

Identifizierung von *Senecio jacobaea* L. im Bestand mit Hilfe unbemannter Luftfahrzeuge

PHILIPP ZACHARIAS ¹

Zusammenfassung: folgt

1 Einleitung

Die Greiskraut-Art *Senecio jacobaea* L. (S.j., auch Jakobskreuzkraut) breitet sich derzeit in Deutschland stark aus und stellt neue Herausforderungen an die extensive Grünlandnutzung. Die Pflanze enthält giftige Pyrrolizidinalkaloide (PA), auf die besonders Pferde und Rinder empfindlich reagieren. Um der Etablierung von S.j. auf Grünlandstandorten wirksam vorbeugen zu können, ist es überaus wichtig die Pflanze möglichst früh bei der Einwanderung aus den Randbereichen zu detektieren. Die Verdrängung auf bereits befallenen Flächen braucht verlässliche Informationen über den Zustand der Flächen. Ein Einsatz teilflächenspezifischer Applikationen von Düngemitteln oder auch mechanischer Maßnahmen basiert auf digitalen Befallskarten. Diese sind nur dann effektiv, wenn sie die zu bekämpfenden Arten in möglichst hoher Auflösung darstellen. Daher ist die Identifikation und Lokalisation der einzelnen S.j. von großem Interesse für die Landwirtschaft.

Unmanned Aerial Vehicles (UAV) können flexibel mit Sensorik ausgestattet werden und bieten so die Möglichkeit, mit geringem Kostenaufwand, schnell und vielseitig Befliegungen über dem Gelände durchzuführen. So können in niedriger Flughöhe hochgenaue Bildaufnahmen von landwirtschaftlichen Flächen und einzelnen Pflanzen erzeugt werden. Diese können durch verschiedene photogrammetrische und statistische Verfahren zu hochauflösenden räumlichen Informationen prozessiert werden. Dadurch wird die gezielte Suche nach unerwünschten Arten, wie etwa S.j. ermöglicht. Die UAV sind daher in der Lage einen wichtigen Beitrag zum standortspezifischen Grünlandmanagement zu leisten.

Ziel war es eine Bildverarbeitungsmethode zu entwickeln, welche Jakobskreuzkraut im Grünland kostengünstig, anhand hochauflösender Bilddaten einer UAV-Befliegung identifizieren kann. Als Ergebnis soll die Position der einzelnen Pflanzen erfasst und für die pflanzenbaulichen Maßnahmen zur Verfügung gestellt werden.

2 Methode

Die Untersuchungen fanden auf einer mit Pferden beweideten Grünlandfläche statt, welche sich in der Nähe von Bad Doberan in Mecklenburg-Vorpommern befindet. Drei UAV-Bildflüge wurden mit einem Falcon 8 von Intel durchgeführt, welche mit einer 36,4 Megapixel Digitalkamera ausgestattet war. Im Fokus der Analyse liegt die Detektion von S.j. zur Blütezeit, daher soll hier

¹ Universität Rostock, Agrar- und Umweltwissenschaftliche Fakultät, Professur für Geodäsie und Geoinformatik, Justus-von-Liebig Weg 6, D-18059 Rostock, E-Mail: philipp.zacharias@uni-rostock.de

nur auf den letzten Bildflug zur Blüte eingegangen werden. Aus einer Flughöhe von etwa elf Metern wurden dabei 69 Aufnahmen angefertigt (Bildüberlappung längs 80%, quer 60%). Vor der Befliegung wurden acht Passpunkte gleichmäßig auf der Fläche verteilt und mit einem RTK-GNSS eingemessen (Lagegenauigkeit circa 2 cm).

Um die in den Bildern enthaltenen räumlichen Informationen zu rekonstruieren, wurde die Software Agisoft Photoscan Professional verwendet (Image Alignment, Dense Matching und Mesh Building). Dazu werden die gewonnenen Bilder, die GNSS-Koordinaten zum Auslösezeitpunkt der Bilder sowie die GNSS-Koordinaten der Passpunkte verwendet. Es wurde ein Orthophoto und ein digitales Oberflächenmodell (DOM) gespeichert. Die effektive Bodenauflösung des Orthophotos liegt bei 1,38 Millimetern pro Pixel. Die Genauigkeit des Rasters im Pixelverband (Root-Mean-Square Reprojection Error) beträgt 0,23 Pixel. Die Bodenauflösung des DOM beträgt 2,75 mm/pixel.

