

**Erfassung struktureller und chromatischer
Pflanzenmerkmale mittels intelligenter,
bildgebender 2D Verfahren im Bereich der
Gemüsezüchtung**

Dissertation

zur

Erlangung des akademischen Grades

Doktor der Agrarwissenschaften (Dr. agr.)

der Agrar-, Ernährungs- und Ingenieurwissenschaftlichen Fakultät
der Rheinischen Friedrich–Wilhelms–Universität Bonn

vorgelegt von

Dirk Jollet
aus Simmerath

Bonn 2025

Referent:

Prof. Dr. Uwe Rascher

Korreferenten:

Prof. Dr. Ribana Roscher

Prof. Dr. Eike Lüdeling

Tag der mündlichen Prüfung: 12.11.2024

Angefertigt mit Genehmigung der der Agrar-, Ernährungs- und Ingenieurwissenschaftlichen
Fakultät der Universität Bonn

Zusammenfassung

Seit Beginn der Landwirtschaft beruht die vom Menschen gesteuerte Entwicklung der Kulturpflanzen auf Auslese und Rekombination. Ziel dieses Prozesses ist es einen Genotyp zu schaffen, der optimal an die jeweiligen Umweltbedingungen angepasst ist und im Phänotyp bestimmte Merkmale, wie beispielsweise einen hohen Ertrag, oder bestimmte strukturelle oder chromatische Merkmale, ausprägt. Die Bewertung dieser Merkmale beruht traditionell auf einer visuellen Beurteilung des Pflanzenmaterials, teilweise unter Zuhilfenahme spezieller Messinstrumente. Einerseits ist es heutzutage möglich, durch die Entwicklung moderner gentechnischer Verfahren, in kürzester Zeit unzählige neue Genotypen zu schaffen. Andererseits nähern sich die Phänotypen moderner Kultursorten in Bezug auf züchtungsrelevante Merkmale immer näher an, sodass heutzutage sehr feine Unterschiede in den Merkmalsausprägungen ausschlaggebend für die Selektionsentscheidung sind. Aufgrund dieser Entwicklungen wird die Bewertung von züchtungsrelevanten Merkmalen mit traditionellen Bewertungsmethoden immer schwieriger, was besonders stark im Bereich der Gemüsezüchtung in Erscheinung tritt, da hier strukturelle und chromatische Pflanzenmerkmale besonders wichtig für die Vermarktung sind.

In dieser Arbeit wurde anhand von vier dedizierten und komplexen Forschungsfragen aus dem Bereich der Gemüsezüchtung untersucht, ob durch die Analyse von 2D-Bilddaten aus RGB-Farbkameras mittels moderner Methoden des maschinellen Sehens die Limitierungen traditioneller Evaluierungsmethoden überwunden werden können.

1. Es wurde ein System, inklusive der Konstruktion einer Fotobox und der Entwicklung einer Bildanalyse Pipeline, etabliert, welches präzise Hochdurchsatzanalysen für die Länge, das Kaliber und die Krümmung länglicher Gemüsefrüchte ermöglicht. Das entwickelte System bietet einen erhöhten Durchsatz bei verbesserter Genauigkeit und Objektivität für die Vermessung der genannten Merkmale.
2. Es wurde erstmalig ein auf transmittierendem Licht beruhendes 2D RGB-Verfahren zur nicht-invasiven Evaluierung der Samenanlagen von Leguminosen entwickelt und evaluiert. Das Verfahren funktioniert für lichtdurchlässige gelbe Sorten, sowie helle grüne Sorten mit gut entwickelten Samen sehr zuverlässig. Hardwareanpassungen können die Analyse der Samenanlagen in dunklen Hülsen mit kleinen Samen zukünftig weiter verbessern.
3. Es wurden umfangreiche anwendungsbezogene Farbanalysen angestellt, um die Möglichkeiten von 2D RGB-Verfahren für die objektive Evaluierung von chromatischen Sorteneigenschaften und Stresssymptomen abzuschätzen. So konnten die Funktionalität von globalen und lokalen Farbkalibrierungsmethoden, und deren mögliche Anwendungsgebiete, aufgezeigt werden. Des Weiteren lieferten die hier etablierten Methoden des maschinellen Sehens grundsätzlich glaubwürdige Ergebnisse zur Abschätzung von Stresssymptomen an Möhrenblättern oder der Farb- und Helligkeitsanalyse von Bohnenhülsen. Nach den Erkenntnissen dieser Arbeit wird es jedoch von Fall zu Fall abhängen, ob 2D RGB-Verfahren die traditionelle visuelle Inspektion ersetzen kann, da der Erfolg von vielerlei Faktoren abhängt.
4. Es wurde ein System, inklusive der Konstruktion eines nicht-motorisierten Feldfahrzeugs und der Entwicklung dedizierter Bildanalyse Algorithmen, entwickelt, welches zukünftig automatisierte und nicht-invasive Hochdurchsatz Ertragsabschätzungen bei kopfbildendem Blattgemüse unter Feldbedingungen ermöglicht. Durch die Nutzung dieses System können deutlich höhere Stichprobengrößen vermessen werden, was grundsätzlich präzisere Rückschlüsse auf den Ertrag einer Sorte ermöglicht und zusätzlich Zeitreihenanalyse einzelner Pflanzen zulässt. Die Präzision der Gewichtsvorhersage einzelne Köpfe ist durch die alleinige Nutzung der projizierten Blattfläche limitiert, kann zukünftig jedoch durch komplexere

Vorhersagemodelle, welche weitere Sorten- oder Standortinformationen hinzuziehen, verbessert werden.

Die Entwicklungen der letzten Jahre im Bereich des maschinellen Sehens sind ein Katalysator für die Merkmalsevaluierung in der modernen Gemüsezüchtung. Beachtet man die technischen Möglichkeiten eines 2D RGB-Verfahrens, stellt dieses in vielen Fällen eine kostengünstige Lösung für eine Objektivierung der Evaluierungen, einen Transfer von invasiven zu nicht-invasiven Verfahren, eine Durchsatzsteigerung, oder eine Steigerung der Genauigkeit dar.

Abstract

Since the beginning of agriculture, the human-controlled development of cultivated plants has been based on selection and recombination. The aim of this process is to create a genotype that is optimally adapted to the respective environmental conditions and expresses certain characteristics in the phenotype, such as a high yield or certain structural or chromatic characteristics. The evaluation of these characteristics is traditionally based on a visual assessment of the plant material, sometimes with the aid of special measuring instruments. On the one hand, today it is possible to create countless new genotypes in a very short time thanks to the development of modern genetic engineering methods. On the other hand, the phenotypes of modern cultivars are coming ever closer together in terms of breeding-relevant traits. Therefore, nowadays very subtle differences in trait expression are decisive for selection decisions. Due to these developments, the evaluation of breeding-relevant traits with traditional evaluation methods is becoming increasingly difficult, which is particularly evident in the field of vegetable breeding, where structural and chromatic plant characteristics are particularly important for marketing.

In this thesis, four dedicated and complex research questions from the field of vegetable breeding were investigated to determine whether the limitations of traditional evaluation methods can be overcome by analyzing 2D image data from RGB color cameras using modern machine vision methods.

1. A system, including the construction of a photo box and the development of an image analysis pipeline, was established that enables precise high-throughput analysis for the length, caliber, and curvature of elongated vegetable fruits. The developed system offers increased throughput with improved accuracy and objectivity for the measurement of these characteristics.
2. For the first time, a 2D RGB method based on transmitting light was developed and evaluated for non-invasive evaluation of legume seed arrangement within the pod. The method works very reliably for translucent yellow varieties and light-colored green varieties with well-developed seeds. Hardware adaptations can further improve the analysis of seed arrangement in darker pods with smaller seeds in the future.
3. Extensive application-related color analyses were carried out to assess the possibilities of 2D RGB methods for the objective evaluation of chromatic variety characteristics and stress symptoms. Thus, the functionality of global and local color calibration methods and their possible fields of application could be demonstrated. Furthermore, the machine vision methods established here provided fundamentally credible results for the estimation of stress symptoms on carrot leaves or the color and brightness analysis of bean pods. However, according to the findings of this work, it will depend on a case-by-case basis whether 2D RGB methods can replace traditional visual inspection, as success depends on many factors.
4. A system, including the design of a non-motorized field vehicle and the development of dedicated image analysis algorithms, was developed to enable automated and non-

invasive high-throughput yield estimation of head-forming leafy vegetables under field conditions in the future. By using this system, significantly larger sample sizes can be measured, which enables more precise conclusions to be drawn about the yield of a variety. Furthermore, it allows time series analysis of individual plants. The precision of the weight prediction of individual heads is limited by the sole use of the projected leaf area, but can be improved in the future by more complex prediction models that include further variety or location information.

The developments of recent years in the field of machine vision are a catalyst for trait evaluation in modern vegetable breeding. Considering the technical possibilities of a 2D RGB method, it is in many cases a cost-effective solution for objectifying evaluations, transferring from invasive to non-invasive methods, increasing throughput or improving accuracy.

Inhaltsverzeichnis

1	Einführung	7
1.1	<i>Die Grundzüge der Pflanzenzüchtung.....</i>	7
1.2	<i>Ertrag und Qualität und deren Sekundärziele in der Gemüsezüchtung</i>	8
1.3	<i>Kontext der Arbeit.....</i>	10
1.4	<i>Forschungsfragen</i>	11
1.4.1	Präzise Hochdurchsatzanalyse wichtiger Formparameter länglicher Gemüsesorten	11
1.4.2	Nicht-invasive Analyse der Samenanlage von Leguminosen	12
1.4.3	Chromatische Analyse von RGB-Bilddaten zur Evaluierung von Sorteneigenschaften sowie Stresssymptomen in der Gemüsezüchtung	12
1.4.4	Automatisierte und nicht-invasive hochdurchsatz-Ertragsabschätzung bei kopfbildendem Blattgemüse.....	13
2	Grundlagen des maschinellen Sehens (MS).....	14
2.1	<i>Elektromagnetisches Spektrum</i>	14
2.2	<i>Farbräume</i>	14
2.2.1	RGB-Farbraum.....	14
2.2.2	HSV-Farbraum	15
2.2.3	CIE-LAB-Farbraum	15
2.3	<i>Klassische Bildverarbeitung</i>	16
2.3.1	Geometrische und chromatische Transformationen	17
2.3.2	Bildfilter.....	18
2.3.3	Texturanalysen	19
2.3.4	Schwellenwertsegmentierung	20
2.3.5	Morphologische Operatoren.....	20
2.3.6	Formdeskriptoren	21
2.4	<i>Ansätze des maschinellen Lernens.....</i>	22
2.4.1	Theorie des statistischen Lernens.....	22
2.4.2	Regression und Interpolation.....	23
2.4.3	Klassifikation	24
2.5	<i>Ansätze des tiefen Lernens</i>	24
2.5.1	Künstliche neuronale Netze mit voller Konnektivität.....	24
2.5.2	Neuronale Faltungsnetzwerke	26
2.6	<i>Statistische Werkzeuge zur Methodvalidierung</i>	28
3	Analyse wichtiger Zuchtmerkmale mittels maschinellen Sehens	30
3.1	<i>Anmerkung zur Zitation der eigenen Veröffentlichungen</i>	30
3.2	<i>Präzise Hochdurchsatzanalyse wichtiger Formparameter länglicher Gemüsesorten</i>	30
3.2.1	Forschungsstand	30
3.2.2	Material und Methoden.....	32
3.2.3	Ergebnisse	41
3.2.4	Diskussion	45
3.3	<i>Nicht-invasive Analyse der Samenanlage von Leguminosen</i>	47
3.3.1	Forschungsstand	47
3.3.2	Material und Methoden.....	48
3.3.3	Ergebnisse	50
3.3.4	Diskussion	51

3.4	<i>Chromatische Analyse von RGB-Bilddaten zur Evaluierung von Sorteneigenschaften, sowie Stresssymptomen in der Gemüsezucht</i>	52
3.4.1	Forschungsstand	53
3.4.2	Farbklassifikation und Helligkeitsanalyse unter kontrollierten Bedingungen zur Sortenbeurteilung	55
3.4.3	Bewertung komplexer biotischer Stresssymptome unter kontrollierten Bedingungen	58
3.4.4	Bewertung chromatischer Pflanzeigenschaften unter Feldbedingungen	66
3.5	<i>Automatisierte und nicht-invasive hochdurchsatz-Ertragsabschätzung bei kopfbildendem Blattgemüse</i>	73
3.5.1	Forschungsstand	73
3.5.2	Material und Methoden	75
3.5.3	Ergebnisse	79
3.5.4	Diskussion	81
4	Fazit und Ausblick	83
4.1	<i>Präzise Hochdurchsatzanalyse wichtiger Formparameter länglicher Gemüsesorten</i>	83
4.2	<i>Nicht-invasive Analyse der Samenanlage von Leguminosen</i>	84
4.3	<i>Chromatische Analyse von RGB-Bilddaten zur Evaluierung von Sorteneigenschaften sowie Stresssymptomen in der Gemüsezucht</i>	85
4.4	<i>Automatisierte und nicht-invasive Hochdurchsatz-Ertragsabschätzung bei kopfbildendem Blattgemüse</i>	87
5	Résumé	89
6	Danksagung	91
7	Abkürzungsverzeichnis	92
8	Literaturverzeichnis	93
9	Abbildungsverzeichnis	104
10	Tabellenverzeichnis	107

1 Einführung

1.1 Die Grundzüge der Pflanzenzüchtung

Der Mensch begann vor ca. 12000 Jahren an mehreren Orten gleichzeitig Landwirtschaft zu betreiben, was durch eine globale Erwärmung des Klimas und dem Rückgang der Eisschilde zu dieser Zeit initiiert wurde (Gignoux et al., 2011; Velimirovic et al., 2021). Im Zuge dessen wurde die natürliche Vegetation zurückgedrängt und der Boden für den Anbau von Wildpflanzen wie beispielsweise Emmer, Einkorn, Erbsen, Linsen und Flax in Europa, Teosinte, Maniok, Bohnen, Kartoffeln, und Tomaten in Amerika, Sorghum und Hirse in Afrika, oder Reis in Asien bearbeitet (Fagan et al., 1994; Zeder, 2011). Dabei ging der Anbau mit einer ständigen Auslese der Wildpflanzen einher, wodurch sich überlebenswichtige Grundeigenschaften der Wildarten veränderten. Beispielsweise vergrößerten sich die Fruchtkörper im Laufe der Zeit im Vergleich zu den abgestammten Wildarten, viele Samen verloren die Dormanz und es wurde gegen die natürlichen Ausbreitungsmechanismen der Samenanlagen selektiert. Letzteres führte beispielsweise bei Gräsern zu einer verringerten Spindelbrüchigkeit oder bei Leguminosen zu Hülsen, welche bei der Trockenreife nicht aufplatzen (Becker, 2019). Die frühen Kulturarten verloren somit wichtige Eigenschaften für eine erfolgreiche Vermehrung, boten dem Menschen jedoch ernährungsphysiologische und anbautechnische Vorteile im Vergleich zu den Wildarten (Becker, 2019). Dies lässt sich gut am Beispiel des Mais verdeutlichen: Der natürliche Vorfahre des Mais ist ein Tieflandgras namens Teosinte (*Zea mays ssp. parviglumis*), welches in Mesoamerika beheimatet ist. Zu Beginn der Domestikation nutzten die Menschen wahrscheinlich natürliche Mutationen, welche an Einzelpflanzen innerhalb einer Population auftraten, zur Auslese. Im weiteren Verlauf der Domestikation wählten die Menschen innerhalb einer Population jene Pflanzen für die Vermehrung aus, welche in bestimmten Merkmalen die gewünschte Ausprägung hatten. Durch stetige Auslese, also einer Verringerung des Genpools einer Population, und Rekombination, einer Erweiterung des Genpools durch Einbringung von Genen einer anderen Population durch Einkreuzung, entstand aus der Wildart Teosinte im Laufe der Zeit die Kulturart Mais (*Zea mays L.*) (Becker, 2019; Yang et al., 2019). Abbildung 1 zeigt die phänotypischen Unterschiede der Pflanzen (a), sowie der Fruchtkörper (b) zwischen Teosinte und Mais. So bildet Teosinte viele Triebe aus, welche mehrere Hundert Fruchtkörper mit 10 Samen tragen, wohingegen Mais nur 1 bis 2 Triebe ausbildet, welche 2 Fruchtkörper mit vielen Hundert Samen tragen. Während Teosinte unter ungünstigen Wachstumsbedingungen weniger Triebe bildet, bleibt die Wuchsform des Mais stabil, was eine einfache Ernte der Fruchtkörper durch den menschlichen Kultivator sicherstellt (Yang et al., 2019).

