



Bundesministerium
für Ernährung
und Landwirtschaft



Schlussbericht zum Thema

„Entwicklung automatisierter Verfahren zur frühzeitigen Felddiagnostik des Falschen Mehltaus, *Plasmopara viticola*, für einen angepassten Pflanzenschutz im Ökoweinbau“

FKZ: 2815NA187

Projektnehmer/Projektnehmerin:

Technische Universität Darmstadt

Gefördert durch das Bundesministerium für Ernährung und Landwirtschaft auf Grund eines Beschlusses des deutschen Bundestages im Rahmen des Bundesprogramms Ökologischer Landbau.

Das Bundesprogramm Ökologischer Landbau (BÖL) hat sich zum Ziel gesetzt, die Rahmenbedingungen für die ökologische Landwirtschaft in Deutschland zu verbessern. Es wird vom Bundesministerium für Ernährung und Landwirtschaft (BMEL) finanziert und in der BÖL-Geschäftsstelle in der Bundesanstalt für Landwirtschaft und Ernährung (BLE) in die Praxis umgesetzt. Das Programm gliedert sich in zwei ineinandergreifende Aktionsfelder - das Forschungs- und das Informationsmanagement.

Detaillierte Informationen und aktuelle Entwicklungen finden Sie unter:

www.bundesprogramm.de
www.oekolandbau.de/forschung

Wenn Sie weitere Fragen haben, wenden Sie sich bitte an:

Bundesanstalt für Landwirtschaft und Ernährung
Bundesprogramm Ökologischer Landbau
Deichmanns Aue 29
53179 Bonn
Tel.: 0228-6845-3280
E-Mail: boel-forschung@ble.de

Schlussbericht AuDiSens

Zuwendungsempfänger: Technische Universität Darmstadt, Fachbereich für Informatik, Fachgruppe Maschinelles Lernen

Förderkennzeichen: 2815NA187

Vorhabenbezeichnung: AuDiSens – Entwicklung automatisierter Verfahren zur frühzeitigen Felddiagnostik des Falschen Mehltaus, *Plasmopara viticola*, für einen angepassten Pflanzenschutz im Ökoweinbau

Laufzeit des Vorhabens: 09.11.2018- 31.10.2021

Gefördert durch:



Bundesministerium
für Ernährung
und Landwirtschaft

BÖLN

Bundesprogramm Ökologischer Landbau
und andere Formen nachhaltiger
Landwirtschaft

aufgrund eines Beschlusses
des Deutschen Bundestages

Kurzfassung Deutsch: (max. 1500 Zeichen)

Innerhalb des Projektes AuDiSens wurde die Implementierung und Erweiterung verschiedener Algorithmen des maschinellen Lernens, speziell des tiefen Lernens, zur Analyse und Krankheitsfrüherkennung hyperspektraler Bildaufnahmen aus dem Weinanbau erreicht. Hyperspektrale Aufnahmen stellen besonders hochdimensionale Daten dar, die deutlich schwieriger sind für das menschliche Auge zu analysieren. Mit den angewandten und erweiterten Methoden können nun hyperspektrale Aufnahmen des Weinbaus in automatisierte Verfahren untersucht, gleichzeitig aber auch leichter interpretiert werden.

Hierbei wurden innerhalb des Projektes verschiedene experimentelle Konfigurationen mit verschiedenen Methoden untersucht. Diese beinhalteten Bildklassifikation mit convolutional neural networks, Pixelreflektanzklassifikation mit self-attention networks, Wahrscheinlichkeitsmodelle mit sum-product networks, simplex volume maximization für unüberwachtes Lernen und verschiedene algorithmische Erweiterungen des eXplanatory Interactive Learning (XIL), die erlauben, dass eine NutzerIn Feedback auf die Erklärungen eines Modells geben kann, um es möglicherweise zu revidieren und somit sicher zu gehen, dass das Model Entscheidungen auf Basis "richtiger Gründe" macht. Dieser letzter Punkt ist besonders wichtig in Anwendungsbereichen, die mit hohem Risiko verbunden sind, wie z.B. der Weinanbau. Wie in unseren Experimenten gezeigt, kann es leicht sein für ein KI Model, ohne Vorkenntnisse oder Bedingungen in der Lösungsoptimierung, auf Störfaktoren in den Daten zu achten, um eine hohe Genauigkeit zu erzielen. Dieses Problem kann noch verstärkter in unkontrollierten Datenaufnahmebedingungen, wie z.B. auf dem freien Feld, auftreten. Daher ist gerade in diesem Umfeld eine besondere Vorsicht und die Verwendung von XIL besonders wichtig.

Auf Grundlage dieser Ergebnisse können zukünftig Pflanzenphänotypisierung im Weinanbau basierend auf hyperspektrale Messungen durchgeführt werden. Bisherige Ergebnisse wurden auf Daten aus kontrollierten Laborbedingungen erzielt. Für Hochdurchsatzverfahren im freien Feld werden zusätzliche Vorverarbeitungssoftware benötigt werden, wie zum Beispiel eine initiale Pflanzensegmentierung. Mit dem Abschluss dieser Arbeit wurde gezeigt, wie Algorithmen aus dem Bereich des maschinellen Lernens behilflich sein können in einer frühzeitigen Krankheitserkennung, sodass langfristig der Einsatz von der Umwelt schädlichen Pestiziden reduziert werden kann.

Kurzfassung Englisch: (max. 1500 Zeichen)

Within the AuDiSens project, the implementation and extension of various machine learning algorithms, especially deep learning, for the analysis and early disease detection of hyperspectral image recordings from vineyards was achieved. Hyperspectral images represent particularly high-dimensional data that are significantly more difficult for the human eye to analyze. With the applied and extended methods, hyperspectral images of viticulture can now be investigated in automated procedures, but at the same time they can be interpreted more easily.

Various experimental configurations were investigated within the project using different methods. These included image classification with convolutional neural networks, pixel reflectance classification with self-attention networks, probability models with sum-product networks, simplex volume maximization for unsupervised learning, and various algorithmic extensions of eXplanatory Interactive Learning (XIL) that allow a user to provide feedback on a model's explanations to possibly revise it and thus ensure that the model makes decisions based on "correct reasons". This last point is particularly important in high-risk application domains, such as viticulture. As shown in our experiments, it can be easy for an AI model, with no prior knowledge or constraints during its optimisation process, to look for confounding factors in the data in order to achieve high accuracy. This problem can be even more pronounced in uncontrolled data acquisition conditions, such as in the open field. Therefore, special care and the use of XIL is particularly important in this environment.

Based on these results, future plant phenotyping in viticulture can be performed based on hyperspectral measurements. The results of this project were obtained on data from controlled laboratory conditions. For high-throughput procedures in the field, additional pre-processing software will be needed, such as initial plant segmentation. With the completion of this work, it has been shown how machine learning algorithms can assist in early disease detection so that in the long run the use of environmentally harmful pesticides can be reduced.

I. <u>Abkürzungsverzeichnis</u>	5
II. <u>Einführung</u>	6
1. Gegenstand des Vorhabens, Ziele und Aufgabenstellung des Projekts, Bezug zu den Zielen des BÖLN	6
2. Planung und Ablauf des Projektes	7
3. Zusammenarbeit mit anderen Stellen.	9
III. <u>Wissenschaftlicher und technischer Stand</u>	9
IV. <u>Ausführliche Darstellung der wichtigsten Ergebnisse</u>	11
A. Beerenklassifikation:	11
B. Probabilistische Verfahren zur Anomaliedetektion:	12
C. Erklärungen und Interaktivität:	12
D. Aufmerksamkeit auf spektraler Ebene (Beeren):	14
E. FieldSpec- und UV-Vorversuchsdaten:	15
F. Prototype Trees:	17
G. Unabhängigkeitsproblem der Daten:	17
H. Simplex Volume Maximization zur Untersuchung der Datenverteilung:	20
I. Weiterentwicklung der Erklärungs- und Interaktivitätsalgorithmen:	20
J. Weiterentwicklung Neuro-symbolischer Modelle:	22
K. qPCR Datensatz:	23
V. <u>Diskussion und Zusammenfassung der Ergebnisse</u>	25
VI. <u>Nutzen und Verwertbarkeit der Ergebnisse</u>	26
VII. <u>Gegenüberstellung der ursprünglich geplanten zu den tatsächlich erzielten Zielen</u>	26
VIII. <u>Literaturverzeichnis</u>	27
IX. <u>Veröffentlichungen zum Projekt</u>	29

I. Abkürzungsverzeichnis

- NAP: Nationalen Aktionsplans zur nachhaltigen Anwendung von Pflanzenschutzmitteln
- CNN: Convolutional Neural Network
- XAI: eXplainable AI
- XIL: eXplanatory Interactive Learning
- SVM: Support Vector Machine
- SiVM: Simplex Volume Maximisation
- SPN: Sum-Product Network
- SAN: Self-Attention Network
- iCSN: interactive Concept Swapping Network
- NeSy: Neuro-Symbolic
- XGB: Gradient Tree Boosting
- ctrl: Kontrolle
- inoc: inokuliert
- norm: normalisiert

II. Einführung

1. Gegenstand des Vorhabens, Ziele und Aufgabenstellung des Projekts, Bezug zu den Zielen des BÖLN

Im Rahmen von AuDiSens wurden Sensor-basierte Frühdiagnoseverfahren am Beispiel des Falschen Mehltaus der Weinrebe neue Handlungsoptionen zur Reduktion von Pflanzenschutzmitteln untersucht. Die erzielten Ergebnisse hiervon dienen einem situationsgerechten Pflanzenschutz und stellen traditionelle und pilzwiderstandsfähige Rebsorten in einen Vergleich.

Die Implementierung der Diagnostik erfolgte stufenweise vom Labor über das Gewächshaus ins Freiland. In einer Machbarkeitsstudie sollte die Übertragbarkeit des im Labor und Gewächshaus erstellten Prognosemodells auf Freilandbedingungen geprüft werden. Hierfür sollten mobil einsetzbare Diagnosesensoren (d.h. Sensorsystem auf einer automatisierten Phänotypisierungsplattform und Temperatur- und Feuchtigkeitssensoren im Weinberg) getestet und ein Konzept zur Verwertbarkeit und Weiterentwicklungen erstellt werden.

Das Projekt AuDiSens ordnet sich in das Fördermodul A der Bekanntmachung über die „Förderung von innovativen Vorhaben für einen nachhaltigen Pflanzenschutz“ ein. Sie stellt einen Beitrag im Kontext des Nationalen Aktionsplans zur nachhaltigen Anwendung von Pflanzenschutzmitteln (NAP) dar (siehe Maßnahme 6.1.1). Das Vorhaben AuDiSens soll dazu beitragen, den technischen Fortschritt zu beschleunigen und verfolgte dabei folgende Ziele der Förderbekanntmachung:

Entwicklung und Optimierung technischer Lösungen:

- Erarbeitung von innovativen Lösungen zur situationsbezogenen und/oder teilflächenspezifischen Anwendung von Pflanzenschutzmitteln (Precision Farming)
- Entwicklung schneller, sensitiver und spezifischer Diagnoseverfahren für wirtschaftlich bedeutsame Schadorganismen, u.a. auch Felddiagnostik
- Entwicklung und Weiterentwicklung von Monitoringverfahren und –Systemen (z.B. Automatisierung der Informationserhebung, Felddiagnostik), Prognosemodellen und Schadschwellenkonzepten zur Bekämpfung von landwirtschaftlich bedeutsamen Schadorganismen.

In Deutschland dient weniger als ein Prozent der landwirtschaftlichen Nutzfläche dem Anbau von Reben. Auf dieser geringen Fläche erfordert allein die Bekämpfung der bedeutendsten Schaderreger der Rebe, Falscher Mehltau (*Plasmopara viticola*) und Echter Mehltau (*Uncinula necator*), einen der höchsten Pflanzenschutzmittelaufwendungen im Vergleich zu anderen

Kulturen. AuDiSens zielt auf eine Reduktion dieser Aufwendungen und befasst sich im Schwerpunkt mit dem Falschen Mehltau.