Zum Zeitpunkt des Fluges am fünften Juli konnten einige Hotspots mit blühendem S.j. aufgenommen werden. Die überwiegende Mehrheit dieser Pflanzen befand sich in der Vollblüte (BBCH-Stadium 64). Im Orthophoto wurden zunächst alle tatsächlich vorhandenen S.j.-Pflanzen visuell identifiziert und in Punktform digitalisiert (Abbildung 1). Eine überwachte Klassifikation wurde verwendet, um alle gelben Bildelemente herauszufiltern. Dazu werden im GIS über dem Orthophoto Trainingspolygone angelegt, die voneinander unterscheidbare Bereiche auf dem Orthophoto in Klassen getrennt spektral beschreiben. Anschließend wird das Orthophoto automatisch nach den definierten Bereichen abgesucht und jedem Pixel wird eine der Klassen zugeordnet. Danach werden alle Klassen außer der Blütenklasse aus dem Raster entfernt sowie Pixelrauschen beseitigt. Schließlich erfolgt eine Konvertierung des Rasters in ein Vektorformat. Unter den so gewonnenen 4669 Polygonen befinden sich nun die gesuchten S.j.-Pflanzen sowie weitere, gelb blühende Pflanzen, Laub und andere gelbliche Objekte. Daher muss ein zweiter Klassifizierungsschritt erfolgen, welcher die vorhandenen S.j. selektiert. Für diese Aufgabe wurden drei statistische Modelle mit Hilfe der Statistik-Software SPSS getestet: Ein neuronales

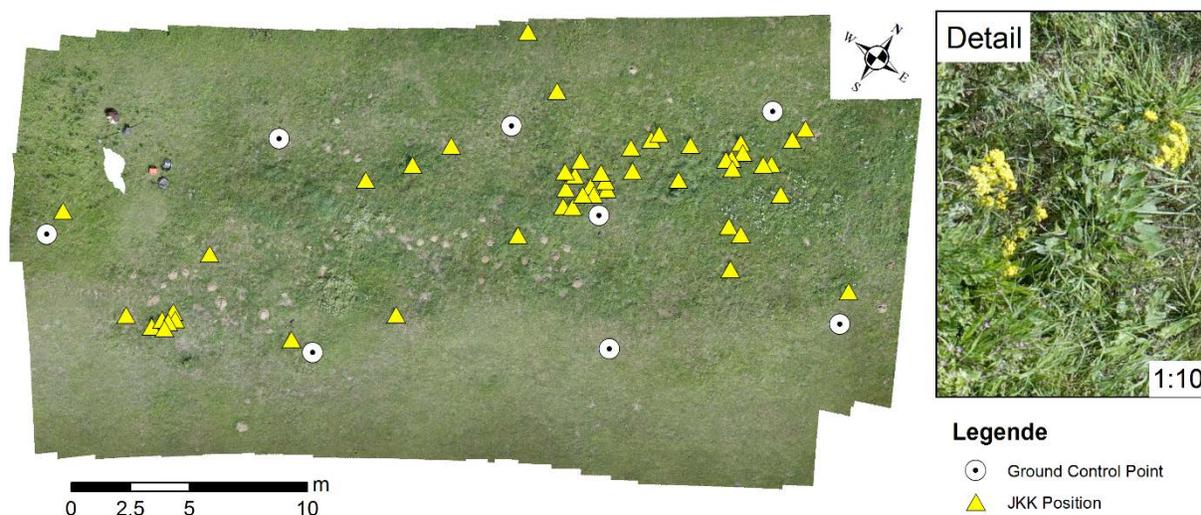


Abb 1: Orthophoto, als Ergebnis der photogrammetrischen Verarbeitung der UAV-Bilddaten. Außerdem abgebildet, Positionen der Ground Control Points und die digitalisierten Positionen von *Senecio jacobaea* L.

Netz (NN) vom Typ „Multilayer Perceptron“, eine logistische binäre Regression (LBR) sowie eine Diskriminanzanalyse (DA). Um einen Trainingsdatensatz zu erzeugen, wurden alle Polygone manuell als „S.j.“- oder „Nicht-S.j.-Polygon“ klassifiziert. Als Modellparameter dienen RGB-Werte aus dem Orthophoto, Wuchshöhen aus dem DOM, sowie verschiedene Werte, die aus den Polygonen abgeleitet wurden (Tabelle 1). Der Datensatz wurde randomisiert und zu je 50% in eine Trainings- und eine Teststichprobe aufgeteilt. Durch die Randomisierung wird sichergestellt, dass die Stichproben räumlich zufällig verteilt sind. Die drei Modelle wurden jeweils mit derselben Stichproben-Einteilung gestartet. Zuvor wurden die Werte standardisiert. Die statistischen Modelle treffen für jedes Objekt eine Ja/Nein-Entscheidung, also ob es sich dabei um S.j. handelt oder nicht.