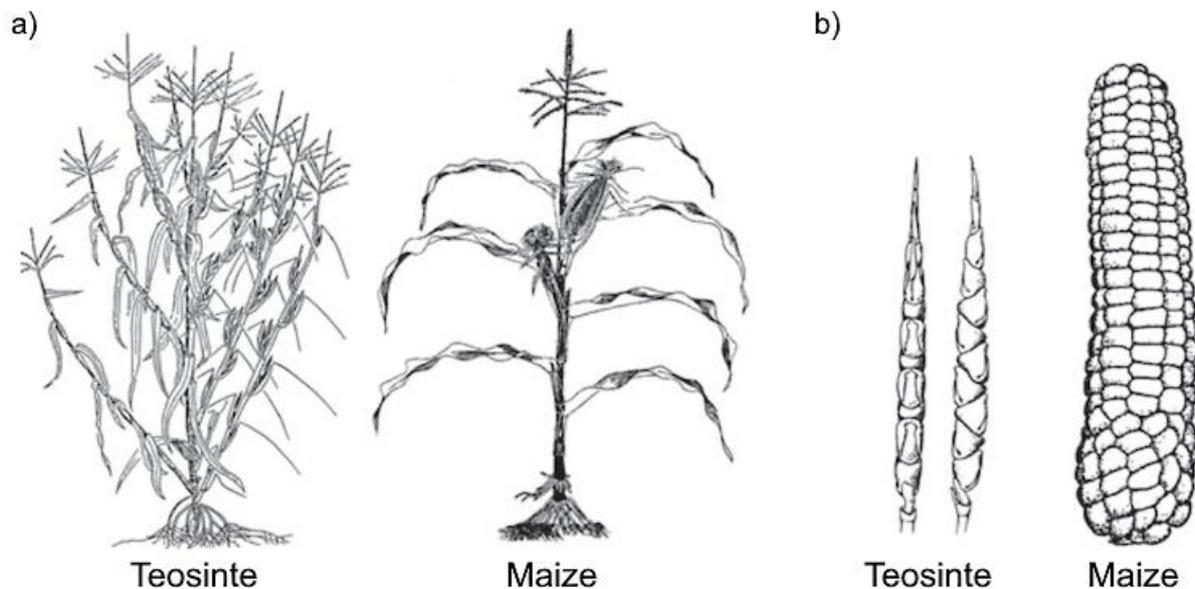


Abbildung 1: Phänotypischen Unterschiede der Pflanzen (A), sowie der Fruchtkörper (B) zwischen Teosinte und Mais. Quelle: (Yang et al., 2019)

Zwar lief die Domestikation für jede Kulturart individuell unterschiedlich ab, im Grunde sorgte jedoch immer ein stetiger Prozess der wiederkehrenden Auslese und anschließenden Rekombination in Kombination mit der Verbreitung der Landwirtschaft im Lauf der Jahrtausende für die Entstehung vielerlei Landsorten der unterschiedlichen Kulturarten (Becker, 2019). Im Mittelpunkt der Auslese standen dabei stets die primären Zuchtziele *Ertrag* und *Qualität*, wobei *Ertrag* die Menge an essbarer Biomasse beschreibt, die geerntet werden kann und *Qualität* für jede Kulturart individuell definiert werden muss (Arakeri & Lakshmana, 2016; Araújo et al., 2015; Becker, 2019; Blasco et al., 2003; Lee et al., 2011; Zhu et al., 2019). *Ertrag* und *Qualität* stellen auch in heutigen Zuchtprogrammen noch die primären Zuchtziele dar, welche um sekundäre Zuchtziele ergänzt werden. Im Falle des Ertrags haben bereits viele Kulturarten die biologisch maximale Größe ihrer Fruchtkörper erreicht, sodass man nun beispielsweise auf die Sekundärziele Krankheitsresistenz oder Toleranz gegenüber abiotischen Stressfaktoren züchtet, da diese Faktoren den Realertrag negativ beeinflussen können (Becker, 2019). Da *Qualität* kultursortenspezifisch definiert wird, müssen auch deren sekundäre Zuchtziele individuell festgelegt werden. Dies lässt sich gut am Beispiel des Weizens verdeutlichen: So stellt in der Weizenzucht die Backfähigkeit das wichtigste Qualitätsmerkmal dar, welche sich jedoch nicht direkt erfassen lässt. Es ist jedoch gut erforscht, dass die Backfähigkeit des Weizenkorns hauptsächlich von der Zusammensetzung der Speicherproteine beeinflusst wird, welche mittels Elektrophorese bestimmt werden können. Bestimmte Elektrophorese Muster werden somit bei der Züchtung berücksichtigt und stellen ein sekundäres Zuchtziel für das primäre Zuchtziel der Backfähigkeit dar (Becker, 2019). Im Folgenden wird nun auf die Besonderheiten in der Gemüsezüchtung eingegangen.

1.2 Ertrag und Qualität und deren Sekundärziele in der Gemüsezüchtung

Grundsätzlich definiert sich der Ertrag auch in der Gemüsezüchtung als Menge an essbarer Biomasse, die eine Kulturart produzieren kann. Die Erhebung dieser Größe lässt sich jedoch, wie bei allen Kulturarten, auf verschiedenste Weise realisieren (Becker, 2019). So kann man den Ertrag über das Gewicht des Ernteguts pro Flächeneinheit bestimmen, oder man untersucht die Anzahl Fruchtkörper pro Pflanze oder das Gewicht einzelner Fruchtkörper. Auch in der Gemüsezüchtung stellen Krankheitsresistenzen oder Toleranzen sekundäre Zuchtziele für den Ertrag dar. Um beispielsweise Resistenzen gegenüber Blattkrankheiten zu evaluieren,

werden Pflanzen unterschiedlicher Sorten künstlich mit dem gewünschten Krankheitserreger infiziert und schließlich zu einem bestimmten Zeitpunkt Bewertungen durchgeführt. Dabei untersucht ein erfahrener Zuchtextperte die Pflanzen visuell und schätzt dabei den Gesundheitszustand des Blattwerks ein. Werden diese Tests an verschiedenen Standorten durchgeführt, kristallisieren sich Sorten mit stärkeren Resistenzen heraus. Zwar gibt es keine klare Definition, wie der Gesundheitszustand eine Pflanze zu bewerten ist, es fließen jedoch Parameter wie das Verhältnis zwischen gesundem, chlorotischem, und nekrotischem Blattmaterial, der Turgordruck der Blätter oder das allgemeine Wachstum in die Bewertung des Züchters mit ein (Bock et al., 2022).

Die Qualität wird bei vielen Kulturarten fast ausschließlich über physiologische Eigenschaften definiert. Beim Weizen sind, wie zuvor erwähnt, die Backfähigkeit oder Proteingehalte des Korns wichtige Qualitätsmerkmale. Ein anderes Beispiel gibt der Raps: So enthielten ursprüngliche Sorten einen hohen Gehalt an Erucasäure, welche für den Menschen als Gesundheitsschädlich eingestuft wurden (Becker, 2019). Ein geringer Gehalt dieser Säure stellt somit ein wichtiges Qualitätskriterium in der Rapszüchtung dar. In der Gemüsezüchtung stehen neben physiologischen Qualitätsmerkmalen wie beispielsweise dem Geschmack auch phänotypische Qualitätsmerkmale des Ernteguts im Vordergrund (Ilakiya et al., 2020; Tripathi & Maktedar, 2020; Zheng et al., 2022). Dies hängt natürlicherweise damit zusammen, dass Gemüse nicht zwangsläufig destruktiv weiterverarbeitet wird und schnell verderblich ist. Dies ist bei Weizen oder Raps Saaten nicht der Fall, und während Weizen hauptsächlich zu Mehl gemahlen und Raps zu Öl gepresst wird, tritt das Erntegut im Gemüsesektor oft un- oder minimalverarbeitet in den Markt ein. Dabei gibt es 3 unterschiedliche Abnehmer für frisch geerntetes Gemüse: Zunächst den Endverbraucher, der das Gemüse entweder direkt vom Erzeuger oder aus Supermärkten bezieht. Neben ökologischen Aspekten wie kurzen Lieferketten, oder nachhaltige Anbaumethoden, beeinflussen vor allem haptisch-visuelle Reize das Kaufverhalten des Endverbrauchers (Herath, 2019; S. M. E. Rahman et al., 2021). So muss das Produkt im Regal optisch ansprechend und homogen erscheinen und frei von Verwachsungen, Schäden, oder Krankheitssymptomen sein. So sollten beispielsweise Bohnenhülsen eine möglichst geringe Krümmung und homogene Ausprägungen in Länge, Kaliber und Farbe besitzen. Der Einzelhandel stellt das Bindeglied zwischen Erzeuger und Endverbraucher dar, welcher besondere Ansprüche in Bezug auf die Logistik stellt. Dies betrifft einerseits die Transportfähigkeit beziehungsweise Lagerfähigkeit des Ernteguts, andererseits sollte die Form des Ernteguts so sein, dass dieses effizient verpackt und transportiert werden kann (Tort et al., 2022). Dies zeigt sich beispielsweise daran, dass bis heute hauptsächlich gerade Gurken in Supermärkten zu finden sind, obwohl die diesbezügliche EU-Vorordnung, welche nur gerade Gurken für die Vermarktung in der EU zuließ, bereits 2009 ausgesetzt wurde (Verordnung (EWG) Nr. 1677/88). Der letzte große Abnehmer für frisches Gemüse ist die verarbeitende Industrie, welche oftmals einen verringerten Anspruch an die Erscheinung des Ernteguts als viel Mehr an dessen Verarbeitbarkeit stellt. Einen wichtigen Zweig stellt hierbei die Verarbeitung im Sinne der Konservierung dar. Hierzu zählen beispielsweise Säfte (z.B. Möhrensaft), Pasten (z.B. Tomatenmark), sowie eingelegte (z.B. rote Bete) beziehungsweise fermentierte Konserven (z.B. Sauerkraut). Für die industrielle Verarbeitung ist es zunächst wichtig, dass das Erntegut möglichst sauber und frei von anderen Pflanzenteilen wie beispielsweise Blütenansätzen oder Stielen in der Fabrik ankommt, um die Anzahl der Schritte während der Verarbeitung zu minimieren¹. Die rückstandslose Erntbarkeit mittels maschineller Erntetechniken kann somit auch ein wichtiges Qualitätsmerkmal in der Gemüsezüchtung sein. Weiterhin muss das Erntegut gut maschinell verarbeitbar sein, wofür eine bestimmte Form oder Größe der Früchte von Nöten sein kann.

¹ Latauschke, 2021 - Mittelfeine Buschbohnen bei sehr guten Bedingungen mit Spitzenerträgen
https://www.gartenbau.sachsen.de/download/Buschbohnen_mittelfein_2021.pdf (20.02.2024)

In der Gemüsezüchtung lassen sich Ertrags- und Qualitätsmerkmale meist auf strukturelle (Form) sowie chromatische (Farbe) Merkmale des Ernteguts oder des Blattwerks reduzieren, welche während eines Züchtungsprogramms untersucht und bewertet (bonitiert) werden. So kann beispielsweise, wie oben bereits beschrieben, die Bewertung der Resistenz einer Sorte auf der Evaluierung der Farbe des Blattwerks beruhen. Zwar stellt die Farbe grundsätzlich ein artenübergreifendes Merkmal dar, jedoch beruht dessen Bewertung traditionell auf der visuellen Bewertung durch einen erfahrenen Zuchtextperten. Diese Art der Evaluierung ist zeitaufwändig und benötigt hochqualifiziertes Personal, wird aber auch bei einigen kulturspezifischen Strukturmerkmalen angewandt. So stellt für viele Kulturarten mit länglichen Fruchtkörpern wie Gurken, Zucchini, Möhren, oder Bohnenhülsen die Krümmung ein wichtiges Qualitätskriterium dar, welches nur visuell evaluiert werden kann, da es keine Messinstrumente gibt, um die Krümmung eines Objekts physisch zu messen. Visuelle Bewertungen unterliegen dabei stets der individuellen Wahrnehmung des Boniteurs und können vor allem bei Bewertungen der Farbe durch Umweltbedingungen wie Lichtstärke oder Lichtqualität und deren Änderungen beeinflusst werden. Folglich können visuelle Bewertungen nicht nur zwischen verschiedenen Boniteuren variieren, sondern auch die Bewertungen eines Boniteurs können merklich (mitunter auch auf einer täglichen Basis) schwanken (Bock et al., 2022; Miller-Butler et al., 2019; Simko & Hayes, 2018). Die Problematik der Subjektivität des Betrachters gilt ebenfalls, wenn Merkmale haptisch evaluiert werden müssen. Ein konkretes Beispiel hierfür ist die Bewertung der Kornanlage von Bohnenhülsen. So ist es für den Frischmarkt wichtig, dass die Samenanlage vollständig ist, da die Hülse sonst eine inhomogene Erscheinung aufweist, was nicht erwünscht ist (King & Vining, 2019). Im Stadium der Grünreife sind die Samen in der Hülse jedoch noch so klein, dass die Hülse meist gering oder kaum markiert, sodass eine haptische Bewertung zwar ebenfalls fehlerbehaftet sein kann, jedoch die einzige sinnvolle Möglichkeit zur Bewertung der Samenanlage darstellt. Anders verhält es sich bei Strukturmerkmalen, welche mit Hilfe von Messinstrumenten gemessen werden können. So können beispielsweise die Dimensionen von Fruchtkörpern mit Linealen oder Schieblehren gemessen werden. Zwar können solche Messungen sehr präzise Ergebnisse liefern, wenn ein klar definiertes Messprotokoll eingehalten wird. Jedoch sind solche Messungen auch immer sehr zeitintensiv, wodurch ein Flaschenhals in Züchtungsprogrammen entsteht (Ilakiya et al., 2020).