2. Planung und Ablauf des Projektes

Die Planung des Projektes AuDiSens basierte auf einen mehrstufigen Arbeitspaket- und Meilensteinplan. Die Arbeitspakete und Meilensteine der TU Darmstadt richteten sich hauptsächlich an die Vorverarbeitung und Analyse der aufgenommenen Hyperspektralbilder mittels Verfahren des tiefen Lernens. Hierbei wurden verschiedene Methoden des Maschinellen und speziell des Tiefen Lernens untersucht und verwendet.

Das Vorhaben AuDiSens gliederte sich in sechs Arbeitspakete (AP):

AP1: Epidemiologie von *P. viticola* auf PIWIs (JKI-OW)

AP2: Sensorevaluierung (JKI-ZR)

AP3: Auswertung der Sensor-Rohdaten (JKI-ZR und TU DA)

AP4: Data mining (TU DA)

AP5: Machbarkeitsstudie für Felddiagnostik (alle Partner)

AP6: Wissenstransfer in die Praxis und Wissenschaft (JKI)

AuDiSens setzte hierbei drei Schwerpunkte: (1) **Epidemiologie (AP1)**, d.h. die Untersuchung der Epidemiologie der Schaderreger auf pilzwiderstandsfähigen Rebsorten als Grundlage für eine gezielte Pflanzenschutzmittelapplikation; (2) Systematische **Erfassung, Auswertung und Interpretation (AP2, AP3, AP4)** von Sensordaten und **Fusion** von Auswerterroutinen zur verbesserten Prognose; (3) **Felddiagnostik (AP5)**, d.h. durch eine Machbarkeitsstudie soll die Übertragbarkeit des im Labor und Gewächshaus erstellten Prognosemodells auf Freilandbedingungen geprüft werden und ein Konzept für die Verwertbarkeit und Weiterentwicklung erstellt werden. Der **Wissenstransfer (AP6)** dient dem direkten Austausch mit der (Weinbau)Praxis, z.B. durch Feldtage. Im folgenden wird detaillierter auf die Arbeitspakete und Meilensteine eingegangen, in denen die TU Darmstadt eingebunden war.

Die Versuche der Arbeitspakete **AP1** und **AP2** wurden zunächst im Labor bzw. auf Sensortestständen im Gewächshaus durchgeführt. Die ausgelesenen Sensordaten mussten vorerst durch Tools der TU Darmstadt vorverarbeitet und referenziert werden. **AP2** wurde hierbei in starker Rückkoppelung der Arbeitspakete **AP3** und **AP4** bearbeitet, um das gesamte Potential der Sensordiagnostik zu erschließen. In einem iterativen Prozess wurden die Algorithmen verbessert und aus den Rohdaten weitere Informationen extrahiert.

AP2 beinhaltete unter anderem den **Meilenstein M-A1 “Proof-of-Principle zur prä-symptomalen Detektion unter kontrollierten Bedingungen”**. Hierbei ging es darum aufgenommene Sensordaten gemeinsam zu untersuchen bzgl. der Verwendbarkeit für weitere Analysen. Zusätzlich ging es darum ein Proof-of-principle zu erarbeiten um den Einsatz der

Sensoren, die Präparierung des Pflanzenmaterials etc. zu erörtern und gegebenenfalls in Absprache aller beteiligten Projektpartner zu überarbeiten. Dieser Meilenstein diente zusammenfassend dazu, erste Analysen auf Daten aus kontrollierten Laboraufnahmen durchzuführen.

AP3: Auswertung der Sensor-Rohdaten

In diesem Arbeitspaket war es notwendig automatisierbare Auswertetools für bildhafte Sensordaten zu entwickeln. Diese dienten dazu, unter anderem, die Rohdaten zu referenzieren und relevante Teile der Daten, z.B. durch cropping, in eine geeignete Repräsentation zu bringen zur anschließenden Bearbeitung mit Deep Learning / Machine Learning Methoden. Diese Auswertetools basierte auf bildverarbeitende Verfahren, wie z.B. Kantendetektoren. Die Auswertetools mussten jedoch für jeden Datensatz individuell angepasst werden, da viele der Filterparameter heuristisch erstellt wurden.

Innerhalb dieses Arbeitspaketes sollten außerdem in **Meilenstein M-A2 “Analyse und Interpretation von Sensordaten”** erste algorithmische Ansätze zur detaillierteren und umfangreichen Analyse und Interpretation der Sensordaten verwendet werden.

AP4: Data Mining

Im Zusammenhang von **Meilenstein M-TU1 “Erkennungs-Algorithmus”** sollten Erkennungs-Algorithmen, vorwiegend aus dem Feld des tiefen Lernens, für die Auswertung der Hyperspektraldaten untersucht und entwickelt werden. Zur eigentlichen Vorhersage des Krankheitsbefalls wurden die vertieften Merkmale zum Trainieren von z.B. Support-Vektor Maschinen (SVM) und anderen Techniken des Maschinellen Lernens benutzt. In einem ersten Schritt kamen Standardverfahren wie z.B. SVMs aber auch die sich für die Hochdurchsatz-Phänotypisierung bewährte Simplex- Volumenmaximierung (SiVM, Kersting et al. 2012) zum Einsatz.

Letztlich sollten in **Meilenstein M-TU2 “Effizienz- und Performancesteigerung”** die Algorithmen und Ansätze aus den vorherigen Arbeitspaketen und Meilensteine in Bezug auf Effizienz und höherer Leistungsfähigkeit untersucht werden. Höhere Genauigkeiten erzielten vor allem Algorithmen aus dem Bereich des Tiefen Lernens. Hier wurden unter anderem Sum-Product Networks (SPNs, Poon & Domingos 2011) und “convolutional neural networks” (CNNs) für die Analyse auf Bildebene, als auch Attention-basierte Algorithmen auf der Ebene einzelner Reflektanzen erfolgreich angewandt.

Gerade in Bereichen des hohen Risikos, wie zum Beispiel in der Krankheitserkennung im Weinanbau, ist es wichtig zu gewährleisten, dass ein Algorithmus sich tatsächlich auf Basis biologisch relevanter Gründe entscheidet und nicht z.B. auf Basis irrelevanter Korrelationen innerhalb der Daten. Aus diesem Grund, wurde in **M-TU2** außerdem auf die Erweiterung und

Entwicklung neuer Ansätze im Bereich der erklärbaren und interaktiven Algorithmen ("eXplanatory Interactive Machine Learning" (XIL)) fokussiert.

3. Zusammenarbeit mit anderen Stellen.

Die Forschungsarbeiten im Bereich XIL wurden an der Technischen Universität Darmstadt in enger Zusammenarbeit mit dem ebenfalls an der TU Darmstadt angesiedelten BMEL/BLE Projekts „DePhenSe“ (FKZ 2818204615) erarbeitet.

III. Wissenschaftlicher und technischer Stand

Hochgenaue Screening-Techniken (HTS) werden auf Einzelpflanzenskala zur Phänotypisierung eingesetzt. Messungen mit Hyperspektral- oder Fluoreszenzsensoren (Mahlein et al., 2012) erlauben die Beobachtung von Änderungen im spektralen Verhalten der Pflanze. Die Ausbildung von Krankheitsbildern lässt sich zeitlich und spektral hoch aufgelöst beobachten. Zur Generalisierung dieser Erkenntnisse sind Experimente mit vielen Wiederholungen und einer großen Zahl von Stichproben nötig. Bei Versuchen in Gewächshäusern oder Wachstumsräumen erfolgt die Vermessung mit hohem Durchsatz, indem die Pflanzen auf Förderbändern durch Messkammern gefahren werden, in denen sie mit geeigneten Sensoren vermessen, z.B. fotografiert, gescannt und gewogen, werden (Fahlgren et al., 2015, Honsdorf et al., 2014).

Diese Hochdurchsatzmessungen erzeugen große Datenmengen und benötigen neben intelligenten Datenspeicher-Organisationssystemen, automatische Auswertemethoden, die aus den aufgenommenen Daten die wichtigen Informationen extrahieren und diese für biologisch/agronomische Modelle zur Verfügung stellen. Klassisch werden drei Hauptprobleme unterschieden, die Klassifikation, die Regression und das Clustering, je nachdem ob der Vorhersagewert (nominell oder reell wertig) zur Lernzeit bekannt ist oder nicht (Hastie et al., 2001). Zur krankheitsspezifischen Diagnostik werden typischerweise zunächst Merkmalsvektoren mittels Bildverarbeitungsverfahren aus den Bildern extrahiert (Bauckhage et al., 2012, Herzog et al., 2014, Wahabzda et al., 2015b). Diese können dann mittels z.B. Neuronale Netze, Generalisierte Linear Modelle, Random Forrests, Support-Vektor Maschinen oder anderen Techniken des Maschinellen Lernens zur Klassifikation von Symptomen benutzt werden. Probabilistische Modelle und Gauss'sche Prozesse erlauben es, die Unsicherheit in der Vorhersage selbst zu quantifizieren, und haben sich zur präsymptomatischen Detektion und Prädiktion von (a)biotischem Stress auch bei Hochdurchsatzmessungen bewährt (Römer et al., 2012, Kersting et al., 2012, Wahabzda et al., 2015a). Auf diese Weise können Informationen gewonnen werden, welche für eine Felddiagnostik an der Rebe einsetzbar sind.

Es ist wichtig anzumerken, dass jedoch tiefe Netzwerke häufig auch als "black box" Modelle bezeichnet werden. Sie können erlauben es hohe Genauigkeiten zu erzielen, aber lassen es nur schwer zu eine Erklärung zu einer Entscheidung zu bekommen. Dies kann zu Problemen führen,

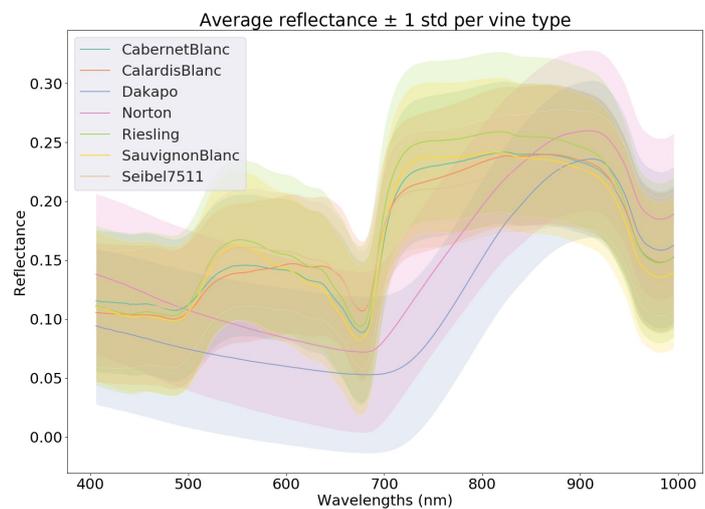
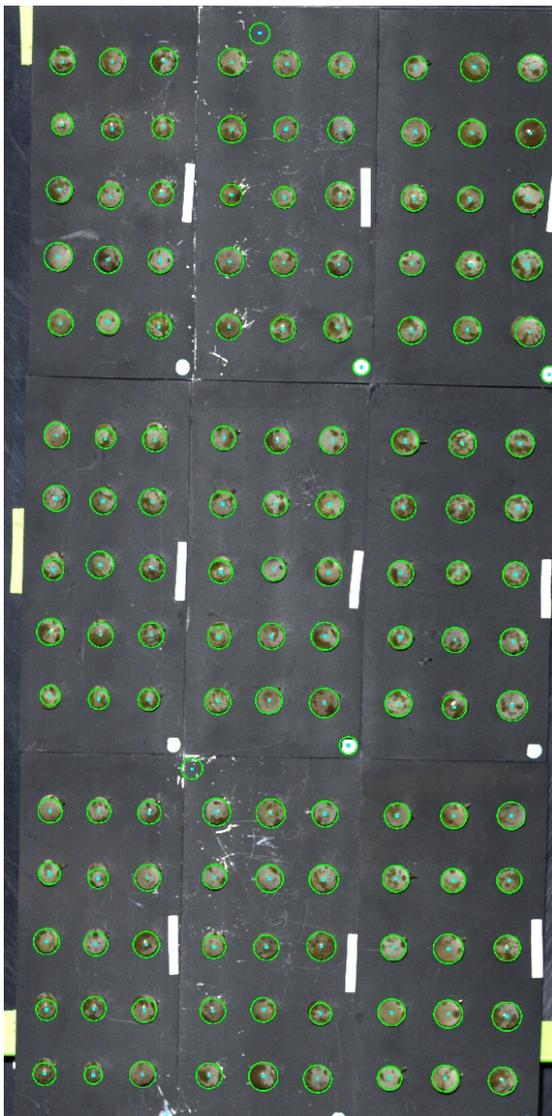


Abbildung 1: (links) Erkennung und Cropping der Beeren aus der Gesamtansicht, hier visualisiert für eine Platte mit CalardisBlancbeeren im Pseudo-RGB. (rechts oben) Extrahierte einzelne Beere in Pseudo-RGB. Diese dient als einzelne Probe, für die Bildklassifikationsalgorithmen. (rechts unten) Die durchschnittliche Reflektanz verschiedener Rebsorten (durchgezogene Linie) mit der Varianz als schattierte Region, um den Mittelwert.

da einige Arbeiten gezeigt haben, dass solche Modelle sich auch auf Störfaktoren in Daten achten können, um auf die hohe Genauigkeit zu kommen. Anhand von Erklärungen eines Modells können solche Fälle identifiziert werden, z.B. mittels einer der vielen Methoden aus dem Bereich explainable AI (XAI). Allerdings ist es noch vorteilhafter, wenn eine NutzerIn Feedback auf die Erklärungen eines Modells geben kann, um es möglicherweise zu korrigieren. Diese Art des Lernens wird explanatory Interactive Learning genannt (XIL) (Teso et al., 2019, Ross et al., 2017) und weist Vorteile auf, besonders für den Kontext wissenschaftlicher Daten.

