegend zwischen Mai und Juli aufgezeichnet wurden (Abbildung 1). Innerhalb jedes Quadrates wurden alle Bäume mit einer Mindesthöhe von 12 m durch visuelle Interpretation angesprochen, im GIS mit einem Punkt markiert und einer Klasse zugewiesen (Abbildung 1).

Bäume unter 12 m Höhe wurden von der Erfassung ausgeschlossen, da eine verlässliche Bestimmung nicht möglich war. Die Höheninformation stammt dabei aus einem amtlichen, normalisierten Digitalen Oberflächenmodell (nDOM) mit 1 m räumlicher Auflösung.

Die Qualität der erfassten Baumpunkte wurde durch Mehrfachbetrachtung im 4-Augen-Prinzip und Attributierung mit sicher/unsicher (ca. 5%) verbessert. Insgesamt wurden ca. 103.000 Baumpunkte (97.200 sicher) gesetzt, wobei die Fichte mit 37,6% die häufigste Baumart in den Referenzdaten ist.

Die gewählten DL-Methoden benötigen flächige Labels, welche die den Baumpunkten zugehörigen Kronenflächen abbilden. Durch Bildsegmentierung wurde versucht, diese flächigen Abgrenzungen automatisiert abzuleiten. Die Ergebnisse stellten sich jedoch besonders in heterogenen Beständen als unzuverlässig heraus; sie wiesen Über-/Untersegmentierung und fehlende Abgrenzungen der Baumarten auf. Eine nachträgliche Korrektur der Segmentierung stellte sich gegenüber einer vollständig manuellen Erfassung der Baumkronen als aufwendiger heraus, so-

dass die flächigen Label auf Basis der RGB-tDOP manuell erzeugt wurden. Da semantische Segmentierung (siehe Folgeabsatz) keine Einzelobjekte betrachtet, konnten dabei aneinandergrenzende Kronen einer Baumart zusammengefasst werden.

Zur späteren Bewertung der Modellierung wurden 10% der Referenzdaten als Testdatensatz vom Modelltraining und -tuning ausgeschlossen. Dabei wurde sichergestellt, dass der Testdatensatz bezüglich Baumartenverteilung, Aufnahmezeitpunkt und räumlicher Verteilung repräsentativ ist. Testdaten wurden vom Modelltraining ausgeschlossen, sodass eine Vorhersage und anschließende Analyse auf dem Modell mit unbekanntem Daten erfolgt.

### Auswahl des passenden Deep Learning Modellansatzes

In den letzten Jahren haben sich Ansätze des DL, speziell sogenannte Convolutional Neural Networks (CNN), in der Erdbebobachtung etabliert. Bemerkenswert sind deren deutlich größere Anzahl an Modellparametern gegenüber herkömmlichen Verfahren des ML sowie die Berücksichtigung lokaler Nachbarschaften der Bildpixel. So fließen in den Klassifikationsprozess neben Spektralinformation (Farbe) auch räumliche Strukturen und Texturen der Bildinhalte ein. Da neben der Erkennung der Baumarten deren exakte räumliche Abgrenzung angestrebt wird, fiel die Wahl auf einen Ansatz der semantischen Segmentierung. Dabei wird jedes Pixel des Eingangsbildes der entsprechenden Klasse zugeordnet. Eine Unterscheidung einzelner Baumindividuen (wie bei Instanzsegmentierung) ist nicht möglich.

Im Rahmen der Modelloptimierung wurden verschiedene Modellarchitekturen verglichen (U-Net, U-Net3P) und Modellparameter, beispielsweise Klassengewichtungen, getestet.

Untersucht wurde auch der Einfluss zusätzlicher Eingangsdaten (z. B. Baumhöhen) oder von Bildaugmentierung. Bei letzterer werden die Originalbilder geometrisch (z. B. Drehen, Spiegeln) oder inhaltlich (z. B. Hinzufügen von Rauschen) manipuliert, um die Robustheit des Modells zu erhöhen und den Datensatz zu vergrößern.

Die Rechendauer eines Trainings betrug je nach Menge der Eingangsdaten ca. 10–12 Stunden auf einer handelsüblichen Grafikkarte. Umfangreiche Hyperparameter Tuning-Experimente dauerten mehrere Tage. Die erzeugten Modelle wurden anschließend anhand von Validierungs- und Testdaten bewertet und das beste Modell ermittelt. Dabei kamen etablierte statistische Metriken wie der F1-Score sowie die Analyse der Fehlermatrix zum Einsatz.

### Die Kombination macht das Rennen

Mit der systematischen Validierung der Trainingsergebnisse konnte ermittelt werden, welche Kombinationen von Parametern und Eingangsdatensätzen zur Verbesserung des DL-Modells führen und schlussendlich das beste Modell bilden. Eine Auswahl relevanter Maßnahmen mit dem Ziel der Verbesserung des Modells zeigt Abbildung 2. Im Gegensatz zu den Erwartungen führten nicht alle Maßnahmen zu einer signifikanten Verbesserung.