1.3 Kontext der Arbeit

Wie im vorherigen Abschnitt beschrieben, ergeben sich durch die Vielzahl an Kulturarten im Gemüsesektor und die vielseitigen Ansprüche der verschiedenen Interessengruppen eine große Anzahl kulturspezifischer Merkmale, die bei der Züchtung neuer Gemüsesorten beachtet und bewertet werden müssen. Aus diesem Grund wird sich diese Arbeit auf vier Forschungsfragen aus der Domäne der Gemüsezüchtung konzentrieren, bei denen traditionelle Bewertungsmethoden entweder in Bezug auf die Geschwindigkeit, Objektivität oder die Genauigkeit an ihre Grenzen stoßen. Im Zuge dieser Arbeit soll untersucht werden, wie mittels 2D-Bilddaten aus RGB-Farbkameras (nachfolgend als 2D RGB bezeichnet) in Kombination mit Methoden aus dem Bereich des maschinellen Sehens die Limitierungen der traditionellen Evaluierungsmethoden für die vorgestellten Forschungsfragen überwunden werden. Hierfür wurde bewusst ein kostengünstiger Sensor gewählt mit dem Potential, das System erfolgreich in die Züchtungspraxis zu überführen. Weitere Vorteile des 2D RGB-Verfahrens für die praktische Pflanzenzüchtung sind (gerade im Vergleich zu anderen bildgebenden Verfahren wie multi- bzw. hyperspektral, thermal, oder 3D) praktische Aspekte wie eine schnelle Bildaufnahme, hohe Auflösung, und eine vergleichsweise einfache Implementierung (Bodner et al., 2017; Ilakiya et al., 2020).

Die Entwicklung der Forschungsthemen dieser Arbeit erfolgte im Rahmen einer Forschungskoooperation zwischen dem Institut für Pflanzenwissenschaften (IBG-2, *Forschungszentrum Jülich GmbH* mit Sitz in 52428 Jülich, West-Deutschland) und drei Züchtungsunternehmen aus dem Bereich der Gemüsezüchtung: *Van Waveren* ist ein Kleinunternehmen mit rund 30 Mitarbeitern am Standort 37124 Rosdorf, Zentral-Deutschland. Das Unternehmen besitzt Expertise in der Zucht von Hülsenfrüchtlern wie Erbsen (*Pisum Sativum*) und Gartenbohnen (*Phaseolus vulgaris*). *Nunhems* ist Teil der BASF-Gruppe und hat sich auf die Zucht verschiedener Gemüsesorten spezialisiert. Zu ihren Kernkulturarten gehören unter anderem Gurken (*Cucumis sativus*), Tomaten (*Solanum lycopersicum*), Gemüsespinat (*Spinacia oleracea*), sowie Möhren (*Daucus Carota*). Das Unternehmen mit Hauptsitz in 6083 AB Nunhem, Niederlande, besitzt weltweit rund 2000 Mitarbeiter. *Rijk Zwaan Marne* ist ein Ableger der Niederländischen Muttergesellschaft *Rijk Zwaan Zaadteelt en Zaadhandel B.V.* mit Hauptsitz in 2678 KX De Lier, Niederlande. Das Unternehmen beschäftigt weltweit ca. 3800 Mitarbeiter und hat sich auf die Zucht von Tomaten, Melonen, Auberginen, Paprika, Salaten (*Lactuca sativa*), sowie Gemüse Kohl (*Brassica oleracea*) spezialisiert. Am Standort 25709 Marne in Nord-Deutschland werden ausschließlich Züchtungsprogramme für den Gemüse Kohl durchgeführt.

Die Kooperation bietet einmalige Einblicke hinter die Kulissen der Züchtungsunternehmen und zeigt aktuelle Bedarfe für das Forschungsgebiet der digitalen Pflanzenphänotypisierung. Ziel dieser Arbeit ist es aktuelle Erkenntnisse aus unterschiedlichen Teilbereichen dieses Forschungsgebiets zu bündeln und zu einer Lösung der Forschungsfragen auszubauen, um somit einen Beitrag für die Gesellschaft zu leisten.

Zunächst wird für jede dieser Forschungsfragen eine Literaturrecherche erfolgen, um den aktuellen Forschungsstand auf dem jeweiligen Teilbereich des Forschungsgebiets der digitalen Pflanzenphänotypisierung im Bereich der Gemüsezüchtung aufzuzeigen. Es werden Methoden aus dem Bereich des maschinellen Sehens identifiziert, welche zur Lösung der Forschungsfrage angepasst oder weiter ausgebaut werden. Schließlich wird eine konkrete Lösung für jede der Forschungsfragen präsentiert und im Kontext des jeweiligen Teilbereichs des Forschungsgebiets bewertet. Zuletzt wird ein Fazit über den Beitrag dieser Arbeit für das Forschungsgebiet der digitalen Pflanzenphänotypisierung in der Gemüsezüchtung gezogen sowie ein Ausblick in Bezug auf die behandelten Forschungsfragen gegeben.

1.4 Forschungsfragen

1.4.1 Präzise Hochdurchsatzanalyse wichtiger Formparameter länglicher Gemüsesorten

An längliche Gemüsesorten wie Möhre, Gurke, Zucchini, Aubergine, oder Bohnenhülsen werden teils sehr ähnliche Ansprüche in Bezug auf die Form gestellt. So ist für alle diese Kulturarten die Information über die Länge und den Durchmesser der Frucht, sowie deren Krümmung, essenziell. Die Messung der Länge und des Durchmessers mit Hilfe physischer Messmethoden ist jedoch sehr zeitintensiv und die Evaluierung der Krümmung durch den Züchterblick erfordert sehr geschultes Personal und ist dennoch, wie oben beschrieben, durchaus fehleranfällig. Dies macht sich auch darin bemerkbar, dass die Krümmung mit dem Auge lediglich auf einer ordinalen Skala bestimmt werden kann (z.B. gerade, leicht gekrümmt, mittel stark gekrümmt, stark gekrümmt), obwohl diese mathematisch gesehen kontinuierlicher Natur ist. Des Weiteren wachsen viele der länglichen Gemüsearten an Sträuchern und besitzen somit einen Blütenansatz, der die Frucht mit dem Rest der Pflanze verbindet. Ob vorzugsweise mit oder ohne diesen Blütenansatz geerntet wird, kann einerseits von der Kulturart oder von der weiteren Nutzung abhängen. Werden beispielsweise Bohnenhülsen frisch weitervermarktet, wird der Blütenansatz vorzugsweise nicht entfernt, da die Hülse so

länger frisch bleibt. Wird sie hingegen zu Konserven weiterverarbeitet, sollen die Hülsen möglichst ohne Blütenansatz geerntet werden, um weitere Verarbeitungsschritte zu vermeiden. Dabei spielt es weiterhin eine Rolle, ob eine Kulturart beziehungsweise eine Sorte einer Kulturart überhaupt maschinell ohne Blütenansatz geerntet werden kann. Für die Züchtung ist es somit interessant zu wissen, ob die geerntete Frucht einen Blütenansatz besitzt oder nicht².

Im Zuge dieser Arbeit soll am Beispiel der Buschbohne untersucht werden, ob mit Verfahren des maschinellen Sehens die Länge, das Kaliber, sowie die Krümmung objektiv und mit erhöhtem Durchsatz im Vergleich zu traditionellen Bewertungsmethoden berechnet und auf einer kontinuierlichen Skala dargestellt werden können. Dabei sollen vergleichbare Ergebnissen wie bei händisch durchgeführte Referenzmessungen bzw. Experteneinschätzungen auf der Grundlage visueller Bonituren erzielt werden, was eine zuverlässige Identifizierung und Lokalisierung der Blütenansätze und Spitzen der Früchte voraussetzt, um diese Strukturen von den weiteren Analysen auszuschließen.

1.4.2 Nicht-invasive Analyse der Samenanlage von Leguminosen

Wie bereits umrissen, ist ein wichtiges Qualitätsmerkmal bei Leguminosen, dass die Hülsen vollständig entwickelte Samenanlagen besitzen, wobei die Bewertung vor allem bei grünreifen Hülsen schwer durchzuführen ist. Die genaueste Methode wäre sicherlich, die Hülsen zu öffnen und die darin enthaltenen Samen zu zählen, beziehungsweise zu schauen, ob es Samenanlagen in der Hülse gibt, welche keine Samen ausbilden. Diese Methode ist jedoch nochmal deutlich zeitintensiver wie eine haptische Evaluierung durch Abtasten der Hülse. Hier steht demnach eine genaue, aber sehr zeitintensive und destruktive einer weniger genauen, immer noch relativ zeitintensiven, aber nicht destruktiven Methode gegenüber. Beide Methoden bieten somit gravierende Nachteile für die moderne Gemüsezüchtung, da jede Hülse einzeln begutachtet werden muss.

Im Zuge dieser Arbeit soll somit eine alternativer Lösungsansatz aus dem Bereich des maschinellen Sehens entwickelt und getestet werden, um die Samenanlage von Leguminosen nicht-invasiv und mit erhöhtem Durchsatz bei hoher Genauigkeit zu untersuchen.

1.4.3 Chromatische Analyse von RGB-Bilddaten zur Evaluierung von Sorteneigenschaften sowie Stresssymptomen in der Gemüsezüchtung

Sorteneigenschaften und die Resilienz gegenüber biotischen und abiotischen Stressoren stellen wichtige Qualitätskriterien für neue Gemüsesorten dar und werden oftmals auf Basis chromatischer Ausprägungen bewertet. Die Bewertung von Farbeigenschaften (auch im Kontext der Resilienzabschätzung) erfolgt traditionell visuell in unterschiedlichsten Umgebungen vom Feld bis hin zu geschlossenen Räumen mit künstlicher Beleuchtung. Die Subjektivität des Begutachtens, welche gerade bei der Wahrnehmung von Farben im Vergleich zu Formparametern noch stärkeren Einfluss hat, da sie im stärkeren Maße von den Umgebungsbedingungen abhängt, ist zuvor schon im Detail als Hauptproblem visueller Bewertungen erläutert worden, weshalb weltweit nach objektiveren Alternativen geforscht wird. Neben kostenintensiveren Multispektral- und Thermal-Bildgebungsmethoden (Chaerle et al., 2004; Cheshkova, 2022; Oerke et al., 2016; Raza et al., 2015; Stoll et al., 2008; Zhu et al., 2017), wird auch immer wieder der Einsatz von RGB-Methoden als kostengünstige Alternative getestet. Letztere haben sich dabei in der Vergangenheit vor allem für einfache Evaluierungen (einzelne Blätter, klare Strukturen, kontrollierte Bedingungen) als nützliches Werkzeug erwiesen (Han & Cointault, 2013; K. Sing & Chetia, 2017; Khadabadi et al., 2016; Mutka & Bart, 2015; Vinoth Kumar & Jayasankar, 2019)

² Latauschke, 2021 - Mittelfeine Buschbohnen bei sehr guten Bedingungen mit Spitzenerträgen
https://www.gartenbau.sachsen.de/download/Buschbohnen_mittelfein_2021.pdf (20.02.2024)

In dieser Arbeit soll das Potenzial günstiger RGB-Methoden in Kombination mit neuen Erkenntnissen aus dem Bereich des maschinellen Sehens für komplexere Anwendungsfälle evaluiert werden. Hierzu werden insgesamt 3 praxisnahe Beispiele bearbeitet:

1. Farbklassifikation und Helligkeitsanalyse unter kontrollierten Bedingungen zur Sortenbeurteilung

Die Klassifizierung der Farbe sowie die Beurteilung der Helligkeit einer Frucht sind wichtige Aufgaben in der Gemüsezücht. Bei manchen Früchten kann es vorkommen, dass sie eine Mischung an Farbausprägungen besitzen, welche fließend ineinander übergehen. Diese zu identifizieren und einzuordnen stellt eine besondere Herausforderung für Methoden des maschinellen Sehens da und soll im Zuge dieser Studie untersucht werden.

2. Bewertung komplexer biotischer Stresssymptome unter kontrollierten Bedingungen

Durch die Bewertung teils minimaler Stresssymptome an feingliedrigen Blättern sollen die technischen Limitierungen eines 2D RGB-Verfahren in Kombination mit neuesten Methoden aus dem Bereich des maschinellen Sehens näher untersucht werden.

3. Bewertung chromatischer Pflanzeigenschaften unter Feldbedingungen

Natürlich fluktuierenden Lichtbedingungen im Feld stellen für die Farbauswertung in RGB-Bildern eine große Herausforderung dar. Aus diesem Grund werden im Zuge dieser Arbeit verschiedene Farbkalibrierungsmethoden zur exakten Darstellung von Farben in Bildern verglichen und die vielversprechendste Methode auf Realbilder, welche unter Feldbedingungen aufgenommen wurden, angewandt.