	RGB	HS	RGB (norm)	HS (norm)
SVM	70,00	79,49	71,48	83,40
CNN	84,78	95,14	80,86	96,58

Tabelle 1: Zielgenauigkeiten verschiedener algorithmischer Ansätze, SVM und CNNs, auf pseudo-RGB Daten des Beerendatensatzes, als auch Hyperspektral Daten im VNIR Bereich. Weiterhin wurde der Effekt einer Weissreferenznormalisierung untersucht.

IV. Ausführliche Darstellung der wichtigsten Ergebnisse

Da im Projektverlauf verschiedene Datensätze aufgenommen wurden sind, entwickelte sich im Projektverlauf ein iterativer Prozess, in dem die verschiedenen Meilensteine und Arbeitspakete über die verschiedenen Datensätze bearbeitet und abgeschlossen wurden.

Im Zusammenhang mit AP1 und AP2 wurden verschiedene Datensätze in kontrollierten Laborumgebungen aufgenommen. Das Ziel hiervon war es optimale Sensorkonfigurationen und einen geeigneten experimentellen Aufbau innerhalb von AP3 und in Rückkopplung mit AP4 und AP5 zu untersuchen.

A. Beerenklassifikation:

In Absprache mit den Projektpartnern wurde hierbei vorerst mit einem Beerendatensatz gearbeitet. Dieser Datensatz besteht aus Hyperspektralaufnahmen von Beeren verschiedener Rebsorten, welche auf einer Ablage präpariert und ausgelegt wurden. Das Ziel dieser Datenaufnahme war es die Beeren in Bezug auf deren Rebsorte klassifizieren zu können. Auf die genauen Arbeitspunkte wird im weiteren Verlauf eingegangen.

Um die Rohsensordaten mittels Verfahren des maschinellen Lernens auswerten zu können, wurde vorerst eine Pipeline entwickelt. Diese beinhaltet die Weissreferenzlokalisierung und -normalisierung und die Gewebeerkennung und -lokalisierung. Abbildung 1 verdeutlicht die Ergebnisse dieses Verfahrens, welche auf Kantendetektion beruht.

Zum Klassifizieren der Daten an Hand der zugrundeliegenden Rebsorten, wurde das Standardverfahren „Support Vector Machine“ (SVM) auf die hyperspektralen, als auch pseudo-RGB Beerendaten angewandt. Diese Ergebnisse wurden mit einem Convolutional neural network (CNN) verglichen (Tabelle 1). Dieses zählt zu den Verfahren des tiefen Lernens, gilt aber mittlerweile auch als Standardverfahren der Bildverarbeitung als auch Pflanzenphänotypisierung (Gutiérrez et al., 2018, Nagasubramanian et al., 2019). Trotz der hohen Genauigkeit des SVMs liegt die des CNNs deutlich höher und berechtigt das weitere Interesse an Verfahren des tiefen Lernens. Tabelle 1 verdeutlicht auch noch mal die Bedeutung einer richtigen Normalisierung mit

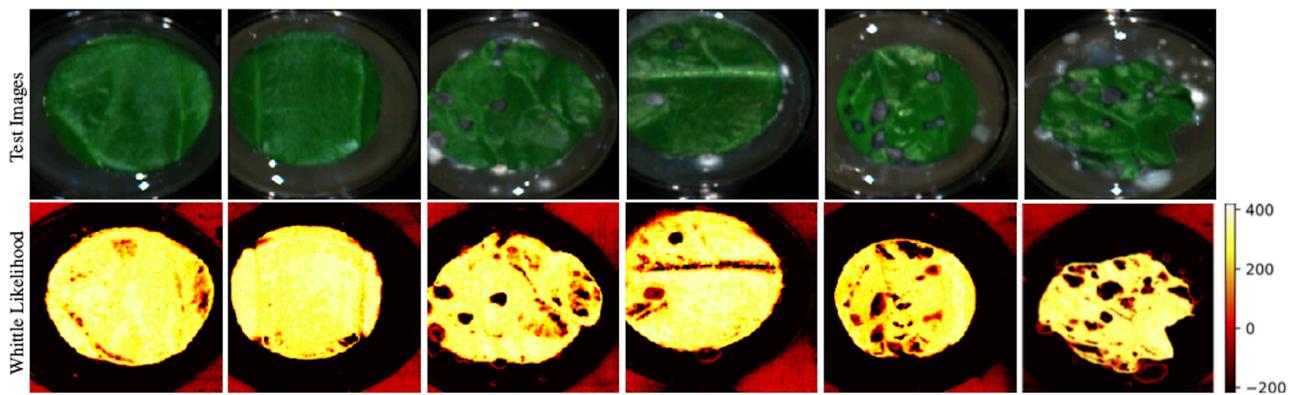


Abbildung 2: Anomaliedetektion an Hand von SPNs. Diese sind Wahrscheinlichkeitsmodelle, welche Sicherheiten angeben können wie wahrscheinlich ein Pixel zum bisher gesehenen Trainingsdatensatz gehört. In diesem Fall wurde ein Whittle-SPN auf Pixel der Klasse gesundes Gewebe trainiert, um an Hand der Wahrscheinlichkeit ungesundes Gewebe auf Pixelebene erkennen zu können.

der Weissreferenz. In der Tabelle erkennt man gerade im Hyperspektralbereich höhere Genauigkeiten, wenn die Daten adäquat normalisiert wurden.

B. Probabilistische Verfahren zur Anomaliedetektion:

In Zusammenarbeit innerhalb der Gruppe in Darmstadt wurde des Weiteren Anomaliedetektion mit Hilfe Wahrscheinlichkeitsmodelle, Sum-Product Networks (SPN), untersucht. Das Ziel war es mit Hilfe der angegebenen Unsicherheiten des Modells gesundes Gewebe von ungesundem Gewebe auf Pixelebene zu unterscheiden. Abbildung 2 verdeutlicht diese Ergebnisse. Für diesen Ansatz wurden diese Ergebnisse auf Hyperspektralaufnahmen des Schwesterprojekts, DePhenSe, visualisiert. Die Ergebnisse zeigen ein Proof of Principle des Ansatzes und die Stärken SPNs für diese Aufgabe.

C. Erklärungen und Interaktivität:

Tiefe neuronale Netze haben in vielen realen Anwendungen hervorragende Leistungen gezeigt. Leider kann es vorkommen, dass sie "Clever Hans"-ähnliches Verhalten zeigen --- unter Ausnutzung von Störfaktoren innerhalb von Datensätzen --- um eine hohe Leistung zu erzielen. In dieser Arbeit stellen wir die neuartige Lernsituation des „erklärenden interaktiven Lernens“ (XIL) vor und veranschaulichen seine Vorteile bei einer Forschungsaufgabe zur Phänotypisierung von Pflanzen. XIL fügt die Wissenschaftlerin in die Trainingsschleife ein, so dass sie das ursprüngliche Modell interaktiv überarbeitet, indem sie Feedback zu seinen Erklärungen gibt. Unsere experimentellen Ergebnisse zeigen, dass XIL dazu beitragen kann, Clever-Hans-Momente beim maschinellen Lernen zu vermeiden und das Vertrauen in das zugrunde liegende Modell zu fördern (oder ggf. zu entmutigen).

Die Ergebnisse dieser Arbeit wurden im Journal Nature Machine Intelligence veröffentlicht (Schramowski et al. (2020)). Eine Version kann unter <https://arxiv.org/abs/2001.05371> gelesen

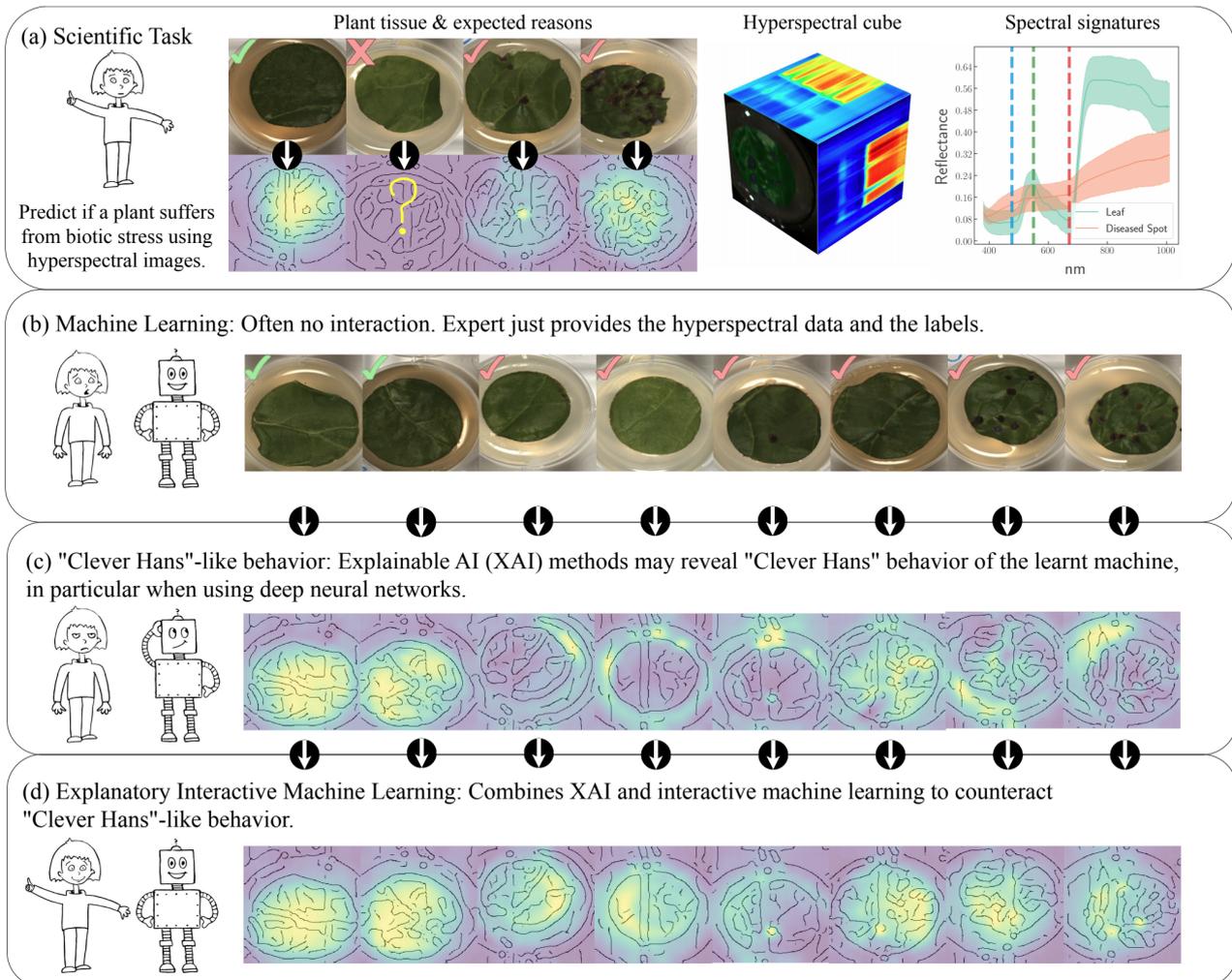


Abbildung 3: Explanatory Interactive Learning (XIL) — Benutzer revidieren lernende Maschinen in Richtung vertrauenswürdige Entscheidungsstrategien.

