

MARCO SEELAND, DAVID BOHO, MICHAEL RZANNY, MARTIN HOFMANN, NEDAL ALAQRAA,
JANA WÄLDCHEN & PATRICK MÄDER

Flora Incognita – Bilder für das Trainieren neuronaler Netzwerke gesucht

Im Forschungsprojekt Flora Incognita werden Bilderkennungsverfahren verwendet, um die Bestimmung der Pflanzenarten Thüringens zu vereinfachen und Laien ein leichtes Erkennen im Gelände zu ermöglichen. Das Projekt

selbst wurde bereits in dieser Zeitschrift detaillierter vorgestellt (WÄLDCHEN et al. 2016). Die Artenbestimmung basiert auf neuesten Methoden des maschinellen Lernens, sogenannten tiefen neuronalen Netzwerken.

Diese Netze sind dem menschlichen Gehirn nachempfunden und bestehen aus einer Vielzahl von kaskadierten Ebenen mit Filtern und mehreren Millionen Neuronen. Für das Trainieren der Netzwerke werden riesige Mengen

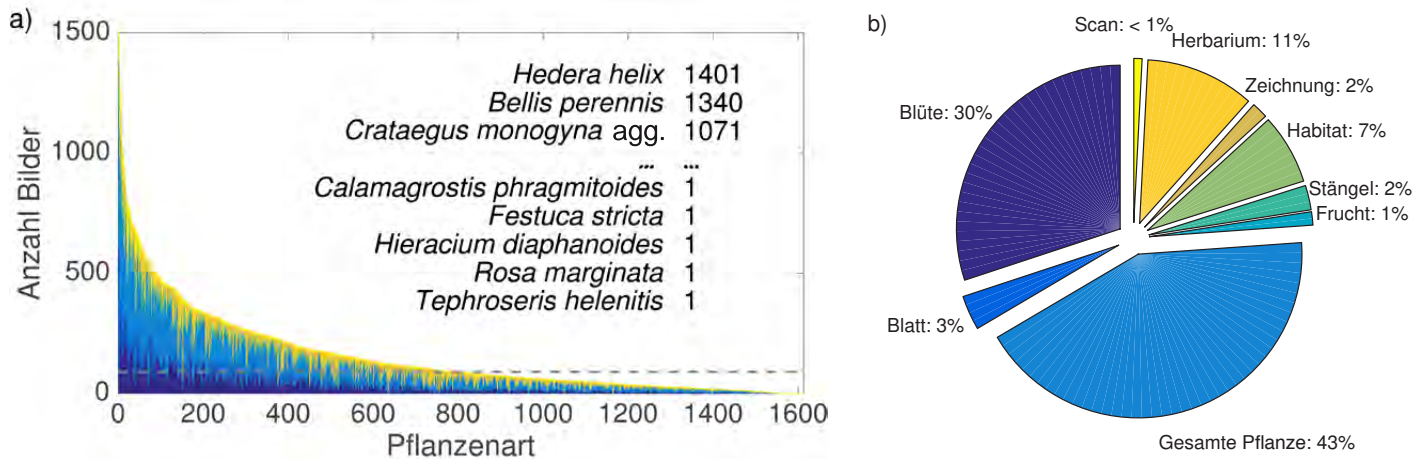


Abb. 1: a) Anzahl der Bilder je Pflanzenart; b) Häufigkeitsverteilung der Bildinhalte über den gesamten Datensatz.

von Beispielbildern der unterstützten Arten benötigt.

Unser aktueller Datensatz umfasst etwa 250.000 Fotos von 1.600 wildwachsenden Blütenpflanzen in Thüringen. Die Arten sind jedoch sehr ungleichmäßig im Datensatz vertreten. Abbildung 1a veranschaulicht die Anzahl der Fotos je Pflanzenart; der Median (gestrichelte Linie) liegt bei 90 Fotos. Abbildung 1b zeigt die Verteilung der Pflanzenperspektiven über den gesamten Datensatz. Mit 43% stellt der Großteil der Bilder den Habitus der Pflanze dar. Derartige Bilder ermöglichen eine grobe automatische Erkennung, für eine Artbestimmung sind jedoch Detailaufnahmen der einzelnen Organe notwendig. Aufnahmen der Blüte sind mit 30% sehr gut vertreten, die der Blätter hingegen nur mit 3%. Für eine genauere automatische Erkennung braucht es jedoch auch Blattmerkmale, wie eine unserer aktuellen Studien bereits andeutet.