1.4.4 Automatisierte und nicht-invasive hochdurchsatz-Ertragsabschätzung bei kopfbildendem Blattgemüse

Unter kopfbildendem Blattgemüse versteht man zunächst Arten oder Unterarten, welche vor der Blüte und der Bildung des Hauptsprosses einen Kopf ausbilden, welcher als Erntegut einen ernährungsphysiologischen Wert besitzt (Meier, 2018). Hier sind beispielsweise Kopfkohl (*Brassica oleracea* convar. *capitata* L. var. *alba* [Weißkohl] / *rubra* [Rotkohl] / *sabauda* [Wirsing]), Chinakohl (*Brassica rapa* subsp. *pekinensis*), Kopf- und Eissalat (*Lactuca sativa* L. var. *capitata*) wie auch Endivie (*Cichorium endivia* L.) als wichtige Kulturarten zu nennen (Meier, 2018). In der Zucht von kopfbildenden Blattgemüse-Arten ist neben Qualitätsmerkmalen wie Geschmack, Nährwert, Lagereigenschaften, sowie der Uniformität der Köpfe innerhalb einer Sorte, auch der Ertrag weiterhin eines der wichtigsten Qualitätsmerkmale (Fang et al., 2004). Dieser ergibt sich aus dem Gewicht des geernteten Kopfes, welches wiederum von dessen Volumen, der Dichte, der Form, sowie der Größe des Strunks innerhalb des Kopfes beeinflusst wird. Die manuelle Evaluierung des Gewichtes, beispielsweise mit einer Federwaage stellt in der Gemüsezücht eine enorme Limitierung dar, da die Köpfe geerntet werden müssen. Durch das destruktive Verfahren können keine Zeitreihenmessungen einzelner Köpfe durchgeführt werden. Außerdem werden nur einzelne Datenpunkte genutzt, um verallgemeinerte Rückschlüsse auf die gesamte Sorte zu ziehen, was Ungenauigkeiten auf Grund geringer Stichprobengröße inkludiert.

Im Zuge dieser Arbeit soll untersucht werden, ob moderne Methoden des maschinellen Sehens nicht-invasive Hochdurchsatz-Ertragsabschätzung bei kopfbildendem Blattgemüse unter Feldbedingungen ermöglichen und somit einen gravierenden Flaschenhals in der Gemüsezücht öffnen können.

2 Grundlagen des maschinellen Sehens (MS)

2.1 Elektromagnetisches Spektrum

Das elektromagnetische Spektrum kategorisiert sämtliche in der Natur vorkommende elektromagnetische Strahlung nach der Wellenlänge (Sloney, 2016). Wie in Abbildung 2 dargestellt reicht dieses Spektrum von der kosmischen Höhenstrahlung mit einer Wellenlänge von ca. 1 Femtometer (10^{-15}m) bis hin zu niederfrequenten Wechselströmen mit einer Wellenlänge von mehreren Tausend km (10^7m). Das für den Menschen sichtbare Licht befindet sich in dem Wellenlängenbereich zwischen 400 und 700 Nanometer ($4\text{e}^{-7}\text{-}7\text{e}^{-7}\text{m}$) und wird von verschiedensten Lichtquellen wie der Sonne, Feuer, oder elektrischen Lampen emittiert (Sloney, 2016). Die emittierten elektromagnetischen Wellen werden oft auch als Lichtteilchen bzw. Photonen interpretiert (Sloney, 2016). Treffen diese emittierten Photonen auf physikalische Objekte, wie beispielsweise Pflanzen, werden diese je nach den Eigenschaften des Objekts entweder absorbiert, gestreut, transmittiert oder reflektiert (Cheshkova, 2022). Photonen können beispielsweise mit einem lichtsensitiven Halbleiter (Fotodiode) gemessen werden (A. Rahman, 2014). Die Photonen regen dabei den Elektronenfluss im Halbleitermaterial an, sodass der gemessene elektrische Strom direkt mit der Lichtintensität des einfallenden Lichts auf die Fotodiode korreliert (A. Rahman, 2014). Ein RGB-Kamerasensor ist so konzipiert, dass er den Photoneneinfall im sichtbaren Wellenlängenbereich misst (Maschal et al., 2010).

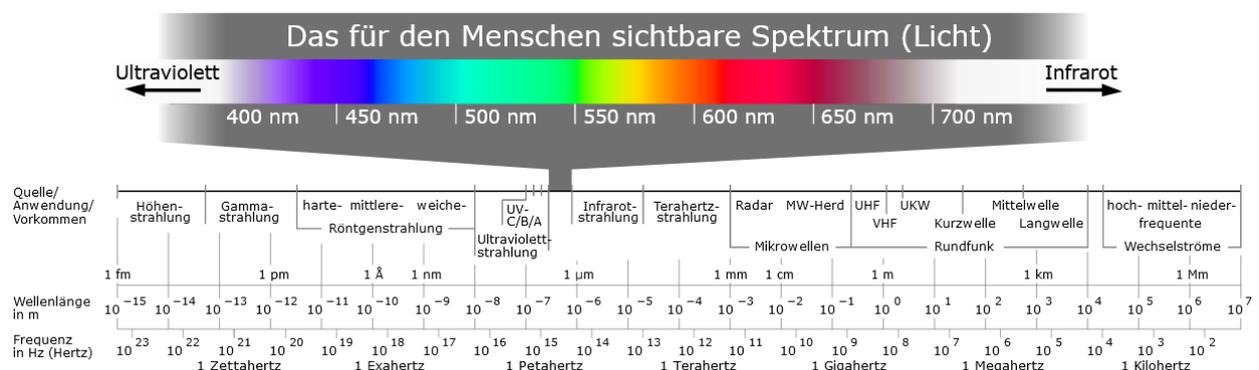


Abbildung 2: Das elektromagnetische Spektrum. Quelle: https://upload.wikimedia.org/wikipedia/commons/6/62/Electromagnetic_spectrum_-_de_c.svg (30.12.2023)

2.2 Farbräume

Die in dieser Arbeit verwendeten Kamera-Sensoren liefern Daten im RGB-Farbraum. Insbesondere im Bereich der Bildverarbeitung ist der RGB-Farbraum häufig jedoch nicht geeignet, da die Intensitäten der einzelnen Kanäle untereinander korreliert sind und die Kanäle nicht als unabhängige Merkmale genutzt werden können. Daher kommen hier in der Regel andere Farbräume zum Einsatz, wie der HSV oder LAB-Farbraum, welche nachfolgend neben dem RGB-Farbraum genauer betrachtet werden sollen.

2.2.1 RGB-Farbraum

Das Modell des Rot-Grün-Blau (RGB) Farbraums wurde speziell für die Darstellung von Bildern auf elektronischen Monitoren entwickelt, welche unterschiedliche Farben durch anteiliges Mixen der monochromen Spektren der 3 Primärfarben Rot, Grün und Blau erzeugen, wie dies auch bei dem Mischen von Malfarben und im Farbdruck funktioniert (Hastings &

Rubin, 2012). Auf Grund technischer Limitierungen wird dabei jedoch ein geringeres Gamut im Vergleich zum gesamten wahrnehmbaren Gamut aller sichtbaren Farben dargestellt (Hastings & Rubin, 2012). Das Gamut beschreibt die Spanne an dargestellten Farben (Smith, 1978). Der RGB-Farbraum kann grundsätzlich als Würfel in einem 3-dimensionalem kartesischen Koordinatensystem dargestellt werden, wobei jede Achse den Intensitätswert einer der 3 Primärfarben darstellt (Abbildung 3a). Der RGB-Farbraum gehört damit zu den additiven Primärfarbräumen und wird hier als Tupel in der Form (R = 0-255, G = 0-255, B = 0-255) dargestellt, wobei jedes Element die Intensität der jeweiligen Primärfarbe darstellt, woraus sich schließlich die darzustellende Farbe ergibt (Da Rugna et al., 2003; Hastings & Rubin, 2012). Obwohl sich das RGB-Modell für visuelle Technologien bewährt hat, ist es für andere Anwendungen eher ungeeignet, da es nicht der menschlichen Wahrnehmung entspricht und nicht intuitiv zu interpretieren ist (Hastings & Rubin, 2012).

2.2.2 HSV-Farbraum

Der Hue-Saturation-Value (HSV) Farbraum ist eine Abwandlung des RGB-Farbraums und unterliegt damit den gleichen Einschränkungen wie Geräteabhängigkeit und visueller Uneinheitlichkeit (Hastings & Rubin, 2012). Der HSV-Farbraum ist jedoch deutlich intuitiver, da er die Luminanz von den chromatischen Komponenten trennt (Hastings & Rubin, 2012; Hidayatullah & Zuhdi, 2015). Die Luminanz wird dabei vom Value (Helligkeitwert) beschrieben, die Chromatik (altgriechisch χρώμα (chróma)=,Farbe') teilt sich auf Hue (Farbtonwert) und Saturation (Sättigungswert) auf. Der Farbtonwert wird in Gradzahlen zwischen 0° und 360° angegeben und beschreibt die Grundfarben Rot (0°/360), Gelb (60°), Grün (120°), Cyan (180°), Blau (240°) und Magenta (300°), sowie die Kombinationen dazwischen. Der Sättigungswert gibt an, wie sehr die reine Farbe mit weißem Licht verdünnt ist. Der Helligkeitwert beschreibt die Luminanz der Farbe von dunkel (schwarz) zu hell (Vollton) (Hastings & Rubin, 2012; Hidayatullah & Zuhdi, 2015)(Abbildung 3b). Im Rahmen dieser Arbeit wird eine Farbe im HSV-Farbraum als Tupel in der Form (H = 0-360, S = 0-255, V = 0-255) dargestellt, wobei H den Farbtonwert, S den Sättigungswert und V den Helligkeitwert wiedergibt. Die Transformation vom RGB- in den HSV-Farbraum berechnet sich wie folgt:

$$\begin{aligned}
 V &\leftarrow \max(R, G, B) \\
 S &\leftarrow \text{if } V \\
 &\neq 0: \frac{V - \min(R, G, B)}{V}; \text{ else: } 0 \\
 H &\leftarrow \text{if } V \\
 &= R: \frac{60(G - B)}{V - \min(R, G, B)}; \text{ if } V \\
 &= G: 120 + \frac{60(B - R)}{V - \min(R, G, B)}; \quad [1] \\
 &\text{if } V \\
 &= B: 240 \\
 &+ \frac{60(R - G)}{V - \min(R, G, B)}; \text{ if } R = B \\
 &= G: 0
 \end{aligned}$$

2.2.3 CIE-LAB-Farbraum

Eine international anerkannte, standardisierte und reproduzierbare Darstellung von Farben stellt der von der *Internationalen Beleuchtungskommission* (CIE = *Commission Internationale de l'Eclairage*) vorgestellte LAB-Farbraum dar. Dieser leitet sich in seinen Ursprüngen von den Spektren des sichtbaren Lichts ab, gehört zu den wahrnehmbaren Farbräumen und umfasst somit das gesamte Gamut aller sichtbaren (Da Rugna et al., 2003; Hastings & Rubin, 2012). Eine 3-dimensionale Repräsentation beschreibt eine vertikale Luminanzachse ,L' die von

Schwarz zu Weiß reicht, sowie 2 horizontale Achsen ‚A‘ und ‚B‘ die von ‚Rot bis Grün‘ und ‚Gelb bis Blau‘ reichen. Per Definition sind Rot und Gelb dabei auf der positiven Seite der jeweiligen Achse und Grün und Blau auf der negativen Seite (Hastings & Rubin, 2012). Schwarz liegt im Nullpunkt des kartesischen Koordinatensystems und reines Weiß liegt im positiven Limit der vertikalen Achse. Das geplottete Gamut aller wahrnehmbaren Farben formt dabei einen bauchförmigen Körper, der die reduzierte Farbwahrnehmung bei sehr hohen und sehr niedrigen Luminanzwerten (Helligkeitswerten) wiedergibt (Abbildung 3c). Das LAB-Farbraum Modell entspricht der menschlichen Wahrnehmung, sodass gemessene und wahrgenommene Abstände zwischen Farben im Vergleich zu anderen Farbräumen gut übereinstimmen (Hastings & Rubin, 2012). In dieser Arbeit wird eine Farbe im LAB-Farbraum als transformierter Tupel in der Form (L = 0-255, A = 0-255, B = 0-255) dargestellt, wobei L den Luminanzwert bzw. Helligkeitswert beschreibt und A und B die Werte auf der entsprechenden Achse. Die Transformation vom RGB- in den LAB-Farbraum berechnet sich wie folgt:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0,412453 & 0,357580 & 0,180423 \\ 0,212671 & 0,715160 & 0,072169 \\ 0,019334 & 0,119193 & 0,950227 \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

$$X \leftarrow \frac{X}{0,950456}$$

$$Z \leftarrow \frac{Z}{1,088754}$$

$$L \leftarrow \text{if } Y > 0,008856: \left(116 * Y^{\frac{1}{3}} - 16\right) * \frac{255}{100}; \text{ else: } (903,3 * Y) * \frac{255}{100} \quad [2]$$

$$A \leftarrow 500 * (f(X) - f(Y)) + 256$$

$$B \leftarrow 200 * (f(Y) - f(Z)) + 256$$

wobei

$$f(T) \leftarrow \text{if } t > 0,008856: T^{\frac{1}{3}}; \text{ else: } 7,787T + \frac{16}{116}$$

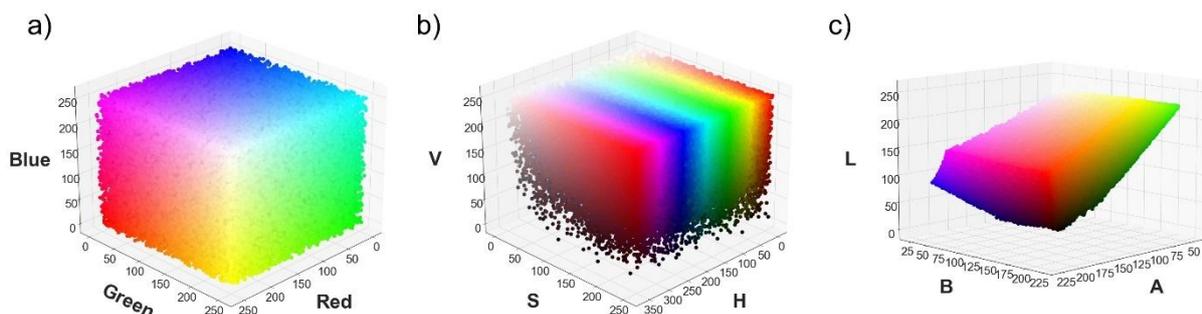


Abbildung 3: Skalierte 3-Dimensionale Repräsentationen des a) RGB, b) HSV und c) CIE-LAB Farbraums.