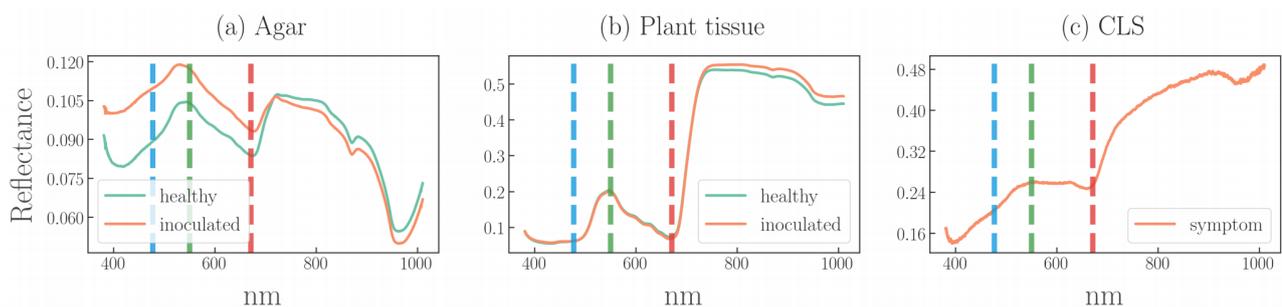


Abbildung 4: Spektrale Signaturen von gemessenen Agarplatten mit Zuckerrübenblattscheiben.

werden. Abbildung 3 veranschaulicht diesen Prozess. Genauer, wurde ein tiefes neuronales Netzwerk darauf trainiert Bilder mit ausgestanzten Blattgewebe zu klassifizieren, je nach dem, ob die Pflanzen inokuliert wurden mit dem Erreger *Cercospora Beticola*, oder nicht. Über die Erklärungen des Netzwerkes, konnte erkannt werden, dass das Netzwerk auf Merkmale des Agars (die einbettende Flüssigkeit in der Ablageschale) fokussierte für die hohen Genauigkeiten, anstatt auf Merkmale des Gewebes. Biologische Untersuchungen bestätigten, dass im

Hyperspektralraum tatsächlich Unterschiede vorhanden waren im Agar der Kontroll- und inokulierten Pflanzen (siehe Abbildung 4). Mit Feedback einer Expertin auf den Erklärungen konnte das Model, jedoch revidiert werden, um hohe Genauigkeiten basierend auf biologisch "richtige Gründe" zu haben.

D. Aufmerksamkeit auf spektraler Ebene (Beeren):

Die Klassifizierungsergebnisse aus Abschnitt A wurden an Hand der Methode aus C untersucht, wobei unter anderem rauskam, dass das CNN auch auf dem Stiel der Beeren achtet, um eine akkurate Vorhersage bzgl. der Rebsorte zu treffen. Anstatt auf der Bildebene die Erklärungen zu korrigieren, wie in C untersucht, wurde in Absprache mit den Projektpartnern direkt eine Klassifikation auf Pixelebene durchgeführt. Speziell wurde ein „Self-Attention Network“ (SAN) (Skrli et al., 2020) benutzt, welches die Eingabedimensionen mit selbst erlernten Gewichten gewichtet. Somit können für bestimmte Klassen bestimmte Merkmale, in diesem Fall bestimmte Wellenlängenbereiche, wichtiger sein, als für andere Klassen. Diese Methodik wurde innerhalb

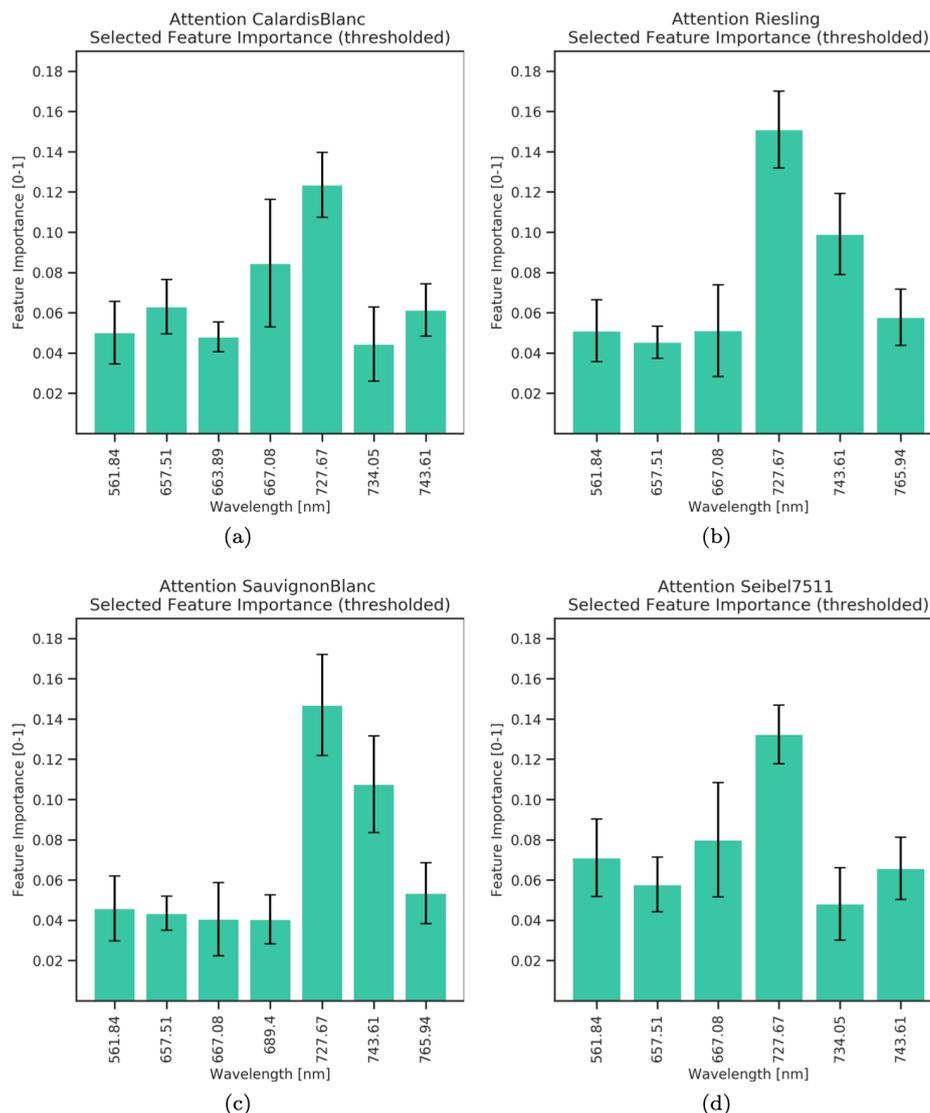


Abbildung 5: Die Gewichte verschiedener Klassen, welche vom SAN erlernt wurden. Seite 14 von 29



Abbildung 6: Aufnahme des neu erstellten UV-Vorversuch Datensatzes. Die hyperspektrale Aufnahme wird hier zur Veranschaulichung als RGB Bild dargestellt (links). Zu sehen ist ein Weinsteckling wobei ein Blatt mit UV-C Licht bestrahlt wurde, zur Untersuchung der Induktion der Stilbensynthese. Das bestrahlte Blatt wurde von den Projektpartnern annotiert und liegt als binäre Masken vor (rechts).

des Projektzeitraums schon erfolgreich im Schwesternprojekt, DePhenSe, angewandt (Brugger et al., 2021).

Die Vorhersagegenauigkeit lag bei 93,19%. Da die Netzwerke nur auf Pixel der Beeren trainiert wurden, konnte somit sicher gestellt werden, dass die Netzwerke nur auf Merkmale der Beeren geachtet haben. In Abbildung 5 sind die Erklärungen der SANS dargestellt für vier Rebsorten. Diese stellen dar welche Wellenlängen wichtig sind für die jeweilige Rebsorte.

E. FieldSpec- und UV-Vorversuchsdaten:

Im Projektzeitraum 2020 wurden zwei weitere Datensätze untersucht. Beide Datensätze sind Vorversuchsdatensätze, die für Untersuchungen vor der Aufnahme der Hauptversuche dienen sollen. Beide dieser Datensätze bestehen aus Hyperspektralaufnahmen ganzer Stecklinge (siehe Abbildung 6 links).

Der FieldSpec Vorversuch stellt eine Zeitreihenaufnahme dar, bei der Bilder von Stecklinge über einen Zeitraum von 7 dpi (days post inoculation) aufgenommen wurden. Hierbei wurden Aufnahmen von 20 Stecklinge erstellt, wobei 14 Stecklinge an bis zu zwei markierte Blätter mit Sporen des Falschen Mehltaus inokuliert wurden und 6 Stecklinge zur Kontrolle dienten. Hierbei entstammt eine Hälfte der Stecklinge der Sorte Solaris und die andere der Sorte Müller-Thurgau.

Der zweite Datensatz, der UV-Vorversuchsdatensatz, stellt keine Zeitreihe dar und beinhaltet nur Stecklinge der Sorte Solaris. In diesem Datensatz wurden bis zu zwei Blätter mit UV-C Licht

	Train	Test
SWIR	50 ± 0	50 ± 0
VNIR	95.08 ± 0.21	95.88 ± 0.5

bestrahlt, um die Induktion der Stilbensynthese zu untersuchen. In diesem Datensatz wurden 5

Tabelle 2: Die Klassifikationsgenauigkeit der SAN für die Pixelklassifikation des UV-Vorversuchs Datensatzes. Hierbei werden die Genauigkeit auf dem Trainings- und Testsatzes dargestellt, sowohl für die SWIR Daten, als auch VNIR Daten.

Stecklinge als Kontrollpflanzen aufgenommen und 10 Stecklinge mit UV-C Licht bestrahlt.

Zu jeder Hyperspektralaufnahme wurde von den Projektpartnern eine Annotation der relevanten Blätter (inokuliert oder bestrahlt) auf Pixelebene erstellt (siehe Abbildung 6 rechts).

Beide Datensätze beinhalten nur wenige Bilder. Daher wurden die Self-Attention Networks auch auf zur Pixelklassifikation dieser Datensätze angewandt. Hierbei wurden die relevanten Pixel mit Hilfe der erstellten Annotationsmasken extrahiert und eine Klassifikation pro Pixel für die Labels Kontrolle und inokuliert/bestrahlt durchgeführt, basierend darauf, ob ein Pixel von einer bestrahlten/ inokulierten Pflanze entstammt oder einer Kontrollpflanze.