Tab 1: Parameter für die statistische Klassifikation

Messquelle	Parameter im GIS
Polygone	Umfang, Fläche, Shape-Index, Breite und Länge der minimalen Begrenzungsgeometrie
Orthophoto	Minimum, Maximum, Mittel, Std-Abw, Range, Varietät, Median, Minderheit und Mehrheit der RGB-Werte
Wuchshöhe	Minimum, Maximum, Mittel, Std-Abw und Range der Höhenwerte

Mit der Diskriminanzanalyse wurde vor den Analysen zunächst eine Variablenselektion durchgeführt. Bei diesem Vorgang werden die Variablen schrittweise hinzugefügt und nur dann beibehalten, wenn sie die Unterscheidung zwischen den beiden Gruppen der abhängigen Variablen (S.j. „ja“ oder „nein“) verbessern. Dabei wurden 7 von den 19 Variablen aussortiert. Dies sind Polygon-Umfang, Range, Mittel und Standardabweichung der Höhenwerte, sowie Minimum, Minderheit und Mittel der RGB-Werte. Damit wurde keine der Informationsquellen (Orthophoto, DOM oder Polygone) vollständig ausgeschlossen. Dies deutet darauf hin, dass in jeder Quelle wichtige Informationen für die Klassifizierung enthalten sind.

Da die tatsächliche Klassifizierung aller Objekte bekannt ist, können nun die Fälle der Wahrheitsmatrix berechnet werden: True Positives (TP), False Negatives (FN), False Positives (FP) und True Negatives (TN). Diese ermöglichen den Vergleich der drei Modelle.

3 Ergebnisse und Diskussion

Die Wahrheitsfälle der drei Modelle, gemessen an der Teststichprobe, sind in Tabelle 2 aufgeführt. Bei dieser Untersuchung sind die Wahrheitsfälle folgendermaßen zu interpretieren: (i) Korrekt als S.j. klassifizierte Objekte (TP), (ii) S.j.-Objekte die fälschlicherweise als Nicht-S.j.-Objekte klassifiziert wurden (FN bzw. „Fehler erster Art“), (iii) Nicht-S.j.-Objekte die fälschlicherweise als S.j.-Objekte klassifiziert wurden (FP bzw. „Fehler zweiter Art“), (iv) Korrekt als Nicht-S.j. klassifizierte Objekte (TN).

Tab 2: Wahrheitsfälle für die drei angewandten statistischen Modelle, berechnet aus der Teststichprobe.

Modell	TP		FN		FP		TN	
	insg.	% ⁴	insg.	% ⁴	insg.	% ⁵	insg.	% ⁴
NN¹	143	61,6	89	38,4	51	2,4	2052	97,6
LBR²	117	50,4	115	49,6	54	2,6	2049	97,4
DA³	162	69,8	70	30,2	185	8,8	1918	91,2

¹Neuronales Netz; ²Diskriminanzanalyse; ³Logistische Binäre Regression⁴prozentualer Anteil an allen Jakobskreuzkraut-Objekten in der Teststichprobe⁵prozentualer Anteil an allen Nicht-Jakobskreuzkraut-Objekten in der Teststichprobe

Tab 3: Bewertungsmaße für die drei angewandten statistischen Modelle, berechnet aus der Teststichprobe

Maß	NN	LBR	DA
False-Positive-Rate	0,024*	0,026	0,088
Precision	0,74*	0,68	0,47
True-Positive-Rate	0,62	0,50	0,70*
Accuracy	0,94*	0,93	0,89
F-Score	0,45*	0,35	0,33

*bester Wert unter den drei getesteten Modellen.