2.3 Klassische Bildverarbeitung

Nachfolgend werden verschiedene traditionelle Verfahren der Bildverarbeitung erläutert, welche in dieser Arbeit zum Einsatz kamen. Diese Verfahren sind grundlegend für Verfahren des maschinellen Sehens und bilden die Basis für klassische maschinelle Lernverfahren. Ansätze des tiefen Lernens grenzen sich insofern von den anderen Verfahren ab, dass sie das Konzept des „End-to-End-Lernens“ verfolgen. Die Erklärungen der traditionellen Verfahren

werden im Folgenden aus Gründen der Verständlichkeit anhand von Graubildern verdeutlicht. Alle vorgestellten Methodiken lassen sich jedoch auf die einzelnen Kanäle eines Farbbilds anwenden (welche auch Graubilder darstellen).

2.3.1 Geometrische und chromatische Transformationen

Bildtransformationen lassen sich in geometrische und chromatische Transformationen einteilen, wobei erstere die Pixelpositionsdaten in einem kartesischen Koordinatensystem und letztere die Pixelwerte transformieren (Mikołajczyk & Grochowski, 2018; Shorten & Khoshgoftaar, 2019).

Skalierung: Bei der Skalierung werden die Größe des Bildes, also die Ausmaße der Höhe und Breite, durch Interpolation verändert (Bradski, 2000). Der Bildausschnitt bleibt jedoch unverändert. Das Bild kleiner zu machen kann ein nötiger Schritt sein, um ein Bild beispielsweise mittels neuronaler Netzwerke weiter prozessieren zu können (K. He et al., 2017) (Abbildung 4b).

Zoom: Beim Zoom schneidet man das Bild auf einen bestimmten Bildausschnitt zurecht indem man die Koordinaten der Bildmatrix angibt, die man behalten möchte (Harris et al., 2020). Dies kann Hilfreich sein um unnötige Informationen von Beginn an von der Bildanalyse auszuschließen (Abbildung 4c).

Rotation: Bei der Rotation wird jede Pixelposition $\begin{bmatrix} x_i \\ y_i \end{bmatrix}$ mit der Transformationsmatrix $\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$ multipliziert, sodass das Bild mit dem Winkel θ um das Bildzentrum gedreht wird (Bradski, 2000) (Abbildung 4d).

Spiegelung: Horizontale beziehungsweise vertikale Spiegelung des Bildes erreicht man durch spaltenweise oder reihenweise Umkehrung der Elementen Reihenfolge der Bildmatrix (Harris et al., 2020) (Abbildung 4e).

Translation: Bei der Translation wird jede Pixelposition $\begin{bmatrix} x_i \\ y_i \end{bmatrix}$ mit der Transformationsmatrix $\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \end{bmatrix}$ multipliziert, wobei t_x die horizontale Verschiebung und t_y die vertikale Verschiebung angibt (Abbildung 4f).

Histogramm-Entzerrung: Das Histogramm eines Bildes gibt die Häufigkeitsverteilung der Pixelwerte an und gibt Aufschluss darüber, wie hell oder dunkel ein Bild ist oder wie groß der Kontrast ist. Je breiter sich das Histogramm darstellt, umso größer ist der Kontrast im Bild. Bei einem Binärbild, welches lediglich aus weißen und schwarzen Pixeln besteht ist der Kontrast maximal. Für eine hohe Bildqualität wird jedoch ein ausgeglichenes Histogramm mit einer breiten Pixelwertverteilung und ausgeglichenen Pixelwerthäufigkeiten angestrebt (Tarika et al., 2014). Um dies zu erreichen, gibt es unterschiedliche Verfahren zur Histogramm-Entzerrung. Neben der globalen Histogramm-Entzerrung gibt es auch lokale Verfahren wie beispielsweise die kontrastlimitierte lokale Histogramm-Entzerrung (Contrast Limited Adaptive Histogram Equalization = CLAHE), welche weniger Bildrauschen im Vergleich zum globalen Ansatz erzeugt (Tarika et al., 2014) (Abbildung 4g).

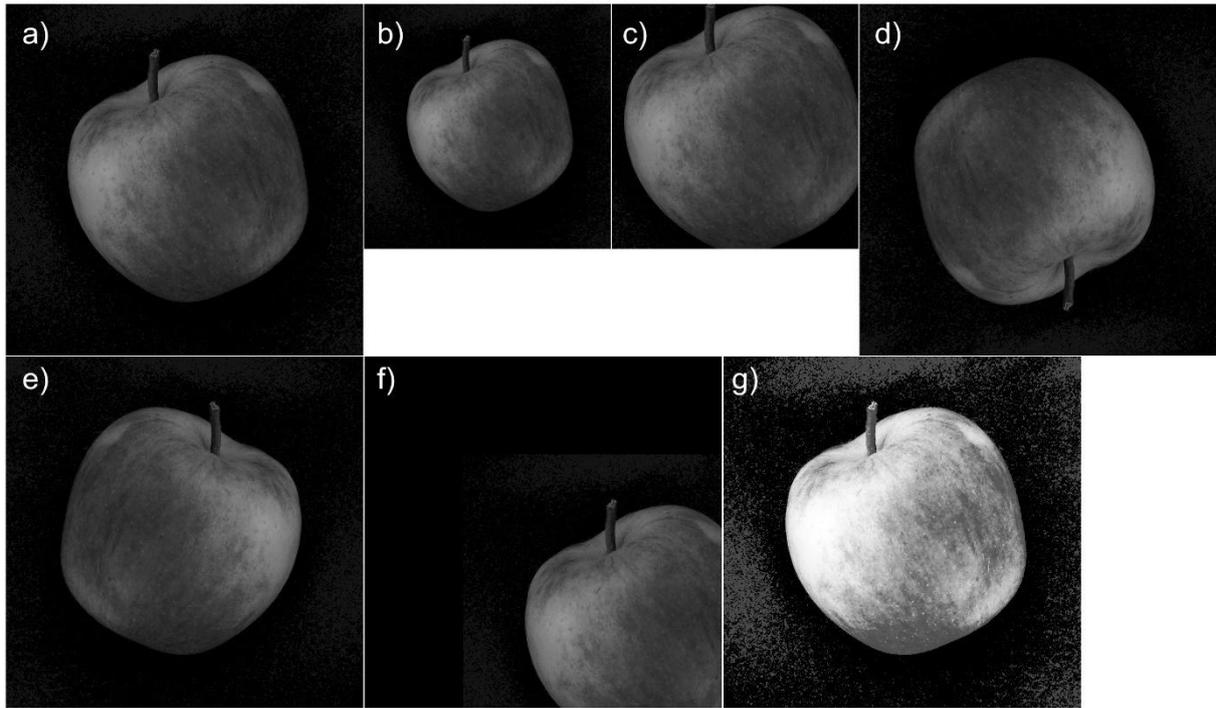


Abbildung 4: Beispiele für klassische Bildtransformationen. a) Originales Graubild; b) Skalierung; c) Zoom; d) Rotation; e) Spiegelung; f) Translation; g) Histogramm Entzerrung

2.3.2 Bildfilter

Ein Bildfilter besteht aus einer sogenannten Faltungsmatrix, üblicherweise mit den Ausmaßen von 3x3 oder 5x5 Pixeln. Die Faltungsmatrix wird nun so über jedes Pixel des Bildes gelegt, dass das zu betrachtende Pixel (hier ‚Initialpixel‘ genannt) im Zentrum der Faltungsmatrix liegt. Das Initialpixel und (im Falle einer 3x3 Faltungsmatrix) die 8 umliegenden Pixel werden nun mit den entsprechenden Elementen der Faltungsmatrix multipliziert und die Ergebnisse aufsummiert, wodurch sich der neue Wert für das Initialpixel ergibt (Abbildung 5a) . Je nach Aufbau des Filters ist dieser in der Lage bestimmte Operationen wie Weichzeichnen, Schärfen, oder Kantenerkennung an Bildern durchzuführen (Abbildung 5b).

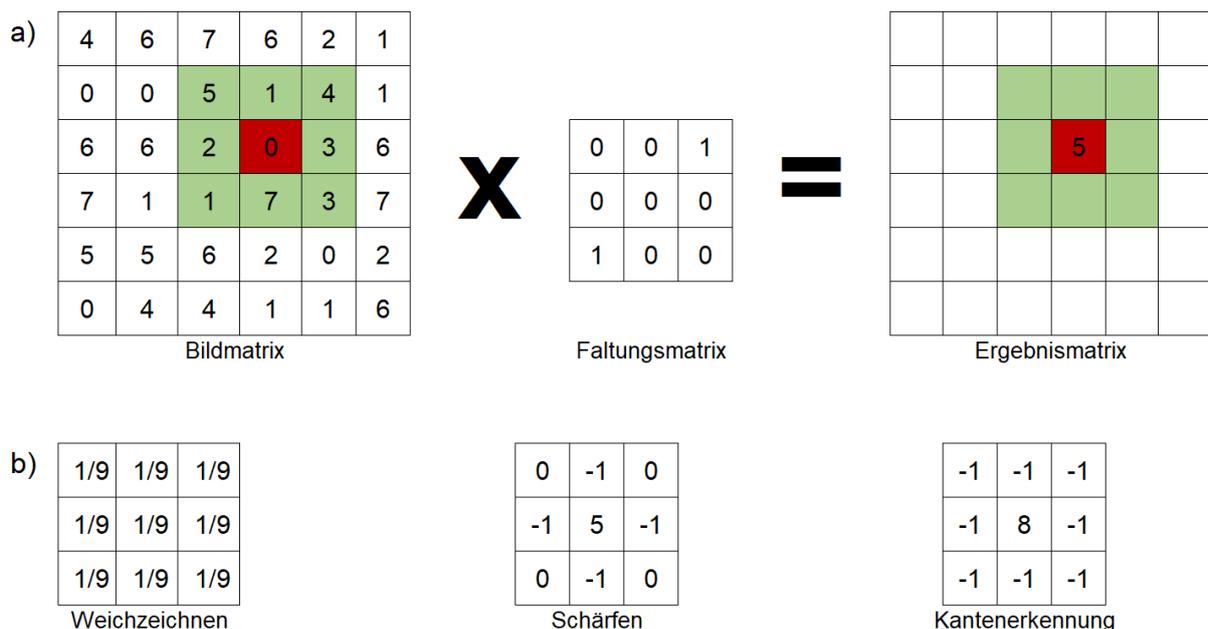


Abbildung 5: a) Schema zur Funktionsweise einer 3x3 Faltungsmatrix. Die Faltungsmatrix wird mit dem Zentrum über das Initialpixel (Rot) gelegt und jeweils die entsprechenden Elemente der Bildmatrix (grün) und der

Faltungsmatrix miteinander multipliziert und anschließend aufsummiert, woraus sich der neue Wert für das Initialpixel ergibt. Abbildung inspiriert aus dem GIMP Benutzerhandbuch (<https://www.gimp-handbuch.de/>). b) Beispiele für 3x3 Faltungsmatrizen für 3 einfache Bildfilter

2.3.3 Texturanalysen

Die räumliche Variation der Pixelwerte sowie die Helligkeitsintensitätsverteilung geben Struktur- wie auch Texturinformationen eines Graubildes, dessen Komplexität mittels einer Vielzahl an Methodiken reduziert werden kann, um gewünschte Informationen aus einem Bild zu ziehen (Armi & Fekri Ershad, 2019; Wiley & Lucas, 2018). Objekte in einem Bild stellen jeweils zusammenhängende Strukturen dar, die sich auf Grund ihrer Kontur, Farbe oder Textur voneinander unterscheiden lassen (Wiley & Lucas, 2018).

Im Bereich des maschinellen Sehens beschreibt Textur die Oberflächenbeschaffenheiten von Objekten wie Glätte, Rauheit und Regelmäßigkeit (Armi & Fekri Ershad, 2019). Diese kann zum Beispiel durch statistische Analyse der Pixelwerte beschrieben werden. So können die Werthäufigkeiten ohne räumliche Zuordnung als Intensitätshistogramm dargestellt werden, um beispielsweise helle Objekte von dunklen zu unterscheiden oder statistische Größen wie den Variationskoeffizienten, die Entropie, die Asymmetrie (Skewness) oder die Kurtosis aus dem Histogramm abzuleiten (Armi & Fekri Ershad, 2019). Der Variationskoeffizient CoV gibt die relative Streuung der Daten um den Mittelwert an und berechnet sich wie folgt aus den Pixelwerten:

$$CoV = \frac{\sigma}{\bar{x}} \quad [3]$$

Dabei gibt σ die Standardabweichung und \bar{x} den Mittelwert an (Brown, 1998). In der Informationstheorie nach Shannon gibt die Entropie den Grad der Unsicherheit einer Verteilung an und berechnet sich folgendermaßen:

$$Entropie = - \sum_{n=1}^{N-1} px_n * \log_2(px_n) \quad [4]$$

Dabei gibt N die Anzahl an Werten der Verteilung (256 bei einem 8-Bit Bild) und px_n die Häufigkeit/Wahrscheinlichkeit mit der ein Wert x_n auftritt (Armi & Fekri Ershad, 2019; Katiyar & Arun, 2014; Pedregosa et al., 2012). Die Asymmetrie beschreibt die Abweichung des Histogramms von der Normalverteilung, welche eine Asymmetrie von 0 besitzt. Eine Asymmetrie von >0 beschreibt eine Verschiebung nach rechts und <0 eine Verschiebung des Histogramms nach links. Die Asymmetrie berechnet sich wie folgt (Dorić et al., 2009):

$$Asymmetrie = \frac{1}{N} \sum_{n=1}^N \left(\frac{x_n - \bar{x}}{\sigma} \right)^3 \quad [5]$$

Die Kurtosis nach *Fisher* beschreibt die Stauchung, bzw. Streckung einer Verteilung (eines Histogramms) und ist 0 bei Normalverteilung, negativ bei einer Streckung und positiv bei einer Stauchung. Die Kurtosis berechnet sich wie folgt (Dorić et al., 2009):

$$Kurtosis = \frac{1}{N} \sum_{n=1}^N \left(\frac{x_n - \bar{x}}{\sigma} \right)^4 - 3 \quad [6]$$

2.3.4 Schwellenwertsegmentierung

Die Umwandlung eines Graubildes in ein Binärbild (1-Bit) stellt eine weitere Möglichkeit dar den Informationsgehalt eines Bildes auf die wesentlichen Elemente zu reduzieren. Eine Möglichkeit ein Graubild in ein Binärbild umzuwandeln, besteht darin einen globalen Schwellenwert festzulegen, sodass jeder Pixel mit einem Wert unterhalb des Schwellenwertes den Wert 0 (Schwarz) zugeteilt bekommt oder 1 (Weiß), wenn dieser darüber liegt. Ein solches Verfahren bietet sich beispielsweise an, wenn ein dunkles Objekt vor einem hellen Hintergrund segmentiert werden soll, ist jedoch empfindlich gegenüber fluktuierenden Lichtbedingungen innerhalb des Bildes (Roy et al., 2014). Um dieses Problem zu umgehen können pixelweise Schwellenwerte berechnet werden (Roy et al., 2014). Dazu wird um jeden Pixel ein lokales Fenster gelegt, dessen Größe variabel ist (z.B. 35 * 35 Pixel). Der lokale Schwellenwert ergibt sich nun aus dem gewichteten Mittel der Pixelwerte minus einer vorher festgelegten Abweichung (van der Walt et al., 2014). Somit können Schattierungseffekte im Bild ausgeglichen werden.