Tabelle 2 stellt die Genauigkeitsergebnisse als Mittelwert und Standardabweichung über eine 5-fache Kreuzvalidierung dar, sowohl für den VNIR und SWIR Bereich, für den UV-Vorversuchs

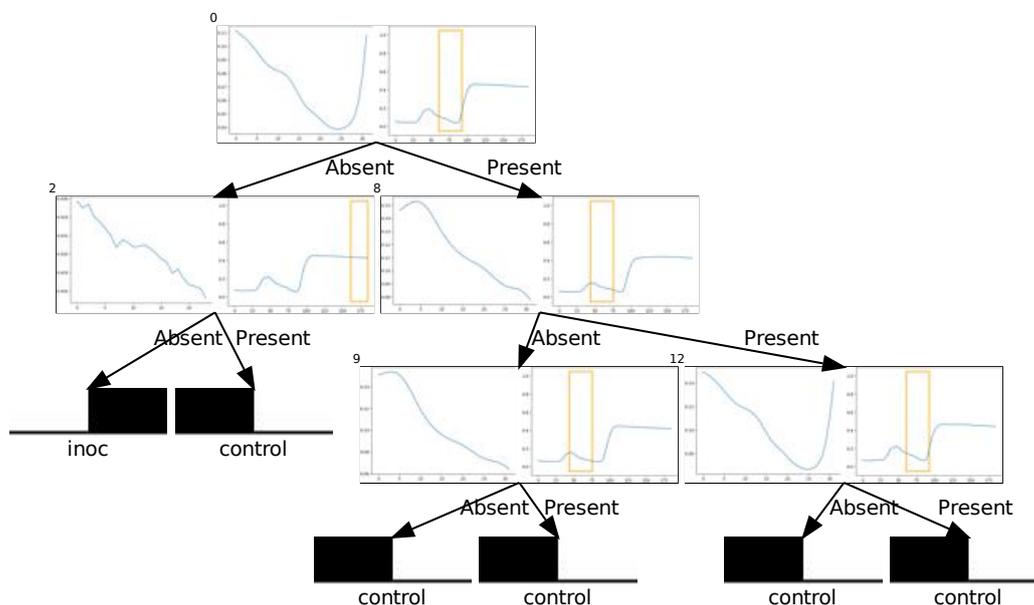


Abbildung 7: Der gelernte Entscheidungsbaum für den VNIR UV-Vorversuchsdatensatz. Dieser kann wie folgt gelesen werden: an jedem Knoten wird entschieden, ob der Part Prototype der jeweils links dargestellt wird und aus dem gelben Bereich eines Trainingbeispiels (im rechten jeweiligen Bild dargestellt), in einem Datenbeispiel enthalten ist. Somit wird ein Pfad im Entscheidungsbaum gefolgt bis zu einem Blatt, aus dem dann die Klassifikationentscheidung ausgelesen werden kann.

	Train	Test
VNIR	83,89 ± 6,29	83,94 ± 6,34

Tabelle 3: Die Klassifizierungsgenauigkeit des ProtoTree auf die VNIR Daten des UV-Vorversuchsdatensatzes.

Datensatz. Pro Kreuzvalidierungslauf wurde der gesamte Datensatz zufällig in Trainings, Validierungs- und Testsatz aufgeteilt.

Man erkennt, dass im VNIR Bereich eine hohe Genauigkeit erzielt werden konnte. Im SWIR Bereich konnte keine Verbesserung über eine Zufallsrate erzielt werden.

F. Prototype Trees:

Zur weiteren Untersuchung der Erklärungen für eine Klassifikation wurde ein weiterer Algorithmus untersucht in Bezug auf den UV-Vorversuchsdatensatz. Der ProtoTree (Nauta et al., 2020) Ansatz basiert darauf ein CNN gefolgt von einem binären Entscheidungsbaum zu trainieren. Der Entscheidungsbaum beinhaltet eine Sammlung an lernbare Prototypensoren. An jedem Knoten des Entscheidungsbaumes wird nun bestimmt, z.B. durch eine Distanzmetrik, ob ein Datenbeispiel nah an dem Prototypen des Knotens ist. Die Prototypen in ProtoTree haben nur die Dimensionen eines Ausschnitts des Datenbeispiels, z.B. ein kleiner Ausschnitt eines Bildes. Somit stellen die Prototypen nur "Prototype Parts" dar.

Ein Prototyp kann nun dargestellt werden als ein Ausschnitt aus dem Trainingsdatensatz, der zu dem Prototypen am nächsten ist. Der Entscheidungsbaum wird nun so gelesen, dass an jedem Knoten entschieden wird, ob der Prototyp im Datenbeispiel vorhanden ist oder nicht, bis zu einem Baumblatt, der die finale Klassifikation vorhersage liefert.

Für die Pixelklassifikation des UV-Vorversuchsdatensatzes wurde ein ProtoTree mit einem 1D CNN, wie oben erwähnt, trainiert. Die Genauigkeitsergebnisse und Erklärungsvisualisierungen werden in Tabelle 3 und Abbildung 7 dargestellt. Diese Art der Erklärungen stellt somit einen interessanten Ansatz dar, der auf größere Regionen und daher potenziell verständliche, konzeptuellere Erklärungen basiert, im Vergleich zu dem Ansatz der SAN Erklärungen, welches eine Relevanz pro Wellenlänge liefert. Allerdings, ist die Vorhersagegenauigkeit mit dem ProtoTree Ansatz weniger akkurat, im Vergleich zum SAN Ansatz.

G. Unabhängigkeitsproblem der Daten:

Im Laufe der Analysen zum ProtoTree wurde sicherheitshalber untersucht in wie weit die bisherige Annahme stimmt, dass die Pixel tatsächlich i.i.d (unabhängig und identisch verteilt) sind. Dies ist eine wichtige Annahme im Deep Learning Bereich und kann zu irreführende Ergebnisse und overfitting führen, falls nicht korrekt.

Ähnlich zu den Ergebnissen mit dem Beerendatensatz, wurden die Pixel in den bisherigen Experimenten einem der drei Datenaufteilungen (Trainings-, Validierungs- oder Testsatzes) zugeteilt, unabhängig davon, von welchem Bild/welcher Pflanze der Pixel stammt. Dies hat zur

	Train	Test
Mixed Plant ID	95.51 ± 0.05	95.42 ± 0.22
Separate Plant IDs	97.11 ± 0.97	55.55 ± 7.15

Tabelle 4: Untersuchung des Effektes des Datenaufteilungsansatzes auf die Klassifikationsgenauigkeit für den SAN Algorithmus auf dem VNIR UV Vorversuchdatensatz. In „Mixed Plant IDs“ wurden Pixel zufällig einem der drei Datensätze (Trainings-, Validierungs-, Testsatz) zugewiesen. In „Separate Plant

	Train	Test
Mixed Berry ID	95.43 ± 0.06	95.30 ± 0.06
Separate Berry IDs	95.51 ± 0.73	93.19 ± 1.08

Tabelle 5: Untersuchung des Effektes des Datenaufteilungsansatzes auf die Klassifikationsgenauigkeit für den SAN Algorithmus auf dem Beerendatensatz. In „Mixed Berry IDs“ wurden Pixel zufällig einem der drei Datensätze (Trainings-, Validierungs-, Testsatz) zugewiesen. In „Separate Berry IDs“ wurde alle Pixel einer Beere einem der drei Datensätze zugewiesen.

Folge, dass Pixel eines Stecklings sowohl im Trainings-, als auch im Testsatz vorhanden sein könnten. Wegen der niedrigen Anzahl an Stecklingen in den neuen Daten wurde daher überprüft, ob die bisherigen Algorithmen fälschlicherweise Eigenschaften der Pflanzen nutzen, statt globalere Eigenschaften, wie z.B. die Folgen einer UV-C Bestrahlung, um die hohen Genauigkeiten zu erzielen.

Dafür wurden die Pixel nun so aufgeteilt, dass alle Pixel einer Pflanze exklusiv in einer der drei Datenaufteilungen sind. Die Algorithmen wurden erneut trainiert und durch eine Kreuzvalidierung untersucht. Für diese Analysen haben wir den vorherigen Datenaufteilungsansatz "Mixed Plant IDs" bezeichnet und den neuen Ansatz „Separate Plant IDs“. Der Vergleich für die VNIR UV Vorversuchsdaten wird in Tabelle 4 beispielhaft für SANs dargestellt.

Zu erkennen ist, dass tatsächlich die Testgenauigkeit deutlich sinkt mit dem Separate Plant ID Ansatz, im Vergleich zum Mixed Plant ID Ansatz, obwohl die Trainingsgenauigkeit sogar leicht ansteigt. Dies bestätigt, dass tatsächlich die Pixel einer Pflanze nicht unabhängig voneinander zu betrachten sind. Die Ergebnisse deuten darauf hin, dass die Modelle eher auf individuelle Eigenschaften der wenigen Stecklinge achten, als generelle Eigenschaften z.B. einer Erkrankung.

Als nächstes wurden die bereits abgeschlossenen SAN Experimente mit dem Beerendatensatz wiederholt, wobei auch hier beide Datenaufteilungsansätze verwendet wurden. Die Genauigkeitsergebnisse sind in Tabelle 5 dargestellt. Es ist zu erkennen, dass es bei dem Beerendatensatz keinen starken Effekt zu erkennen gibt. Dies lässt darauf schließen, dass durch die höhere Anzahl an Beeren pro Klasse (ca. 125 Beeren pro Klasse) genug unabhängige Datenbeispiele vorliegen, damit der Algorithmus besser generalisieren kann, im Vergleich zu den maximal 10 unabhängige Stecklinge des UV Vorversuchdatensatzes.

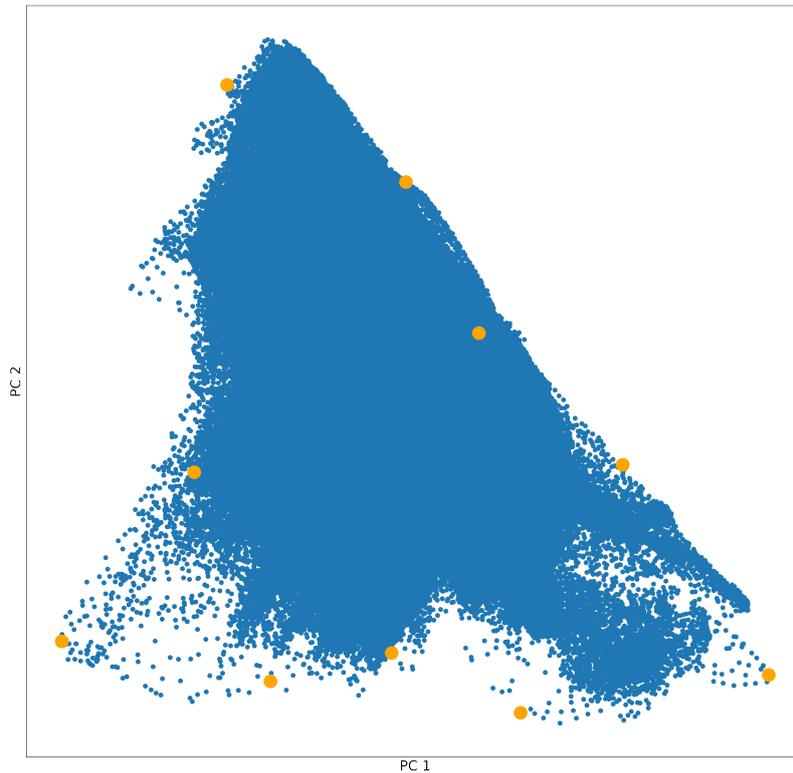


Abbildung 8: Die VNIR UV Vorversuchsdaten im Raum der zwei Hauptkomponenten einer PCA dargestellt. In Orange sind die Basisvektoren dargestellt, die durch den SiVM Algorithmus bestimmt wurden.

Ein weiterer wichtiger Unterschied zwischen den beiden Datensätzen ist, dass alle Pixel einer Beere eindeutig einer Weinsorte zugehören. Im Vergleich: im UV Vorversuchsdatensatz wird ein Blatt inokuliert, bzw. bestrahlt, aber es ist nicht eindeutig, dass jeder Bereich des Blattes relevante Symptome entwickelt, bzw. aufweist. Somit können auch auf einem inokulierten Blatt gesunde Bereiche vorkommen. Das macht eine eindeutige supervised Klassifikation schwierig auf Pixelebene schwierig, selbst mit einem Superpixel Ansatz wie in Brugger et al. (2021).