In der Teststichprobe befanden sich 232 S.j.-Objekte (TP + FN) und 2103 Nicht-S.j.-Objekte (FP + TN). Die Prozentangaben in Tabelle 2 beziehen sich jeweils auf diese Teilsummen, während sich die Gesamtwerte auf alle Objekte der Teststichprobe beziehen. Ein optimales Ergebnis wäre die korrekte Erkennung aller S.j.-Objekte als TP und aller Nicht-S.j.-Objekte als TN. Um die räumliche Verteilung der Wahrheitsfälle zu visualisieren, wurde für das Ergebnis jedes Modells eine Kartendarstellung angefertigt (Abb 2). Es fällt auf, dass die überwachte Klassifikation sehr viele Objekte im westlichen Bereich der Fläche erzeugt hat, wo sich laut Referenz kein S.j. befindet. Diese stellen sich nun als eine Ansammlung vieler TN in diesem Bereich dar, da der Großteil von den Modellen korrekt aussortiert wurde.

Das neuronale Netz und die Diskriminanzanalyse erkennen S.j. beide recht gut (TP 61,6% bzw. 69,8%), während die Regression hier schlechter abschneidet (50,4%). Allerdings ist der Anteil der FP bei der Diskriminanzanalyse mit 8,8% viel höher als bei den anderen beiden Modellen. Dadurch wird die Qualität der durch die Diskriminanzanalyse getroffenen Vorhersage stark herabgesetzt. Das neuronale Netz ist am besten im Erkennen von Nicht-S.j.-Objekten, was an der hohen Zahl der TN erkennbar ist. Hier schneidet die Diskriminanzanalyse am schlechtesten ab.

Um die Wahrheitsfälle zusammenzufassen wurden Bewertungsmaße nach FAWCETT (2006) berechnet (Tabelle 3). Dies sind False-Positive-Rate, Precision, True-Positive-Rate, Accuracy und F-Score. Die Maße bestätigen die oben gewonnenen Erkenntnisse und heben das neuronale Netz

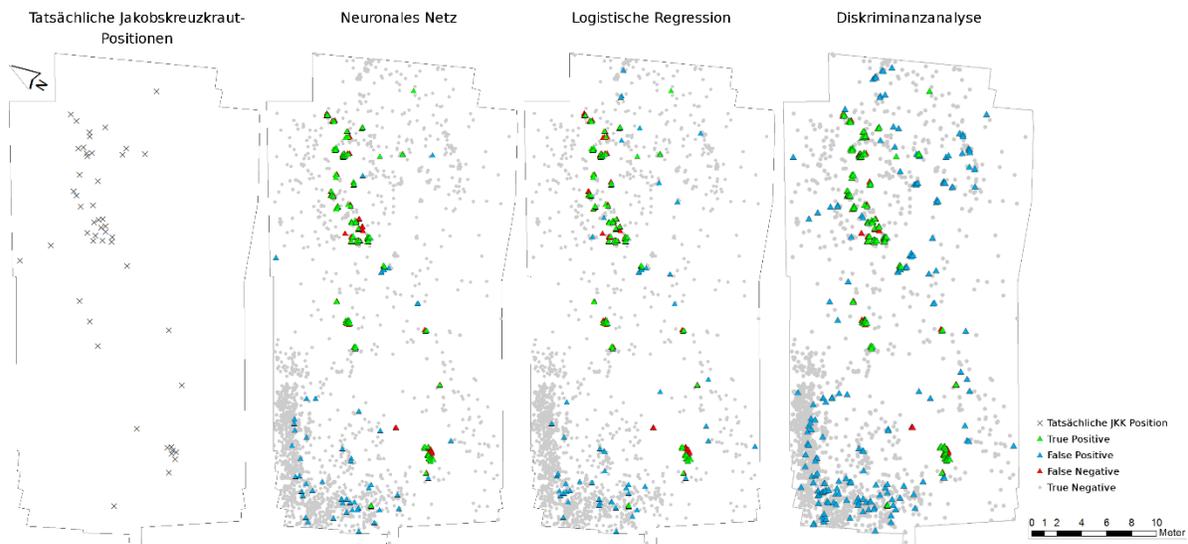


Abb 2: Kartendarstellung der tatsächlichen Jakobskreuzkraut-Positionen und den Ergebnissen der drei statistischen Modelle

eindeutig als Favoriten hervor. Die Regression nimmt eine Mittelposition ein, ist aber durch den recht hohen FN-Anteil kritisch zu betrachten. Die Diskriminanzanalyse kann anhand der Werte, wie auch mit Blick auf die Kartenansicht, als schlechtestes der drei getesteten Modelle angesehen werden. Die besten Klassifikations-Ergebnisse liefert das neuronale Netz. Dies ist auch anhand der Karten ersichtlich, da vor allem im Hotspot-Bereich im Norden der Fläche vergleichsweise wenig FP auftreten. Hinsichtlich des Anteils der FN befindet sich das neuronale Netz zwischen den beiden anderen Modellen.