Parameter / Eingangsdatensatz	Einfluss	Ziel der Maßnahme
Bildaugmentierung	++	Menge der Trainingsdaten künstlich vergrößern
nDOM	++	Baumhöhe als zusätzliche Information nutzen
Klassengewichte	+	Gewichtung der Klassen beim Training (Verlustfunktion) soll Ungleichverteilung der Trainingsdaten ausgleichen
Vergrößerung des Trainingsdatensatzes	+	Trainingsdatensatz vergrößern anhand zusätzlicher tDOP derselben Gebiete aus anderen Zeitpunkten
Ausschluss schwer interpretierbarer Bäume	+	Schwer erkennbare Bäume aus dem Trainingsdatensatz ausmaskieren (Klasse »unsicher«)
Abgrenzung der Baumkrone durch automatische Segmentierung	o	Baumkronen durch einen Segmentier-Algorithmus im Vergleich zur manuellen Abgrenzung genauer abgrenzen.
Entfernung von Daten mit schlechter Bildqualität	o	Verbesserung durch höheren Anteil an Bildern mit hoher Bildqualität, muss gegen die Verkleinerung des Datensatzes abgewogen werden.

2 Empirisch ermittelter Einfluss von Parametern und Eingangsdatensätzen auf die Genauigkeit des Modells

Einzelklasse		Kombinierte Klasse	
Buche	0,51	0,83	Laubholz
Eiche	0,21		
Sonst. Laubholz	0,38		
Kiefer	0,75	0,77	Kiefer
Fichte	0,67	0,76	Fichte und sonst. Nadelholz
Sonst. Nadelholz	0,30		
Totholz	0,41	0,41	Totholz
Hintergrund	0,67	0,67	Hintergrund
Gesamt	0,57	0,76	Gesamt

3 Pixelbasierte F1-Werte der Baumarten. Der F1-Score bildet das harmonische Mittel aus Precision und Recall und kombiniert damit Genauigkeit und Vollständigkeit in einer einzigen Metrik; Der Ergebniswert liegt zwischen 0 und 1, wobei höhere Werte besser sind.

Die erreichten Genauigkeiten (F1-Scores) des besten Modells sind in Abbildung 3 dargestellt. Die höchsten Genauigkeiten erreichen die Klassen Kiefer und Fichte mit einem F1-Score von 0,75 und 0,67. Neben dem überdurchschnittlichen Anteil am Trainingsdatensatz ist dies auch auf deren markante und gut erkennbare Struktur in den Bilddaten zurückzuführen. Sehr niedrige Werte erzielt dagegen die Eiche (0,21), was mit der vergleichsweise geringen Menge an Trainingsdaten dieser Klasse und ihrer geringen Differenzierbarkeit gegenüber anderen Laubholzklassen zu erklären ist. Ferner fiel auf, dass es vor allem innerhalb der Laub- und Nadelgehölze zu Verwechslungen kommt, weshalb ein weiteres Training mit zusammengefassten Klassen durchgeführt wurde. Mit der Zusammenfassung einzelner Klassen konnte insbesondere für das Laubholz eine deutliche Verbesserung der Genauigkeitswerte erreicht werden. Das zeigt erneut die Schwierigkeit der Differenzierung innerhalb der Laubholzklassen und die gute Trennbarkeit zum Nadelholz. Somit eignet sich das Modell mit den kombinierten Klassen als pixelgenauer Laub-/Nadelholz-Klassifikator und erreicht für die Klassen Laubholz einen F1-Score von 0,83, die weiteren Klassenergebnisse sind Abbildung 3 (rechts) zu entnehmen. Einen visuellen Vergleich zwischen Ergebnissen der Einzelbaumartenklassen und kombinierten Klassen gibt Abbildung 4. Neben der pixelbasierten Validierung, welche die korrekte Abgrenzung der Bäume berücksichtigt, wurden die Ergebnisse des Modells auch punktbasiert ausgewertet. Damit soll gezeigt werden, ob der Kernbereich eines Einzelbaumes, unabhängig von dessen Abgrenzung im Rand-

bereich, korrekt erkannt wurde. Für die Kombinationen Laubholz, Kiefer sowie Fichte/sonstiges Nadelholz konnten F1-Scores von größer als 0,8 erreicht werden. Die zusätzlich erhobene Klasse Totholz weist geringe Genauigkeitswerte auf, was mit der ausgeprägten Heterogenität der Klasse begründet werden kann (geschädigt, abgestorben).