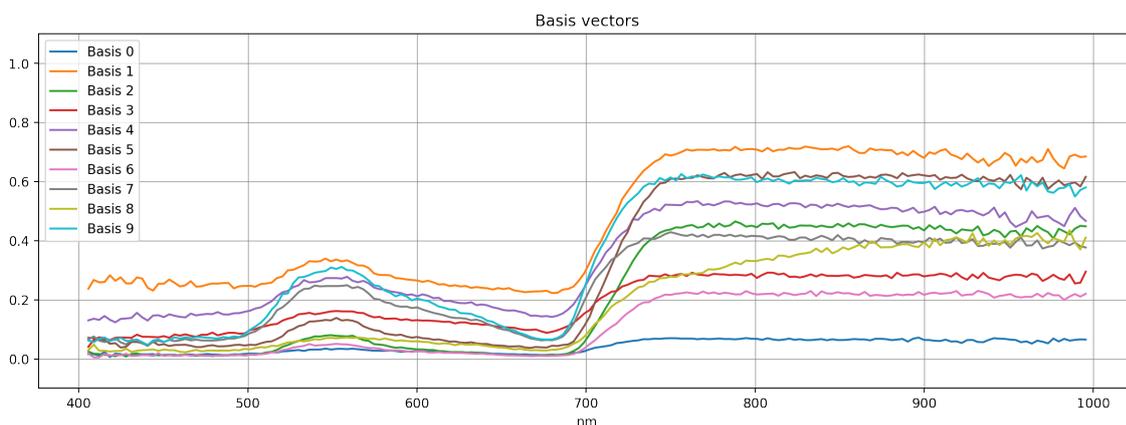


Abbildung 9: Die 10 Basisvektoren für den VNIR UV Vorversuchsdatensatz durch SiVM bestimmt wurden.

H. Simplex Volume Maximization zur Untersuchung der Datenverteilung:

Um die Verteilung der Daten besser zu verstehen in Bezug auf die Variabilität der Pixelreflektanzen wurde daher ein unsupervised Algorithmus angewandt, der Simplex Volume Maximization Algorithmus (SiVM). Dieser bestimmt eine vom Nutzer vorgegebene Anzahl an Datenbeispielen, die den hochdimensionalen Raum der Daten am besten umschließt. Diese „Extrema“ der Daten, auch Basisvektoren bezeichnet, können dann visualisiert werden.

Abbildung 8 zeigt alle Daten im Hauptkomponentenraum der zwei Hauptkomponenten einer PCA Analyse (Principal Component Analysis). In Orange sind die Basisvektoren markiert, die durch SiVM bestimmt wurden. Abbildung 9 stellt diese 10 Basisvektoren im Spektralraum dar.

In Zusammenarbeit mit den Projektpartnern wurden diese Basisvektoren (Abbildung 9) untersucht bzgl. der hohen Variabilität. Es wurde bestimmt, dass mehrere dieser Basisvektoren Pixel in verdunkelte Bereiche der Bilder entstammen, z.B. in denen das Blatt gekrümmt ist.

I. Weiterentwicklung der Erklärungs- und Interaktivitätsalgorithmen:

Im weiteren Verlauf wurde unabhängig von den Hyperspektraldaten weiter an einer Methodik gearbeitet, um die Qualität und Verwendbarkeit von Erklärungen im Deep Learning Bereich zu verbessern. Bisherige Ansätze konzentrieren sich stark darauf, dass dem Menschen eine Gewichtung für eine Vorhersage pro Eingabevariabel geliefert wird (Schramowski et al., 2020). Im Bildbereich besteht dies z.B. aus einer Matrix, die darstellt wie wichtig jeder Pixel war für eine Vorhersage.

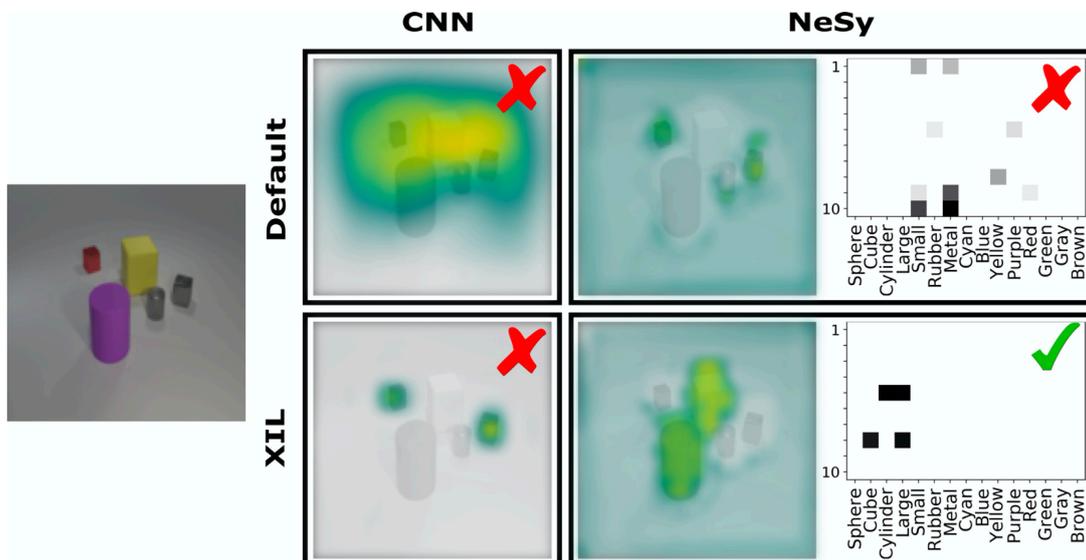


Abbildung 10: Mit unserem Neuro-Symbolic Concept Learner ist es möglich eine symbolische als auch visuelle Erklärung zu erhalten (oben rechts), welches darstellt welche Bildregion relevant ist für eine Entscheidung und aus welchen konzeptuellen Gründen. Mit unserem NeSy XIL Ansatz ist es nun auch möglich das Netzwerk effizient zu korrigieren (unten rechts). Dies ist nicht möglich mit einem Standard CNN Model (Mitte oben und Mitte unten).

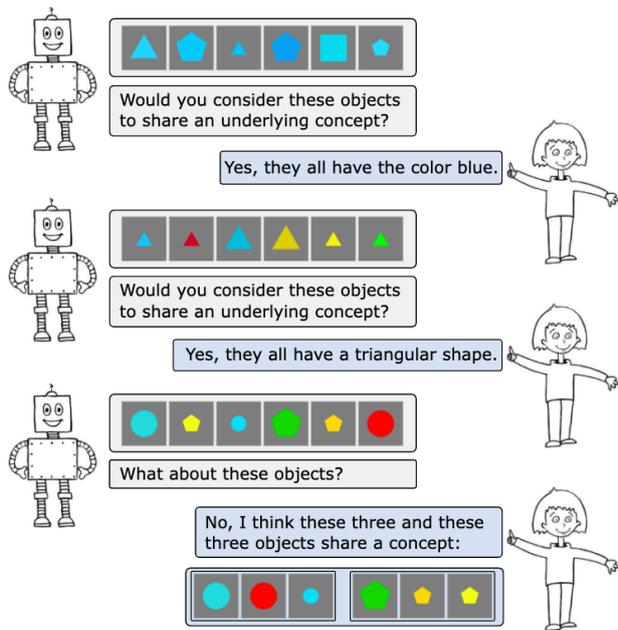


Abbildung 11: Ein trainiertes Modell (links) fragt den menschlichen Benutzer (rechts) ab ob die Konzepte, die es aus den Daten extrahiert hat, mit dem mit dem Wissen des Benutzers übereinstimmen. Anschließend kann das Modell Revisionen vom Benutzer erhalten.

Klassifikation auf Basis dieser symbolischen Repräsentation durchführt, anstatt für den menschen unverständliche gelernte Feature zu nutzen, wie dies bei einer standard CNN Klassifikation oft der Fall ist.

Mit Hilfe dieser verständlichen symbolischen Repräsentation, ist es nun möglich für einen Menschen Feedback zu einer Erklärung nicht nur auf der Bildebene zu geben (z.B. mit Hilfe von Explanatory Interactive Learning Ansätze, siehe Ross et al. (2017), Teso und Kersting (2019) und Schramowski et al. (2020)), sondern auch auf diese symbolische Ebene. Somit kann der Einfluss von Störfaktoren, die zu verfälschte Vorhersageergebnisse eines Deep Learning Models führen können, reduziert werden. Dies wäre mit den bisherigen Ansätzen im Deep Learning Bereich nicht möglich gewesen. Abbildung 10 zeigt ein beispielhaftes Ergebnis aus der Arbeit.

Diese Ergebnisse sind von hoher Relevanz für AuDiSens für eine verständliche automatische Entdeckung von Wissen bzw. Mustern. Für eine Anwendung einer Krankheitsfrüherkennung im wissenschaftlichen und wirtschaftlichen Kontext, ist es wichtig, nicht nur den relevanten Bildbereich zu erhalten den das Diagnosesystem zur Vorhersage nutzt, sondern auch welche konzeptuelle Informationen sich in diesem Bildbereich, vor allem im Spektralraum, befinden.

Eine Schwierigkeit bei solch einem Neuro-symbolischen Ansatz ist, dass ein gewisser Satz an annotierten Attributen notwendig ist, um den symbolischen Teil des Modells zu trainieren. Gerade im Bereich der Wissensentdeckung und Wissenschaftliche Daten, zu denen es häufig an geeignete Vorkenntnisse fehlt, ist es schwierig ein symbolisches Model, und in dem

Diese Art der Erklärung lässt noch viel Unklarheit offen, da es nicht klar wird, warum genau ein Pixel wichtig ist. Es wird also nicht ersichtlich welches Konzept aus diesem Pixel extrahiert wurde, um eine Entscheidung zu fallen. Wichtiger noch, ohne dem Wissen, welches Konzept genutzt wurde, ist es schwierig für die Endnutzerin ein Model zu korrigieren, das zwar eine relevante Bildregion für die Entscheidung verwendet, aber möglicherweise falsche Konzepte nutzt.

In der Arbeit Stammer et al., 2021, die bei der renommierten Konferenz Computer Vision and Pattern Recognition 2021 publiziert wurde, wird eine Lösung zu diesem Problem dargestellt (<https://arxiv.org/abs/2011.12854>). Es wird ein Neuro-Symbolic Concept Learner vorgestellt, welcher eine Szene aus Objekten zuerst in eine symbolische Repräsentation abbildet und eine

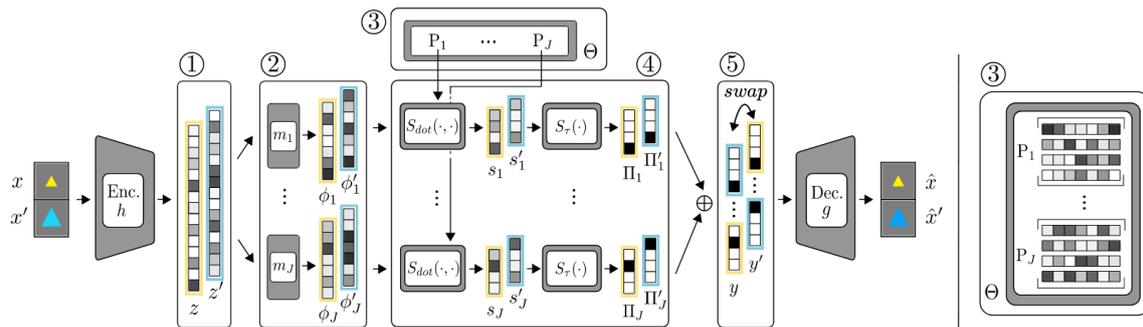


Abbildung 12: Interactive Concept Swapping Network. Ein iCSN basiert auf einer deterministischen Autoencoder-Struktur, mit einer anfangs verstrickte latente Kodierung (1). Mehrere Auslese-Encoder (2) extrahieren relevante Informationen aus diesem latenten Raum und vergleichen ihre extrahierten Konzeptkodierungen mit einem Satz von Prototyp-Slots (3) über ein gewichtetes, Softmax-basiertes Punktprodukt (4). Diese führt zu einem diskretisierten Code, der den ähnlichsten Prototyp-Slot jeder Konzeptkodierung angibt. iCSNs werden über eine einfache Rekonstruktionsverlustfunktion, schwache Überwachung durch Match Pairing und einen Swapping-Ansatz, der die latenten Konzeptrepräsentationen für gemeinsame Konzepte vertauscht (5) und die Bindung semantischer Informationen an spezifische Prototyprepräsentationen erzwingt.