Mit dem bisherigen Datensatz erreichen wir eine Erkennungsrate von 75,8%. Es hat sich gezeigt, dass diese Erkennungsrate stark von Anzahl und Informationsgehalt der Bilder je Pflanzenart abhängt. Je weniger Bilder vorliegen, umso schlechter kann das Aussehen einer Art gelernt werden. Als Bildquellen binden wir unter anderem Online-Bilddatenbanken wie „Wikimedia Commons“ und „Encyclopedia of Life“ ein. Die Qualität der enthaltenen Bilder variiert sehr stark und ist ein wesentlicher Grund für die noch nicht zufriedenstellenden Erkennungsraten. Kritische Punkte bei den verfügbaren Bildern umfassen:

- Unpräzise Aufnahmen: Viele Bilder sind aus zu großer Entfernung aufgenommen, sodass die relevante Art nicht mehr eindeutig zu erkennen ist.
- Unvollständige Perspektiven: Eine Erkennung wird umso genauer, je mehr Bildperspektiven der gleichen Pflanze verfügbar sind, die Datenbanken beinhalten aber häufig nur wenige Perspektiven einer Art.
- Fehlerhafte Bestimmungen: In stichprobenartigen Untersuchungen haben wir durchschnittlich 5% Fehlbestimmungen in den vorhandenen Daten identifiziert.

Um die Datengrundlage für unsere Bilderkennungstechniken zu verbessern, haben wir die App „Flora Capture“ entwickelt und setzen auf interessierte Enthusiasten, welche Pflanzen mit ihrem Mobiltelefon nach vorgegebenen Kriterien fotografieren. Viele der jetzt vorhandenen Funktionalitäten werden in der finalen Anwendung enthalten sein und können so getestet und verbessert werden. Für die Mitarbeit ist es nicht notwendig, eine Pflanze korrekt bestimmen zu können. Falls Sie die Art jedoch kennen, erleichtern Sie uns damit die Zuordnung und Aufbereitung der Daten. Auch haben wir bereits sehr gute Erfahrungen mit privaten Bildsammlungen gemacht, die uns von interessierten Helfern zur Verfügung gestellt wurden. Falls auch Sie Ihre Pflanzenbilder bereitstellen möchten, senden wir Ihnen gerne einen Datenträger zu oder holen diese persönlich bei Ihnen ab.

Werden Sie ein Teil unseres Teams und leisten Sie einen wesentlichen Beitrag

zur Entwicklung einer Bestimmungshilfe für die Blütenpflanzen Thüringens! Weitere Informationen zu unserem Projekt und zur App finden Sie auf unserer Webseite www.floraincognita.de/.



WÄLDCHEN, J., A. THUILLE, M. SEELAND, M. RZANNY, E.-D. SCHULZE, D. BOHO, N. ALAQRAA, M. HOFMANN & P. MÄDER (2016): Flora Incognita – Halbautomatische Bestimmung der Pflanzenarten Thüringens mit dem Smartphone. – Landschaftspflege und Naturschutz in Thüringen 53 (3): 121–125

Projektlaufzeit: 2014–2019
 Förderkennzeichen: BMBF: 01LC1319A, 01LC1319B
 BfN / BMBU: 3514 685C19
 Stiftung Naturschutz Thüringen:
 SNT-082-248-03/2014
 Projektkoordination:
 Prof. (JP) Dr. PATRICK MÄDER

Dr. Marco Seeland
 Prof. Dr. (JP) Patrick Mäder
 M. Sc. David Boho
 M. Sc. Nedal Alaqraa
 M. Sc. Martin Hofmann
 TU-Ilmenau · Helmholtzplatz 5
 98693 Ilmenau
 patrick.maeder@tu-ilmenau.de

Dr. Jana Wäldchen
 Dr. Michael Rzanny
 Max-Planck-Institut für Biogeochemie
 Hans-Knöll-Str. 10 · 07745 Jena
 jwald@bgc-jena.mpg.de