4 Fazit & Ausblick

Die Ergebnisse der Klassifizierung zeigen, dass mit der entwickelten Methode eine Erkennung von S.j. zur Vollblüte (etwa BBCH-Stadium 64) grundsätzlich möglich ist. Die berechneten Bewertungsmaße der Ergebnisse deuten darauf hin, dass neuronale Netze besonders geeignet sind, um die Klassifizierung der Objekte vorzunehmen. Im Test zeigte es die besten Werte für Accuracy (Genauigkeit der Ergebnisse) und Precision (Reproduzierbarkeit der Ergebnisse). Dies zeigt, dass die Methode vergleichsweise genaue Ergebnisse bei geringer Überanpassung an die untersuchte Fläche liefert. Eine Bodenauflösung von 1,38 Millimetern pro Pixel im Orthophoto war ausreichend, um 61,6% der Jakobskreuzkraut-Objekte zu erkennen. Unter Einbeziehung der korrekt erkannten Nicht-S.j.-Objekte ergibt sich eine Erkennungsrate von 94%. Dies bestätigt sich auch durch die visuelle Kontrolle der Vorhersagen mit kartographischen Darstellungen (Abb 2).

Die vielen Nicht-S.j.-Objekte im westlichen Bereich der Fläche könnten durch weitere Bearbeitung des Ergebnisses der überwachten Klassifikation reduziert werden. Es handelt sich vorrangig um Objekte, die kleiner sind als typische S.j.-Blüten. Daher würden diese sich leicht durch Angabe eines minimalen Wertes für die Polygonfläche aus dem Datensatz entfernen lassen. Dadurch würden die Modelle wahrscheinlich bessere Ergebnisse liefern. Allerdings ist mit steigender Fläche der Übergang zu S.j.-Objekten fließend. Es besteht die Gefahr, dass dadurch viele noch nicht voll aufgegangene Blüten, die dementsprechend kleiner sind, frühzeitig aus dem Datensatz entfernt werden.

Auf einer Grünlandfläche mit stärkeren Bewuchs der Begleitvegetation wäre die Erkennung vermutlich schwieriger, da mehr gelbe Blüten von anderen Pflanzenarten hinzukommen würden. Dadurch würde sich die Zahl der Fehler zweiter Art vermutlich erhöhen. Wie die Ergebnisse bei Anwendung der gezeigten, statistischen Methode auf anderen Flächen ausfallen, soll in zukünftigen Arbeiten untersucht werden. Auf neuen Flächen kann theoretisch auch ohne eine manuell erstellte Referenz-Shapedatei eine Klassifizierung vorgenommen werden. Dafür müssen aber ausreichend Trainingsdaten aus vorherigen Untersuchungen vorliegen. Die statistischen Modelle können dann so konfiguriert werden, dass sie die neuen Objekte mit dem zuvor erlernten „Wissen“ klassifizieren. Um diesen Vorgang effektiv zu gestalten, kann eine Datenbank angelegt werden, die umfangreiche Trainingsdaten enthält. Auf diese können die statistische Modelle dann zugreifen, um ihr Training durchzuführen und neue Objekte zu klassifizieren. Auch DAMMER ET AL. (2013) empfehlen die Pflege einer spektralen Datenbank, welche die Wachstumsstadien der Pflanzen berücksichtigt.

Die Identifikation von S.j. vor der Blüte, also im Rosettenstadium, ist Schwerpunkt eines aktuellen Forschungsprojektes an der Professur für Geodäsie und Geoinformatik der Universität Rostock.

5 Literaturverzeichnis

- DAMMER, K.-H.; INTREB, J. & USTYUZHANIN, A., 2013: Hyperspektrale Bildanalyse zur Unterscheidung von *Ambrosia artemisiifolia* und *Tagetes* ssp. *Ambrosia* in Deutschland - lässt sich die Invasion aufhalten? Tagung vom 10. - 12.09.2013 in Berlin, Starfinger, U.; Sölter, U.; Verschwele, A. (Hrsg.), Bundesforschungsinstitut für Kulturpflanzen (JKI) (1. Juli 2014), 131-138.
- FAWCETT, T., 2006: An introduction to ROC analysis. *Pattern Recognition Letters*, **27** (8), 861-874.