2.3.5 Morphologische Operatoren

Morphologische Operatoren stellen nützliche Werkzeuge dar, um Binärbilder, welche auch als Bildmasken bezeichnet werden, weiter aufzubereiten. Die grundlegendsten Operatoren sind die Dilatation, sowie die Erosion. Eine Dilatation lässt den Vordergrund, also die Anzahl an weißen Pixeln, wachsen, während der Hintergrund, also die Anzahl der schwarzen Pixel, schrumpft. Die Erosion stellt den gegenteiligen Prozess dar, trägt also den Rand des Vordergrunds ab (Bradski, 2000; Srisha & Khan, 2013). Die Funktionsweise dieser Methoden beruht dabei auf sogenannten strukturierenden Elementen und logischen Operationen (Srisha & Khan, 2013). Ein strukturierendes Element stellt dabei eine kleine Binärmatrix (meistens 3x3 oder 5x5) dar, die wie eine Faltungsmatrix über jeden Pixel des Binärbilds gelegt wird. Im Gegensatz zu einer Filteroperation wird dann jedoch ein logischer Vergleich zwischen den entsprechenden Elementen des strukturierenden Elements und dem Bild angestellt (Srisha & Khan, 2013). Für die Dilatation bedeutet dies, dass das Initialpixel den Wert 1 erhält, wenn mindestens eine Übereinstimmung bei jenem Vergleich auftritt. Für die Erosion ist dies nur der Fall, wenn das strukturierende Element mit dem Bildausschnitt übereinstimmt. Andernfalls bekommt das Initialpixel den Wert 0 zugeteilt. Die Struktur sowie die Größe des strukturierenden Elements bestimmen dabei maßgeblich das Ausmaß der morphologischen Operationen (Bradski, 2000; Srisha & Khan, 2013). Eine Erosion gefolgt von einer Dilatation nennt man Öffnung, welche zur Rauschreduzierung in Binärbildern genutzt werden kann. Eine Dilatation gefolgt von einer Erosion schließt Löcher im Vordergrund der Bildmaske (Bradski, 2000; Srisha & Khan, 2013; van der Walt et al., 2014).

Ein zusammenhängender Verbund von Vordergrundpixeln in einem Binärbild wird auch als verbundene Komponente bezeichnet (Abbildung 6a), aus denen sich weiterhin die Kontur oder das Skelett ableiten (L. He et al., 2017; Suzuki & Abe, 1985; T. Y. Zhang & Suen, 1984). Die Kontur bildet den äußeren Rand der verbundenen Komponente ab, und wird nach Suzuki & Abe (1985) ermittelt (Bradski, 2000) (Abbildung 6b). Die Skeletonisierung reduziert die

verbundene Komponente in eine 1 Pixel breite Repräsentation, wofür unterschiedlichste algorithmische Ansätze genutzt werden können (van der Walt et al., 2014). Nach Zhang & Suen (1984) werden in einem iterativen Prozess solange, unter Beibehaltung der Konnektivität der Komponente, die Konturpixel entfernt, bis eine 1 Pixel breite Repräsentation der Komponente übrigbleibt (Abbildung 6c). Eine andere Möglichkeit besteht darin, die mediale Achse der Komponente zu bestimmen, dessen Pixel sich dadurch auszeichnen, dass sie mehr als einen nächstgelegenen Punkt zur Kontur haben (van der Walt et al., 2014).

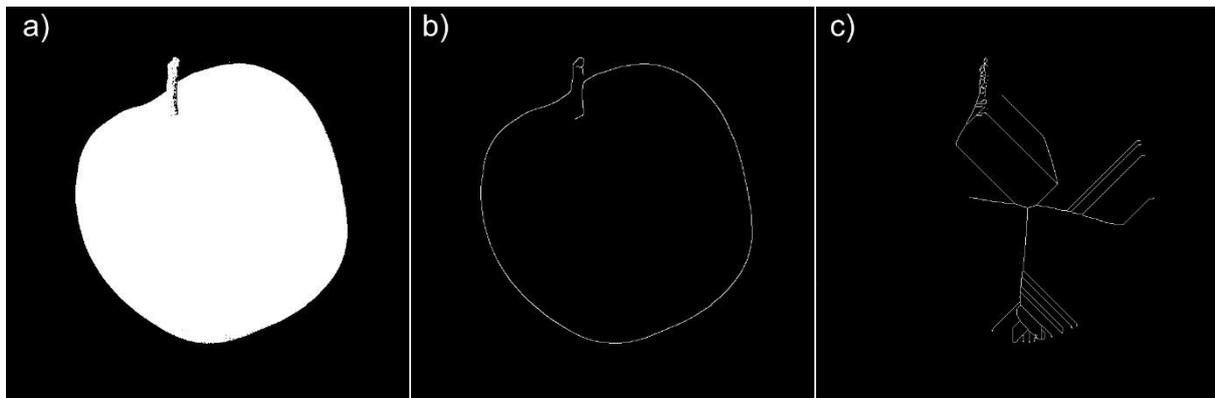


Abbildung 6: a) Binärbild eines Apfels; b) Kontur eines Apfels; c) Skelett eines Apfels

2.3.6 Formdeskriptoren

Aus der verbundenen Komponente, der Kontur und dem Skelett lassen sich folgende Parameter ableiten: Zunächst kann man jeweils die Anzahl der Vordergrundpixel bestimmen. Aus den Ausmaßen der verbundenen Komponente lässt sich weiterhin ein rechtwinkliges umschließendes Rechteck, ein gedrehtes Rechteck, ein Dreieck, oder eine Ellipse mit minimal umschließender Fläche, eine konvexe Hülle oder ein Kreis ableiten. Von diesen Formen können die Ausmaße und Fläche berechnet werden. Die Fläche der verbundenen Elemente lässt sich jedoch auch direkt durch die Kontur und dem *Satz von Green* berechnen (Bradski, 2000). Momente m eines Bildes mit den Dimensionen $M * N$ werden wie folgt berechnet:

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p * y^q * I(x, y) \quad [7]$$

p und q geben den Grad des Moments, x und y die Pixelkoordinate im Bild und I die Intensität des Pixels (x,y) an. Daraus ergibt sich beispielsweise der Schwerpunkt der Kontur einer verbundenen Komponente wie folgt:

$$\text{Schwerpunkt} = [\bar{x}, \bar{y}] = \left[\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right] \quad [8]$$

Bindet man den Schwerpunkt in die obige Formel [7] ein und normalisiert die Gleichung, erhält man zentrale Momente die invariant gegenüber Translation, Spiegelung, Rotation und Skalierung sind (Gonzales & Woods, 2018; Z. Huang & Leng, 2010; Ming-Kuei Hu, 1962):

$$\eta_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \frac{(x - \bar{x})^p * (y - \bar{y})^q * I(x, y)}{I(x, y)^{\frac{p+q}{2}+1}} \quad [9]$$

Die Länge oder Krümmung einer Kontur oder eines Skeletts stellen weitere wichtige Parameter dar. Die Länge berechnet sich aus der Summe der Abstände zwischen allen benachbarten Pixeln. Für alle horizontal bzw. vertikal verbundenen Pixel beträgt der Abstand 1 Pixel, für alle diagonal verbundenen Pixel beträgt der Abstand nach dem *Satz des Pythagoras* ca. 1,41 Pixel (Bradski, 2000).

2.4 Ansätze des maschinellen Lernens

2.4.1 Theorie des statistischen Lernens

Die Theorie des statistischen Lernens bildet die Grundlage für maschinelles Lernen und beschreibt das Konzept, wie ein Computer aus empirischen Daten bestimmte Aufgaben „erlernen“ kann, indem er spezifischen Lernalgorithmen folgt (Luxburg & Schölkopf, 2011). Dabei wird grundsätzlich zwischen überwachten (englisch *supervised*) und unüberwachten (englisch *unsupervised*) Lernalgorithmen unterschieden (Evgeniou & Pontil, 2000; Luxburg & Schölkopf, 2011).

Ein überwachter Lernalgorithmus nutzt annotierte Trainingsdaten, um den funktionalen Zusammenhang zwischen den Eingangsdaten X und den Ausgangsdaten Y in einem iterativen Trainingsprozess zu ermitteln (Evgeniou & Pontil, 2000). Unter die überwachten Lernmethoden fallen klassischerweise Regressions- und Klassifikationsaufgaben, wobei die Ausgangsdaten bei einer Regression kardinal (metrisch) oder ordinal (nach Rang) und bei einer Klassifikation nominal (ohne Rang) skaliert sind (Sarstedt & Mooi, 2014; Sathya & Abraham, 2013; Uyanık & Güler, 2013). Ein Beispiel für eine Regressionsaufgabe wäre die Abhängigkeit des Ertrags eines Weizenfeldes (resultierende/abhängige Variable) in kg/ha von der Niederschlagsmenge, dem Düngereinsatz und der Bestandsdichte (erklärende/unabhängige Variablen). Ein Beispiel für eine Klassifikationsaufgabe wäre die Unterscheidung zwischen Bildern, die eine Weizenähre bzw. eine Gerstenähre darstellen. Hierbei stellt die Pixelmatrix des Bildes die Eingangsdaten und die entsprechende Klassifizierung in Weizen- oder Gerstenähren die Ausgangsdaten dar.

Nach der Theorie des statistischen Lernens wird zur Lösung von Regressions- und Klassifikationsaufgaben zunächst eine Verlustfunktion (englisch *loss function*) definiert, die den Fehler des Lernalgorithmus als empirisches Risiko angibt. Als nächstes wird eine verschachtelte Folge von Hypothesen H_m , $m = 1, \dots, M$ definiert, wobei jede Hypothese eine endliche Komplexität besitzt, welche jedoch größer ist als die Komplexität der vorausgegangenen Hypothesen in H_m . H_m könnte zum Beispiel eine Menge der Polynome vom Grad m sein, eine Menge von Splines mit m Knoten oder eine kompliziertere nichtlineare Parametrisierung. Zuletzt wird das empirische Risiko für jede Hypothese in H_m berechnet und schließlich jene Lösung mit dem besten Kompromiss zwischen empirischen Risiko und Komplexität ausgewählt. Diese Kombination nennt sich *strukturierte Risikominimierung* und erlaubt es dem Algorithmus besser zu generalisieren und *Overfitting* zu vermeiden, was eine Annäherung des empirischen Risikos gegen Null bedeuten würde (Evgeniou & Pontil, 2000; Luxburg & Schölkopf, 2011). *Overfitting* bezeichnet im Kontext des maschinellen Lernens das Phänomen, dass sich ein Modell zu eng an die Trainingsdaten anpasst und dadurch seine Fähigkeit verliert, auf neuen und unbekanntem Daten gut zu generalisieren.

Ein unüberwachter Lernalgorithmus hingegen untersucht die grundlegende Struktur eines Datensatzes ohne die Hilfe von annotierten Trainingsdaten. Die Ausgangsdaten Y ergeben sich somit aus Zusammenhängen innerhalb der Eingangsdaten X (Evgeniou & Pontil, 2000). Klassische Aufgaben für unüberwachte Lernalgorithmen sind beispielsweise das Clustern der Eingangsdaten X , mehrere Variablen aus X zusammenzufassen um die Dimensionen des Datensatzes zu reduzieren oder Beziehungen zwischen Variablen aus X festzustellen

2.4.2 Regression und Interpolation

Grundsätzlich werden für Regression und Interpolationen dieselben überwachten Lernalgorithmen genutzt, jedoch unterscheiden sich die beiden Ansätze in ihrer Zielsetzung.

Die Regression wird genutzt um den Zusammenhang zwischen einer abhängigen Variable und einer oder mehrere unabhängigen Variablen zu bestimmen. Wird die Beziehung zwischen einer abhängigen Variablen und eine unabhängige Variable untersucht spricht man von einer univariaten Regression, wobei die einfachste Form die lineare Regression darstellt. Weitere Formen der univariaten Regression sind zum Beispiel die polynomische, logistische oder Spline Regression (Maalouf, 2011; Madani et al., 2010). Unter einer univariaten Spline Interpolation versteht man eine Sequenz polynomischer Funktionen n -ten Grades, die über sogenannte Knoten linear miteinander verbunden sind und sich einerseits durch ihre Einfachheit, andererseits durch ihre hervorragenden Annäherungseigenschaften auszeichnen (Boor, 1986; Madani et al., 2010).-Wird der Zusammenhang einer abhängigen Variablen zu mehreren unabhängigen Variablen untersucht spricht man von einer multivariaten Regression. Ein Beispiel ist die Regression der partiellen kleinsten Quadrate (PLSR), welche ebenfalls zu den multivariaten linearen Regressionen gehört. Diese wurde entwickelt, um die abhängige Variable von den unabhängigen Variablen abzuleiten und deren zugrundeliegende Struktur zu beschreiben – besonders, wenn die Daten eine kleine Stichprobe haben, die unabhängigen Variablen eine hohe Co-Linearität aufweisen, oder Werte fehlen (Pirouz, 2012). Grundsätzlich weist sie Ähnlichkeiten zur Hauptkomponentenregression auf. So werden die unabhängigen Variablen zunächst zusammenfasst und somit die Anzahl der Eingangsvariablen reduziert. Hierzu werden die unabhängigen Variablen gewichtet miteinander addiert, wobei die quadrierte Summe aller Gewichte 1 ergibt (del Val & de Agreda, 1993). Die Herleitung der Gewichte geschieht dabei unter Maximierung der Kovarianz. Die Variablen des reduzierten Merkmalraums werden auch Komponenten der partiellen kleinsten Quadrate genannt, und die PLSR findet lineare Kombinationen der Komponenten, welche die Abweichungen sowohl in den Eingangs- wie auch in den Ausgangsvariablen erklären (Liu et al., 2022; Pirouz, 2012).