Zusammenhang symbolische Erklärungen zu erhalten. In einer weiteren Arbeit der TU Darmstadt, welches bei der Konferenz Computer Vision and Pattern Recognition 2022 publiziert wird, wird eine Lösung zu diesem Problem dargestellt (Stammer et al. 2022, <https://arxiv.org/pdf/2112.02290.pdf>). Speziell, wird hier ein Verfahren vorgestellt mit dem über schwache Überwachung einem Model prototypische Repräsentationen bestimmter Konzepte gelernt werden können und über ein Diskretisierungsverfahren für den Menschen interpretierbar gemacht werden. Da der latente Raum, die das Model lernt, semantisch fundiert und diskret sind, kann ein Mensch diese leicht untersuchen und gegebenenfalls durch gezieltes Feedback revidieren. Abbildung 11 skizzenhaft mögliche Interaktionen zwischen dem Model und Menschen, während Abbildung 12 einen Überblick über die Module des Models darstellt. Mit diesem Verfahren ist es nun möglich einen semantisch bedeutungsvollen Satz an diskreten Konzepten zu lernen, welche nun für Neuro-symbolische Argumentation verwendet werden können.

J. Weiterentwicklung Neuro-symbolischer Modelle:

Wie in Abschnitt I dargestellt, haben Neuro-symbolische Ansätze viele Vorteile. Eine andere Art Neuro-symbolische Ansätze zu entwickeln, ist über sogenannte "deep probabilistic programming languages" (DPPL). In Skryagin et al. 2021 (<https://arxiv.org/abs/2110.03395>), welches zu Zeit noch in Begutachtung steht und in Zusammenarbeit an der TU Darmstadt entwickelt wurde, wurde eine neue DPPL entwickelt, SLASH, welches neuronale, probabilistische und logische Module verbindet, um logische Anfragen am Model beantworten zu können. Besonders hier dran, ist, dass

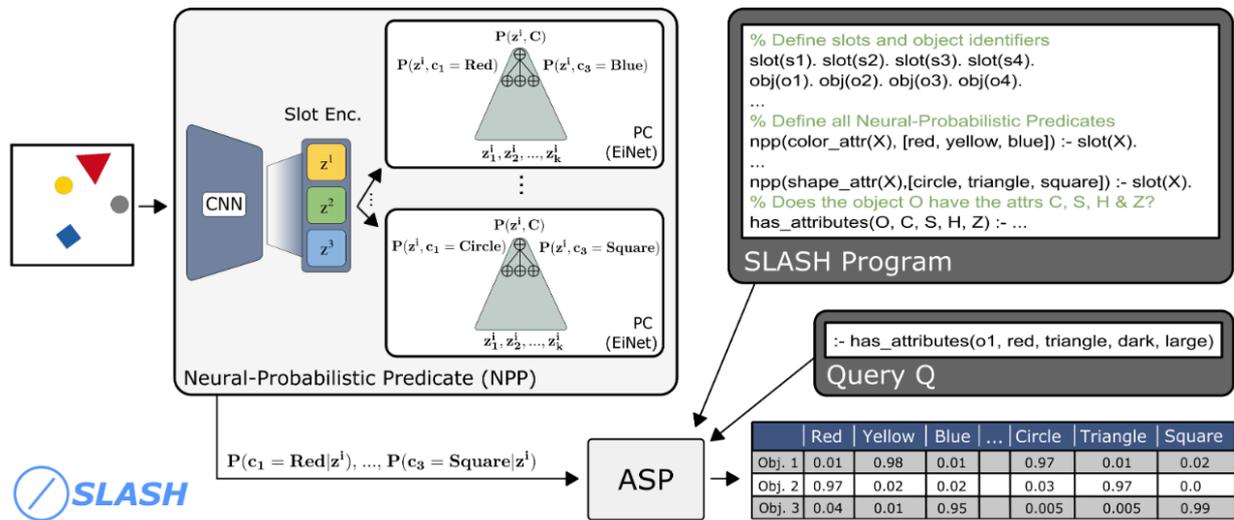


Abbildung 13: SLASH Attention illustriert für eine visuelle Schlussfolgerungsaufgabe. SLASH mit neuronal-probabilistischen Prädikaten, bestehend aus einem Slot-Attention-Encoder und probabilistischen Schaltkreisen (PCs), die über EiNets realisiert werden.

	Ctrl vs Inoc	Ctrl vs Res. vs Susc.	Wine type
XGB	73,33 ± 1,92	58,15 ± 1,47	42,62 ± 1,15
CNN	85,75 ± 1,24	76,59 ± 6,09	90,24 ± 1,07

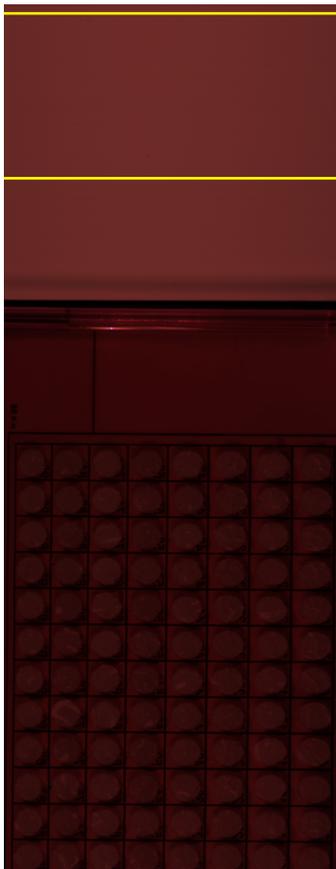
Tabelle 6: Die Klassifizierungsgenauigkeit der gradient boosting Methode (XGB) und eines CNNs auf die Bilder des qPCR Datensatzes. Hier wurden verschiedene Klassifikationen durchgeführt.

das Model nun über Unsicherheiten folgern kann und ein Mensch Gegenfalls vorhandene Vorkenntnisse über die Programmiersprache an das Model geben kann. Abbildung 13 stellt eine Instanziierung SLASHs dar, welche für Objekt-basiertes Konzeptlernen entwickelt wurde. Diese Arbeit bietet Möglichkeiten Vorkenntnisse im Weinbau an ein Model weiterzugeben.

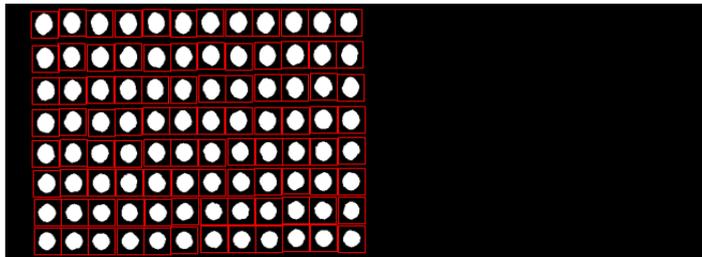
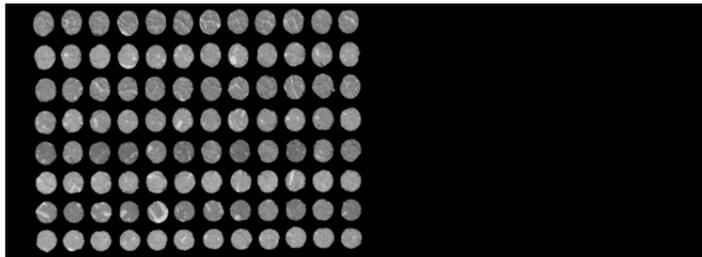
K. qPCR Datensatz:

Um die Probleme des experimentellen Aufbaus zu umgehen, die in vorherigen Abschnitten zu den FieldSpec- und UV-Vorversuchsdaten identifiziert wurden sind, wurde der experimentelle Aufbau etwas verändert. In einem neuen Datensatz (weiter, als qPCR Datensatz benannt), wurden ausgestanzte Blattscheiben auf Ablagepaletten gelegt und mit einer SWIR Kamera aufgenommen. Dieser Aufbau ähnelt dem der Beerendaten und ermöglicht es deutlich mehr Beispiele aufzunehmen. Des Weiteren, weisen die Blattscheiben nicht die Probleme von Schatten- oder Blattformigkeiten auf. Der Datensatz stellt auch wieder eine Zeitreihenmessung dar, wobei für die folgenden Analysen Aufnahmen von 0 und 96 dpi verwendet wurden. Abbildung 14 stellt die Datenvorverarbeitung dar.

Weissreferenzierung



Extraktion einzelner Blattscheiben



Klassen Aufteilungsschema

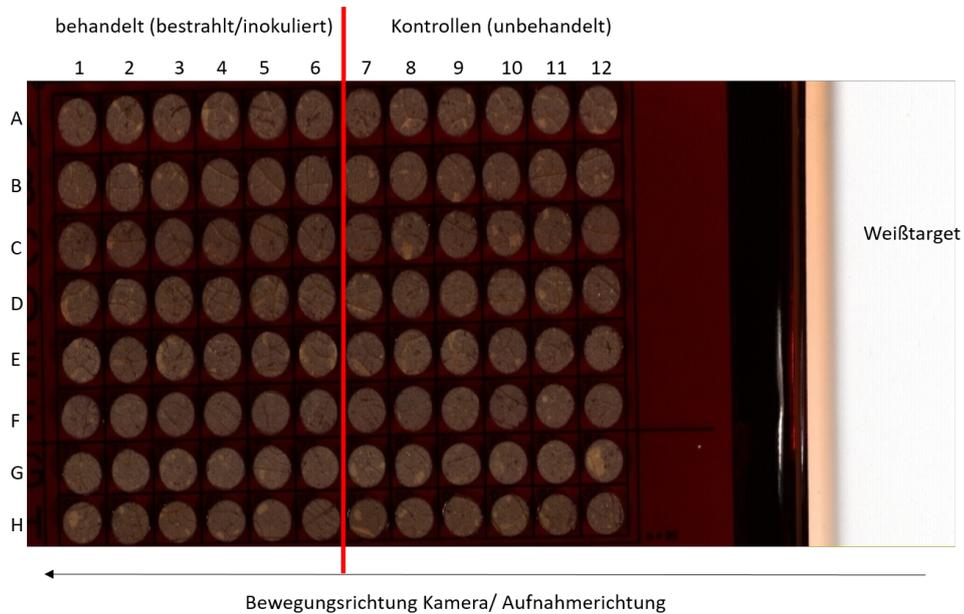


Abbildung 14: Automatisierte Extraktion der Blattscheiben des qPCR Datensatzes. Nach einer Weissnormalisierung (oben links) wurde die einzelnen Blattscheiben durch Farbfilter identifiziert, herausgeschnitten (oben rechts) und nach dem experimentellen Aufbauschema (unten) mit verschiedenen Label versehen (Kontrol vs. Behandelt, Rebsorte, etc.)

Durch die erhöhte Anzahl an Beispielen, konnten Algorithmen nun wieder für Untersuchungen auf der Bildebene verwendet werden, sodass ein Modell Bilder mit jeweils einer Blattscheibe klassifizieren musste. Tabelle 6 stellt die Ergebnisse der Klassifikationsgenauigkeiten des Gradient Tree Boosting (XGB) und eines CNNs (wie in Abschnitt A) für verschiedene Klassifikationsaufgaben. Genauer wurden hier zwischen Kontrolle und behandelte Pflanzen entschieden (Ctrl vs Inoc), Kontrolle und behandelte resistente und behandelte nicht resistente Pflanzen (Ctrl vs Res. Vs Susc) und zuletzt zwischen den Rebsorten unterschieden (Wine type). Anhand dieser Ergebnisse erkennt man die Vorteile des Ansatzes des tiefen Lernens. Allerdings, in Verbindung mit den Ergebnissen zu den UV-Vorversuchsdaten, erkennt man auch, dass es sehr von Vorteil ist genügend Beobachtungsdaten aufzunehmen, um z.B. die Effekte einer Überanpassung des Modells zu mindern.

Es ist jedoch nennenswert, dass man verschiedene Massnahmen gegen einer Überanpassung der Modelle anwenden muss, wie zum Beispiel Dropout-layer und Datenaugmentierungen (z.B. Rotationen, Cropping). Trotz der relativ hohen Anzahl an Beispielen pro Klasse in diesem Datensatz, liegt ein Problem in der hohen Dimensionalität der Bilder. So kann es sein, dass ein hochparameterisiertes Model, wie ein CNN, sich auf Eigenschaften der einzelnen Beispiele konzentriert, statt auf zugrunde liegenden Merkmale einer Klasse.