### Fazit und Ausblick

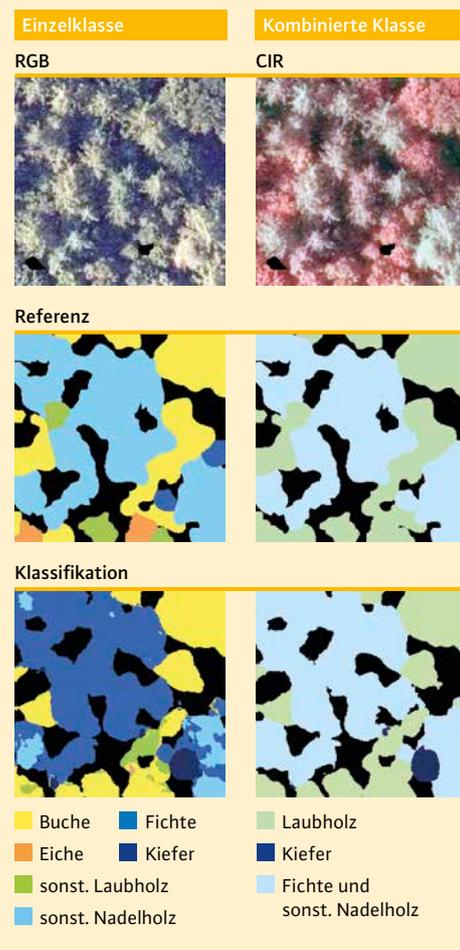
In den eingangs erwähnten Projekten BeechSAT und IpsSAT erzielte DL gegenüber ML bessere Ergebnisse bei der Unterscheidung zwischen gesunden und geschädigten oder abgestorbenen Bäumen. Jedoch war dabei die Klassenanzahl gering und die Klassen deutlich abgrenzbar. In unserer Studie zeigte die hohe Anzahl an Klassen, insbesondere die der Laubhölzer, nur geringe spektrale und strukturelle Differenzen, sodass diese auch durch Nutzung von DL nicht zuverlässig zu trennen waren. Eine Unterscheidung von Laub- und Nadelhölzern war jedoch zuverlässig möglich.

Eine Analyse der Verwechslungen zwischen diesen beiden Klassen zeigte den großen Einfluss des Bildmaterials auf, da hier Artefakte (künstliche Strukturen wie Kanten oder Verzerrungen aus der Orthorektifizierung) häufig waren und somit offenbar einen direkten Einfluss haben.

Einen großen Mehrwert besitzen die auf den tDOP basierenden Referenzdaten (hohe Lagegenauigkeit). Diese können mit geringem Aufwand aktualisiert und zur Validierung von weiteren Fernerkundungsprodukten (z. B. Baumartengruppenkarte Thünen-Institut (Blickensdorfer et al. 2023)) verwendet werden. Eine Anwendung des Modells auf größerer Fläche ist derzeit in Bearbeitung, um den

### Zusammenfassung

Das Ziel des Projekts KIHBA ist eine automatisierte Klassifizierung der Hauptbaumarten Buche, Eiche, Fichte, Kiefer auf Einzelbaumebene durch Nutzung von Deep Learning Verfahren und hochaufgelöster Fernerkundungsdaten. Die genutzte Modellarchitektur ist ein U-Net. Die Ergebnisse zeigen die Grenzen der Methodik, da insbesondere Laubhölzer nicht verlässlich unterschieden werden können. Eine klassenscharfe Trennung zwischen den aggregierten Klassen »Laubholz«, »Kiefer« und »Fichte/sonst. Nadelholz« zeigte sich jedoch als möglich. Es besteht ein fortschreibbarer, großer Referenzdatensatz und die Klassifizierung kann flächig auf Bayern angewendet werden.



4 Ansicht des tDOP in RGB und CIR (NIR, R, G) an einer Trakteecke. Darunter erfasste Referenzlabels und Klassifikationsergebnisse

Einfluss von räumlichen Strukturen, das Auftreten möglicher Randeffekte und Berechnungszeiten der Vorhersage zu evaluieren. Bei Erfolg wäre es möglich, eine flächenscharfe Baumartengruppenkarte in 0,2 m räumlicher Auflösung erzeugen zu können, die bezogen auf die Pixelauflösung eine 100-Fach höhere Auflösung besitzt als die bereits an der LWF vorliegenden, auf Sentinel-2-Daten basierenden Baumartengruppenkarten.

### Literatur

Das Literaturverzeichnis finden Sie unter [www.lwf.bayern.de](http://www.lwf.bayern.de) in der Rubrik »Publikationen«.

### Autoren

Dr. Adelheid Wallner und Steffen Günster sind Mitarbeitende und Rudolf Seitz Leiter der Abteilung »Informationstechnologie« der Bayerischen Landesanstalt für Wald und Forstwirtschaft (LWF). Bettina Felten, Nicole Schmeitzner und Peter Schauer sind Mitarbeitende der IABG mbH Geospatial Solutions in Dresden.  
**Kontakt:** [adelheid.wallner@lwf.bayern.de](mailto:adelheid.wallner@lwf.bayern.de)

### Projekt

Das Projekt »KIHBA« wird durch das BMWK gefördert (Laufzeit 1.5.2021–31.1.2024) und in Kooperation mit IABG mbH Geospatial Solutions durchgeführt.