Das Ziel der Interpolation ist es auf Basis einer oder mehrerer Eingangsvariablen eine Ausgangsvariable vorherzusagen. Die Thin Plate Spline (TPS)-Interpolation stellt eine multivariate Spline Interpolation im 2-dimensionalen Raum dar, die auf einer radialen Basisfunktion und der euklidischen Distanz zwischen Datenpunkten beruht (Keller & Borkowski, 2019). Sie lässt sich physikalisch mit einer dünnen flexiblen Metallplatte vergleichen: Lenkt man die Platte in z -Richtung, orthogonal zur Ebene, aus, kann man das Heben der Platte als univariate Verschiebungen der x - und y -Koordinaten ansehen, wobei die mittlere quadratische Auslenkung proportional zum Quadrat der Krümmung der Metallplatte ist. Auf die TPS-Interpolation bezogen, transformiert diese die Kontrollpunkte in Zielpunkte im Koordinatensystem unter Minimierung der Energie, wobei die Energie als das Integral der Summe der quadrierten Ableitungen zweiten Grades definiert ist (Keller & Borkowski, 2019; Menesatti et al., 2012). Dieser lokale Interpolationsansatz macht TPS besonders unempfindlich gegenüber Ausreißern und optimal für Anwendungen der räumlichen Verzerrung (Menesatti et al., 2012).

2.4.3 Klassifikation

Im Zuge dieser Arbeit wird vor allem der Random Forest (RF) Algorithmus für Klassifikationsaufgaben Anwendung finden. Random Forests gehören zu den überwachten Ensemble-Lernmethoden und nutzen mehrere Entscheidungsbäume, um Klassifikations- oder Regressionsaufgaben zu lösen (Kullarni & Sinha, 2013). Ensemble Lernmethoden zeichnen sich dadurch aus, dass sie oftmals genauer sind wie, bezogen auf den Random Forest, ein einzelner Entscheidungsbaum (Kullarni & Sinha, 2013). Der Aufbau eines Random Forest beginnt mit der Auswahl einer zufälligen Teilmenge der Trainingsdaten und einer zufälligen Auswahl an Eingangsvariablen, auf dessen Grundlage dann ein Entscheidungsbaum erstellt wird (Kullarni & Sinha, 2013). Ein einzelner Entscheidungsbaum wird durch rekursive Aufteilung der annotierten Trainingsdaten erstellt. An jedem internen Knoten wählt der Algorithmus jenes Merkmal aus, welches die Daten am besten aufteilt, um das gewünschte Ergebnis zu erzielen. Der Entscheidungsbaum wächst nun so lange weiter, bis ein Haltekriterium, wie beispielsweise eine vorher festgelegte maximale Tiefe des Entscheidungsbaums, erfüllt ist (Ali et al., 2012). Im gesamten RF wird nun, jedes Mal mit einer anderen Teilmenge der Trainingsdaten, eine vorher festgelegte Anzahl an Entscheidungsbäumen erstellt. Die endgültige Vorhersage des RF wird nun anhand eines Mehrheitsvotums der Vorhersagen der einzelnen Entscheidungsbäume getroffen (Ali et al., 2012; Kullarni & Sinha, 2013). Einige Vorteile von Random Forests sind, dass sie mit hochdimensionalen multivariaten Datensätzen umgehen können und sowohl binäre, kategorische wie auch kontinuierliche Eingangsvariablen akzeptieren. RFs weisen eine hohe Genauigkeit und gute Generalisierungseigenschaften auf und sind als Ensemble-Methode robuster gegenüber *Overfitting* (Ali et al., 2012).

2.5 Ansätze des tiefen Lernens

2.5.1 Künstliche neuronale Netze mit voller Konnektivität

Künstliche neuronale Netze (engl. ANN) mit voller Konnektivität (engl. FCANN – fully connected artificial neural networks) umfassen eine Reihe verschiedenster Methoden des maschinellen Lernens, welche von der Struktur und Funktion eines biologischen Gehirns inspiriert sind. Sie bestehen aus Schichten miteinander verbundener Recheneinheiten, den sogenannten Neuronen, welche, ähnliche wie Neuronen im menschlichen Gehirn, Informationen empfangen, verarbeiten und weiterleiten können (Dastres & Soori, 2021). Abbildung 7 zeigt ein einfaches künstliches neuronales Netz bestehend aus 3 grundsätzlichen Arten von Schichten: Die Eingabeschicht ist die erste Schicht des neuronalen Netzes, welche die Eingabedaten, in der Form von Bildern, Text oder Zahlen, erhält, wobei jedes Eingangsneuron in dieser Schicht einem Merkmal in den Eingabedaten entspricht (Dastres & Soori, 2021). Die Ausgabeschicht generiert das finale Ergebnis, wobei die Anzahl der Neuronen in dieser Schicht von der Fragestellung abhängt. Während bei einer binären Klassifikationsaufgabe die Ausgabeschicht aus einem Neuron besteht, besteht sie beispielsweise bei Mehrklassen-Klassifizierungsaufgaben aus so vielen Neuronen wie es Klassen gibt, wobei jedes Neuron die Wahrscheinlichkeit für die jeweilige Klasse angibt (Lecun et al., 2015). Zwischen der Eingabe- und Ausgabeschicht befinden sich eine oder mehrere versteckte Schichten, welche die Eingabedaten prozessieren und Merkmale extrahieren, wobei die Anzahl der Schichten sowie die Anzahl der Neuronen innerhalb der Schichten die Komplexität und Tiefe des Netzwerks beeinflussen (Lecun et al., 2015). Jedes Neuron empfängt die Ausgaben von Neuronen der vorangegangenen Schicht, wobei jede Verbindung zwischen einzelnen Neuronen eine einstellbare Gewichtung erhält. Ein einzelnes Neuron berechnet nun also die gewichtete Summe aus den Ausgaben der vorangegangenen Schicht und den entsprechenden Gewichten (Dastres & Soori, 2021; Lecun et al., 2015). Anschließend wird das Ergebnis in eine mathematische, nicht-lineare Aktivierungsfunktion gegeben,

wodurch das Netzwerk nicht-lineare und abstrakte Zusammenhänge erkennen kann, welche für die Ausgabeschicht jedoch linear separierbar werden (Lecun et al., 2015). Typische Aktivierungsfunktionen sind beispielsweise die gleichgerichtete lineare Einheit (ReLU), die Sigmoid-Funktion, oder die Hyperbeltangens-Funktion (Lecun et al., 2015). Aus Gründen der Verständlichkeit lassen wir den Verzerrungswert (englisch *bias*) an dieser Stelle unbeachtet. Die Ausgabe des Neurons wird nun an die Neuronen der nächsten Schicht weitergeleitet (Dastres & Soori, 2021; Lecun et al., 2015).

Ziel des Netzwerktrainings ist es die Gewichte, welche die Neuronen miteinander verbinden so anzupassen, dass der Fehler zwischen dem vorhergesagten Ergebnis und dem tatsächlichen Ergebnis minimiert wird (Dastres & Soori, 2021). Wie bei allen überwachten Trainingsmethoden wird hierfür ein Trainingsdatensatz genutzt, der möglichst vielfältige Beispiele für die Eingangsdaten sowie die entsprechenden Ausgangsdaten beinhaltet. Ein diverser Trainingsdatensatz ermöglicht es dem Netzwerk adäquat zu generalisieren und *Overfitting* zu vermeiden. Dabei kann ein Trainingsdatensatz künstlich durch sogenanntes Augmentieren erweitert werden. Im Kontext der Bildanalyse können dafür beispielsweise geometrische und chromatische Bildtransformationen genutzt werden, um beispielsweise verschiedene Orientierungen eines Objekts unter variierenden Lichtbedingungen im Bild zu simulieren. Zur Berechnung des Fehlers zwischen Vorhersage und Referenz wird eine, für die Fragestellung angemessene, Verlustfunktion genutzt (Lecun et al., 2015). Die Anpassung der Gewichte geschieht in der Regel mittels Back-Propagation, bei der der Gradient des Fehlers in Bezug auf die Gewichte des Netzes berechnet wird und die Parameter mit einem Optimierungsalgorithmus wie dem stochastischen Gradienten Abstieg (englisch *stochastic gradient descent*, SGD) aktualisiert werden (Ding et al., 2013; Lecun et al., 2015). Der Durchlauf einer Trainingsiteration inklusive Verlustberechnung und *Back-Propagation* wird auch Epoche genannt. Die Trainingsdaten werden pro Epoche also einmal durch das gesamte Netzwerk prozessiert und anschließend mit Hilfe der Referenzdaten und den Vorhersagen des Netzwerks ein Verlust berechnet, auf dessen Basis die *Back-Propagation* durchgeführt wird. Die Ausgabeschicht gibt nun für jede Klasse eine Wahrscheinlichkeit an, welche auch als Konfidenzwert bezeichnet wird. Jene Klasse mit dem höchsten Konfidenzwert ist somit die vorhergesagte Klasse.

FCANNs (Abbildung 7) bieten erweiterte Möglichkeiten im Vergleich zu klassischen maschinellen Lernverfahren, da sie hochabstrakte Zusammenhänge identifizieren können und robuster gegenüber Rauschen in den Daten sind (Lecun et al., 2015).

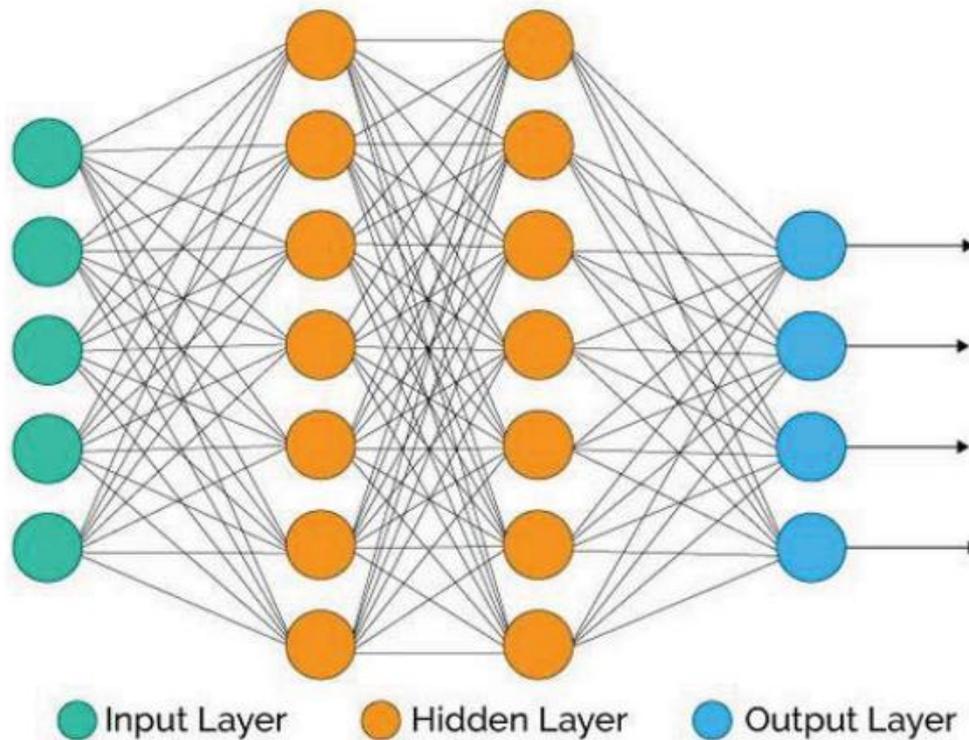


Abbildung 7: Struktur eines vollständig vernetzten künstlichen neuronalen Netzwerks (Dastres & Soori, 2021)

2.5.2 Neuronale Faltungsnetzwerke

Bildklassifikation: Neuronale Faltungsnetzwerke (engl.: CNN – convolutional neural networks) sind eine Gruppe von ANNs, die entwickelt wurden, um Daten zu prozessieren, die als multiple Arrays auftreten und bei denen man davon ausgeht, dass lokale Gruppen von Werten hochgradig miteinander korrelieren. Bilder beispielsweise bestehen aus drei 2D Arrays, welche jeweils die Pixelintensitäten für die 3 Farbkanäle enthalten und bei denen man davon ausgehen kann, dass der Wert eines Pixels stark mit den Werten der umliegenden Pixel zusammenhängt (Lecun et al., 2015). Ein Faltungsnetzwerk zur Bildklassifikation besteht, wie in Abbildung 8 dargestellt, aus Faltungsschichten (englisch *convolutional layers*), zusammenfassenden Schichten (englisch *pooling layers*), sowie Schichten mit voller Konnektivität (englisch *fully-connected layers*, wie bei klassischen FCANNs). Diese ist nämlich bei Faltungsschichten nicht gegeben, um die Kosten für die Berechnung zu minimieren (O’Shea & Nash, 2015). Faltungsschichten nutzen, genau wie zum Beispiel Kantendetektoren, Faltungsmatrizen um sogenannte Merkmalskarten (englisch *feature maps*) zu erstellen, wobei das Ergebnis der lokalen gewichteten Summe der Faltungsmatrix auch hier durch eine nicht-lineare Funktion wie zum Beispiel eine ReLU gegeben wird. Für jede Merkmalskarte nutzt die Schicht eine individuelle Faltungsmatrix, wobei die trainierbaren Gewichte die Elemente der Faltungsmatrix sind. Geht man von einem Bild mit den Dimensionen $64 * 64 * 3$ (RGB-Bild mit den Dimensionen $64 * 64$) und 3 Faltungsmatrizen mit den Dimensionen $3 * 3$ aus, besitzt die gesamte Faltungsschicht 27 trainierbare Gewichte ($3 * 3 * 3$) anstatt 12288 ($64 * 64 * 3$), wie es bei einer Schicht mit voller Konnektivität der Fall wäre. Alle Neuronen einer Merkmalskarte teilen sich also dieselben Gewichte der entsprechenden Faltungsmatrix. Dieses Teilen von Gewichten spart eine große Menge an benötigter Rechenleistung ein (Lecun et al., 2015; O’Shea & Nash, 2015). Auf eine Faltungsschicht folgt normalerweise eine zusammenfassende Schicht, welche die Merkmalskarten der Faltungsschicht in ihren Dimensionen reduziert, um die Anzahl an zu trainierenden Parametern weiter zu reduzieren (O’Shea & Nash, 2015). Gegeben sei eine Merkmalskarte mit den Dimensionen von $8*8$ Pixeln und ein

zusammenfassendes Fenster von 2*2 Pixeln, dann passt das zusammenfassende Fenster 16-mal in die Merkmalskarte und reduziert die Dimensionen von 8*8 auf 4*4 Pixel. Dabei ist der resultierende Pixel entweder der maximale Wert aller unter dem zusammenfassenden Fenster liegenden Pixel, oder das Mittel aus diesen (O'Shea & Nash, 2015; Paszke et al., 2019). Faltungsschichten und zusammenfassende Schichten werden normalerweise mehrfach (10-20 Schichten) hintereinander geschichtet, bis die Dimensionen der Merkmalskarten so klein sind, dass die Daten mittels einer Schicht mit voller Konnektivität in die Form eines klassischen FCANNs überführt werden kann, um das finale Ergebnis vorherzusagen (O'Shea & Nash, 2015). Ein solches Netzwerk ist in der Lage Bilder als Ganzes zu klassifizieren, um beispielsweise die Art der Blattkrankheit zu bestimmen, welche auf dem Bild zu sehen ist.