Die Ergebnisse dieser Evaluierungen sollen im Zusammenarbeit mit den Projektpartnern zu einer wissenschaftlichen Publikation führen. Dies befindet sich in laufender Arbeit.

V. Diskussion und Zusammenfassung der Ergebnisse

In Zusammenfassung, wurden über dem Projektzeitraum verschiedene Datensätze an Hand verschiedener Algorithmen analysiert. Je nach Art der Daten und experimenteller Aufbau der Aufnahme konnten verschiedene algorithmische Ansätze gefunden werden, mit denen ein Proof-of-Principle der Nutzen der Ansätze zeigen konnten. Des Weiteren, konnten mittels detaillierter Untersuchungen der Ergebnisse, Probleme im experimentellem Aufbau erkannt werden, um Analysen mittels Machine Learning Ansätze, speziell deep learning, zu optimieren.

Initiale Algorithmen, die am Beerendatensatz angewandt wurden, konnten erweitert und verbessert werden. Auch die Methoden zu erklärbare und interaktive Algorithmen konnten deutlich erweitert werden. Somit konnten AP3 und AP4, sowie die Meilensteine M-TU1, M-TU2 und M-A2 abgeschlossen werden. In Absprache mit den Projektpartnern konnten die Ergebnisse dieser Meilensteine verwendet werden, um den experimentellen Aufbau zu optimieren, sodass auch M-A1 erfolgreich abgeschlossen werden konnte.

Die Ergebnisse aller Untersuchungen zusammen, stellen den Vorteil vor allem von Algorithmen aus dem Bereich des tiefen Lernens dar. Somit konnten erfolgreich Proof-of-Principle Untersuchungen durchgeführt werden auf die Eignung solcher Methoden auf Hyperspektralphänotypisierungsdaten des Weinanbaus. Allerdings zeigen die Ergebnisse auch, dass man mit Vorsicht die Ergebnisse eines Models betrachten sollte, da ein Modell auf "falsche" Eigenschaften in den Daten fokussieren kann. Ein weiterer wichtiger Punkt ist auch, dass es von

Vorteil ist mehr Beispieldaten zu haben für diese Ansätze. Wie auch aus unseren Untersuchungen hervorgehend, könnte eine Alternative zu “viele Beispieldaten” sein, durch Feedback auf die Erklärungen eines Modells, wie in XIL, Erfahrungen und Wissen einer ExpertIn dem Model mitzuteilen, sodass der Suchraum der möglichen Optima für das Model eingeschränkt werden kann.

VI. Nutzen und Verwertbarkeit der Ergebnisse

Innerhalb des Projektes wurden verschiedene Ansätze des Tiefen Lernens für hyperspektrale Messungen von Pflanzen des Weinanbaus etabliert und bewertet. Des Weiteren wurden neue Methoden entwickelt, um mit Modellen des tiefen Lernens über ihre Erklärungen Feedback zu geben und somit sicher gehen zu können, dass die Modelle ihr Entscheidungen basierend auf wichtige Faktoren des Datensatzes treffen. Diese Methoden erlauben es, dass die Nutzer direkt in den Optimierungsprozess mit einbezogen werden.

Auf Grundlage dieser Ergebnisse können weitere wichtige Ansätze des vertrauenswürdigen tiefen Lernens entwickelt werden. Die Ansätze des tiefen Lernens zeigten vielversprechende Ergebnisse in der Rebsortenerkennung und in der Erkennung von vom falschen Mehltau befallenen Pflanzen. Mit noch mehr Datenzufuhr können diese Ansätze noch verbessert werden und auf einen praktischen Einsatz im Feld vorbereitet werden. Da die Ergebnisse bisher auf Datensätze aus einer kontrollierten Laborumgebung aufgenommen wurden, wird es weiterhin sehr wichtig sein gute Vorverarbeitungstools für den Feldeinsatz zu entwickeln (zum Beispiel eine automatische Weinpflanzendetektion, welches die Pflanzen aus der Umgebung segmentiert), als auch noch wachsamer in Bezug auf Störfaktoren in den Daten zu sein.

VII. Gegenüberstellung der ursprünglich geplanten zu den tatsächlich erzielten Zielen

Das Hauptziel von AuDiSens war es Sensor-basierte Frühdiagnoseverfahren am Beispiel des Falschen Mehltaus der Weinrebe zur Reduktion von Pflanzenschutzmittel zu untersuchen. Hierfür wurden verschiedene Algorithmen des Maschinellen Lernens untersucht, z.B. zur Krankheitserkennung aus Pflanzenbilder. Innerhalb des Projektraumes konnten alle Ziele über die Arbeitspakete und Meilensteine für die TU Darmstadt erreicht werden. Wegen Sensor-basierte Probleme konnten im Zeitraum des Projektes keine geeigneten Daten aus dem freien Feld aufgenommen werden. So konnten allerdings viele Proof-of-Principle Ergebnisse auf verschiedene Labordatensätze erzielt werden, die vielversprechende Resultate lieferten für die Anwendbarkeit der verwendeten Methoden für Krankheitsfrühdiagnose auf dem Feld. Zusätzlich konnten in der Zeit wichtige Arbeiten zur Entwicklung interaktiver und erklärbarer Algorithmen abgeschlossen werden, die für wissenschaftliche Daten viele Vorteile aufweisen.

VIII.Literaturverzeichnis

Bauckhage, C., Kersting, K., Sifa, R., Thurau, C., Drachen, A., Canossa, A.; 2009: Computational Intelligence and Games (CIG), 2012 IEEE conference, 139-146.

Brugger, A., Schramowski, P., Paulus, S., Steiner, U., Kersting, K., Mahlein, A.K. (2021): Spectral signatures in the UV-range can be combined with secondary plant metabolites by deep learning to characterise barley – powdery mildew interaction. *Plant Pathology*.

Cao, Xiangyong, Zhou, Feng, Xu, Lin, Meng, Deyu, Xu, Zongben, und Paisley, John (2017). Hyperspectral image segmentation with markov random fields and a convolutional neural network.

Fahlgren, Noah; Feldman, Maximilian; Gehan, Malia A.; Wilson, Melinda S.; Shyu, Christine; Bryant, Douglas W. et al. (2015): A Versatile Phenotyping System and Analytics Platform Reveals Diverse Temporal Responses to Water Availability in *Setaria*. In: *Molecular Plant* 8 (10), S. 1520–1535. DOI: 10.1016/j.molp.2015.06.005.

Gutiérrez, S., Fernández-Navales, J., Diago, M. P., & Tardaguila, J. (2018). On-the-go hyperspectral imaging under field conditions and machine learning for the classification of grapevine varieties. *Frontiers in plant science*, 9, 1102.

Hastie, T., Tibshirani, R., Friedman, J.(2001): The Elements of statistical learning – data mining, inference and prediction. Springer Series in Statistics. New York, NY, USA: Springer New York Inc.

Herzog, K.; Roscher, R.; Wieland, M.; Kicherer, A.; Läbe, T.; Förstner, W.; Kuhlmann, H.; Töpfer, R.; 2014: Initial steps for high-throughput phenotyping in vineyards. *Vitis*, 53, 1- 8.

Honsdorf, Nora; March, Timothy John; Berger, Bettina; Tester, Mark; Pillen, Klaus; Zhang, Tianzhen (2014): High-Throughput Phenotyping to Detect Drought Tolerance QTL in Wild Barley Introgression Lines. In: *PLoS ONE* 9 (5), e97047. DOI: 10.1371/journal.pone.0097047.

Ilse, M., Tomczak, J., & Welling, M. (2018, July). Attention-based deep multiple instance learning. In *International conference on machine learning* (pp. 2127-2136). PMLR.

Kersting, K.; Xu, Z.; Wahabzada, M.; Bauckhage, C.; Thurau, C.; Römer, C.; Ballvora, A.; Rascher, U.; Leon, J.; Plümer, L.; 2012: Pre-Symptomatic Prediction of Plant Drought Stress Using Dirichlet-Aggregation Regression on Hyperspectral Images. In *Proceedings of the AAAI Conference on Artificial Intelligence*.

Kuska, Matheus, Wahabzada, Mirwaes, Leucker, Marlene, Dehne, Heinz-Wilhelm, Kersting, Kristian, Oerke, Erich Christian, Steiner, Ulrike, und Mahlein, Anne-Katrin (2015): Hyperspectral

phenotyping on the microscopic scale: towards automated characterization of plant-pathogen interactions. *Plant Methods*.

Mahlein, A.-K.; Steiner, U.; Hillnhütter, C.; Dehne, H.-W.; Oerke, E.-C.; 2012: Hyperspectral imaging for small-scale analysis of symptoms caused by different sugar beet diseases. *Plant Methods*, 8, 1-13.

Nagasubramanian, K., Jones, S., Singh, A. K., Sarkar, S., Singh, A., & Ganapathysubramanian, B. (2019). Plant disease identification using explainable 3D deep learning on hyperspectral images. *Plant methods*, 15(1), 98.

Nauta, M., Bree, R. v., & Seifert, C. (2020). Neural Prototype Trees for Interpretable Fine-grained Image Recognition. *CoRR*, 2012.02046

Römer, C.; Wahabzada, M.; Ballvora, A.; Pinto, F.; Rossini, M.; Panigada, C.; Behmann, J.; Leon, J.; Thurau, C.; Bauckhage, C.; Kersting, K.; Rascher, U.; Plümer, L.; 2012: Early Drought Stress Detection in Cereals: Simplex Volume Maximization for Hyperspectral Image Analysis. *Functional Plant Biology*, 39, 878-890.

Ross, A. S., Hughes, M. C., & Doshi-Velez, F. (2017). Right for the Right Reasons: Training Differentiable Models by Constraining their Explanations. *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, 2662–2670

Skrlj, B., Dzeroski, S., Lavrac, N., & Petkovic, M. (2020). Feature Importance Estimation with Self-Attention Networks. *Proceedings of ECAI 2020 - 24th European Conference on Artificial Intelligence*, 325 ,1491–1498

Teso, S., & Kersting, K. (2019). Explanatory Interactive Machine Learning. *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 239–245

Wahabzada M, Mahlein AK, Bauckhage C, Steiner U, Oerke EC, Kersting K.; 2015a: Metro maps of plant disease dynamics-automated mining of differences using hyperspectral images. *PLoS One*. 2015 Jan 26;10(1):e0116902. doi: 10.1371/journal.pone.0116902.

Wahabzada, M.; Paulus, S.; Kersting, K.; Mahlein, A.-K.; **2015b**: Automated Interpretation of 3D Laserscanned Point Clouds for Plant Organ Segmentation. *BMC Bioinformatics*, 16:248.

Zhong, Zilong, Li, Jonathan, Ma, Lingfei, Jiang, Han, und Zhao, He (2017). Deep residual networks for hyperspectral image classification.

IX. Veröffentlichungen zum Projekt

Patrick Schramowski, Wolfgang Stammer, Stefano Teso, Anna Brugger, Franziska Herbert, Xiaoting Shao, Hans-Georg Luigs, Anne-Katrin Mahlein, Kristian Kersting (2020): Making deep neural networks right for the right scientific reasons by interacting with their explanations. Nature Machine Intelligence.

Wolfgang Stammer, Patrick Schramowski, Kristian Kersting (2021): Right for the Right Concept: Revising Neuro-Symbolic Concepts by Interacting with their Explanations. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Xiaoting Shao, Arseny Skryagin, Patrick Schramowski, Wolfgang Stammer, Kristian Kersting (2021): Right for better reasons: Training differentiable models by constraining their influence function. In Proceedings of 35th AAAI Conference on Artificial Intelligence (AAAI).

Wolfgang Stammer, Marius Memmel, Patrick Schramowski, Kristian Kersting (2022): Interactive Disentanglement: Learning Concepts by Interacting with their Prototype Representations. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Folgende Publikation ist bei einem wissenschaftlichen Journal oder Konferenz eingereicht:

Arseny Skryagin, Wolfgang Stammer, Daniel Ochs, Devendra Singh Dhami, Kristian Kersting (2021). SLASH: Embracing Probabilistic Circuits into Neural Answer Set Programming. arXiv preprint arXiv:2110.03395.