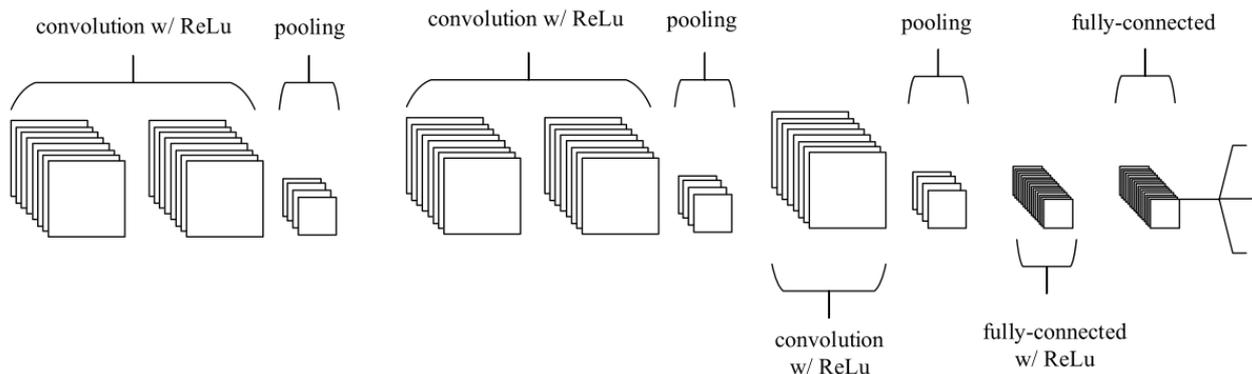


Abbildung 8: Struktur eines neuronalen Faltungsnetzwerks zur Bildklassifikation (O'Shea & Nash, 2015)

Objekterkennung und Segmentierung: Zur Erkennung und Segmentierung von Objekten in Bildern können unterschiedliche Netzwerkarchitekturen genutzt werden:

Encoder-Decoder Netzwerke stellen eine Erweiterung der zuvor beschriebenen Encoder-Einheit zur Bildklassifikation dar, um zusätzlich sogenannte semantische bzw. Instanz Segmentierungen oder beides (panoptische Segmentierungen) durchzuführen. Dabei stellt die semantische Segmentierung eine Aufteilung des Bildes in Vordergrund und Hintergrund dar, während die Instanz Segmentierung einzelne Objekte im Bild maskiert. Hierzu wird die Encoder-Einheit, um eine sogenannte Decoder-Einheit erweitert. Die Architektur und Funktionsweise solcher Encoder-Decoder Netzwerke soll hier exemplarisch am sogenannten U-Net Netzwerk beschrieben werden, welches in Abbildung 9a dargestellt ist (Ronneberger et al., 2015). Während der Encoder die Kanäle der Merkmale unter Reduzierung der Dimensionen des Bildes erhöht, verringert der Decoder wieder die Anzahl der Kanäle der Merkmale, erhöht aber die Dimensionen des Bildes. Der Decoder stellt somit die ursprünglichen Dimensionen des Bildes wieder her und erstellt auf Grundlage der extrahierten Merkmale eine Maske des Objekts. Dabei besteht die Decoder-Einheit ebenfalls aus Faltungsschichten, besitzt anstatt zusammenfassender Schichten jedoch erweiternde Schichten (Abbildung 9a, grüne Pfeile). Als Besonderheit besitzen U-Net Architekturen noch *Skip-Connections*. Um feinere Details des Eingabebildes beizubehalten, gibt es direkte Verbindungen zwischen der Encoder- und Decoder-Einheit, welche die erweiterten Daten der Decoder-Schicht mit den Merkmalskarten der entsprechenden Encoder-Schicht verbinden (Abbildung 9a, graue Pfeile) (Ronneberger et al., 2015).

Sogenannte Regionen basierte Faltungsnetzwerke (engl.: R-CNN – region based convolutional neural network) verfolgen einen anderen Ansatz. Dort werden Regionen

innerhalb eines Bildes mittels sogenannter Ankerpunkte vorgeschlagen, welche schließlich einzeln klassifiziert werden (Ren et al., 2017), wodurch man die umschließenden Rechtecke sowie die Klassifizierung einzelner Objekte im Bild erhält. In diesem Fall spricht man auch von Objekterkennungsnetzwerken. Ein R-CNN besteht aus mehreren verbundenen neuronalen Faltungsnetzwerken, welche somit gemeinsam trainiert werden (Abbildung 9b). Das sogenannte *Region Proposal Network* dient zur Generierung geeigneter Bildausschnitte, die vom darauffolgenden Netzwerk klassifiziert werden (Ren et al., 2017).

CNNs erzielten bereits herausragende Ergebnisse in der Bildanalyse (Lecun et al., 2015), weshalb verschiedene CNNs zur Klassifikation, Objekterkennung und Instanz-Segmentierung im Zuge dieser Arbeit Anwendung fanden. Dabei sind vor allem *Faster R-CNNs*, welche eine Weiterentwicklung der klassischen R-CNNs darstellen, und *Mask R-CNNs* zu nennen, welche wiederum auf den *Faster-RCNNs* aufbauen. So besteht das Grundgerüst dieser Netzwerke aus einem sogenannten *Feature Pyramid Backbone* Netzwerk, einem *Region Proposal* Netzwerk, sowie dem Detektion- und Klassifikationszweig, sowie, im Falle des *Mask R-CNNs*, einem zusätzlichen Segmentierungszweig. Das *Feature Pyramid Backbone* Netzwerk dient der Extraktion von Merkmalen und besteht aus vielen *Residual Neural Network (ResNet)* Schichten in Kombination mit einem Top-Down-Pfad, welcher die Erfassung umfangreicher räumlicher Details ermöglicht, welche für eine präzise Lokalisierung der Objekte im Bild entscheidend sind (X. Li et al., 2019).

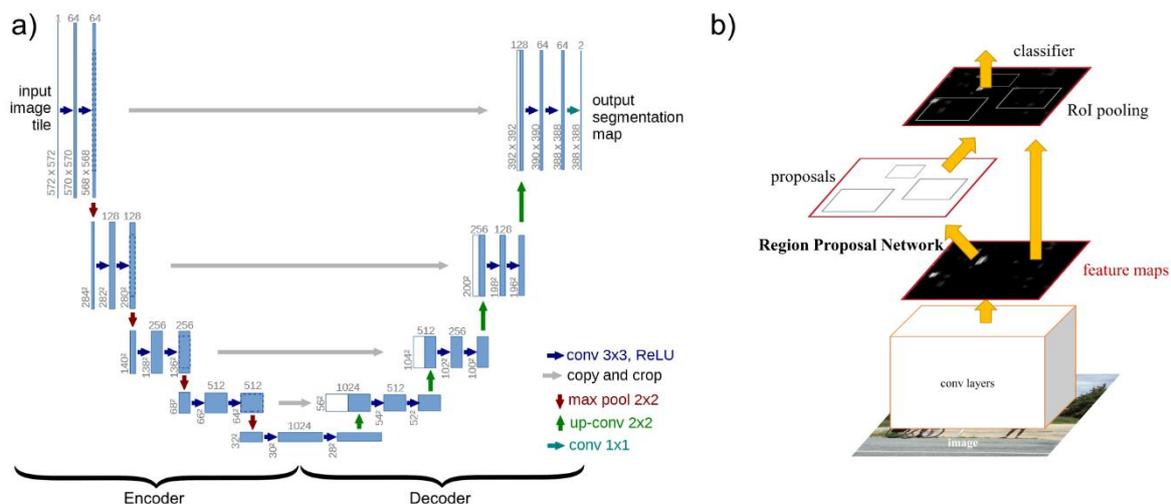


Abbildung 9: a) Struktureller Aufbau eines U-Net CNNs zur Segmentierung von Bildern; b) Struktureller Aufbau eines R-CNNs. Dieses besteht aus mehreren miteinander verbundenen CNNs zur Objekterkennung und Klassifizierung.

2.6 Statistische Werkzeuge zur Methodvalidierung

Nachfolgend werden verschiedene Verfahren zur Beurteilung der korrekten Klassifikation, der richtigen Lokalisierung und des statistischen Zusammenhangs mittels Regressionen eingeführt.

Klassifikationen: Zur Beurteilung der Klassifikationsgüte eines Modells bediente man sich einer Wahrheitsmatrix und verschiedener Gütemaße. Eine „richtig positive (RP)“ Vorhersage bedeutet, dass das Modell einen bestimmte Fall angibt, welcher wahrheitsgemäß gegeben ist. Eine „richtig negative (RN)“ Vorhersage bedeutet, dass das Modell ein Nicht-Vorhandensein eines bestimmten Falls angibt, welcher wahrheitsgemäß auch nicht gegeben ist. Eine „falsch positive (FP)“ Vorhersage bedeutet, dass das Modell einen Fall angibt, welcher wahrheitsgemäß jedoch nicht gegeben ist. Eine „falsch negative (FN)“ Vorhersage bedeutet,

dass das Modell ein Nicht-Vorhandensein eines Falls angibt, welcher wahrheitsgemäß aber gegeben ist (Visa et al., 2011). Die Korrektklassifikationsrate (englisch: accuracy score) gibt an, wie viele der insgesamt klassifizierten Fälle richtig erkannt wurden und berechnet sich wie folgt (Pedregosa et al., 2012):

$$\text{Korrektklassifikationsrate} = \frac{RP + RN}{RP + RN + FP + FN} \quad [10]$$

Die Spezifität (englisch precision) gibt die Fähigkeit eines Modells an, einen bestimmten Fall zu erkennen und gleichzeitig andere Fälle auszuschließen und berechnet sich wie folgt (Pedregosa et al., 2012):

$$\text{Spezifität} = \frac{RP}{RP + RN} \quad [11]$$

Die Sensitivität (englisch recall) beschreibt die Fähigkeit eines Modells, alle Vorkommen eines bestimmen Falls zu finden, auch wenn dadurch „falsch positive“ Vorhersagen gemacht werden. Die Sensitivität berechnet sich wie folgt (Pedregosa et al., 2012):

$$\text{Sensitivität} = \frac{RP}{RP + FN} \quad [12]$$

Lokalisation: Gilt es ein (klassifiziertes) Objekt im Bild zu lokalisieren, wird die Güte der Lokalisation unter Nutzung des umschließenden Rechtecks des Objekts bestimmt. Als Gütemaß wird hier der mAP_{50} (englisch für *mean average precision at 50 percent overlap*) verwendet, welcher sich auf die durchschnittliche Präzision der Lokalisation bezieht, bei der das vorhergesagte umschließende Rechteck mindestens 50 Prozent mit dem tatsächlichen Rechteck des Objekts in einem Bild übereinstimmt. Der mAP_{50} nimmt Werte zwischen 0 und 1 an, wobei ein höherer Wert eine höhere Vorhersagegenauigkeit des Modells wiedergibt (Lopez-Marcano et al., 2021).

Regression: Um den Zusammenhang zwischen einer unabhängigen und einer abhängigen Variable zu untersuchen, werden verschiedene Regressionsstatistiken verwendet. Regression wurden in dieser Arbeit auf Basis des quadratischen Fehlers berechnet (Harris et al., 2020). Der Korrelationskoeffizient diente als ein Indikator für die Genauigkeit der Methode. Weitere Indikatoren bildeten je nach Anwendungsfall der mittlere absolute Fehler (MAE) oder der mittlere absolute Fehler in Prozent (MAPE). Galt es zu untersuchen, ob sich verschiedene Gruppen unterschieden, wurde im Falle metrischer Daten und unter Einhaltung der Normalverteilung der Daten eine Varianzanalyse (ANOVA) durchgeführt. Handelte es sich um ordinal skalierte Daten, gepaarte Variablen oder lag keine Normalverteilung vor, nutze man stattdessen den nicht-parametrischen *Wilcoxon-Vorzeichen-Rang-Test*. Zur Quantifizierung der Unterschiede zwischen verschiedenen Gruppen nutzte man die Effektgrößenanalyse nach Cohen. Die Effektgröße D misst die standardisierte Differenz zwischen zwei Mittelwerten und wird wie folgt kategorisiert: klein $<0,5$, mittel $0,5-0,8$, groß $>0,8$ (Cohen, 1988).

3 Analyse wichtiger Zuchtmerkmale mittels maschinellen Sehens

3.1 Anmerkung zur Zitation der eigenen Veröffentlichungen

Im Zuge dieser Arbeit war der Autor an insgesamt 3 Veröffentlichungen beteiligt. In der Veröffentlichung Wilhelm et al., 2022 war der Autor mit einer Co-Autorschaft beteiligt. Diese Veröffentlichung ist jedoch thematisch nicht relevant für diese monographische Abhandlung der Doktorarbeit. Des Weiteren war der Autor an 2 Veröffentlichungen als Hauptautor beteiligt (Jollet et al., 2021, 2023). Die Zitate aus der früheren Veröffentlichung (Jollet et al., 2021) sind im Fließtext gegeben. Die Veröffentlichung Jollet et al., 2023 bündelt einen Großteil der Ergebnisse und Erkenntnisse, die über die Forschungen der qualitativen Merkmalsanalyse von Bohnenhülsen erzielt wurden. Aus diesem Grund wurde auf einzelne Zitierungen im Fließtext abgesehen und es werden hier alle Kapitel aufgelistet, welche übersetzte Zitate aus dem Englischen aus dieser Veröffentlichung enthalten. Große Teile dieser Abschnitte sind aus der Veröffentlichung übernommen worden. Die Übersetzung erfolgte frei und nicht Wort für Wort. Oftmals wurden Textelemente neu angeordnet. Falls in Abbildungen, welche aus den Veröffentlichungen übernommen wurden, ebenfalls textliche Elemente vorhanden waren, wurden diese ebenfalls aus dem Englischen ins Deutsche übersetzt. Die folgenden Kapitel sind von dieser Anmerkung betroffen:

3.23.2 Präzise Hochdurchsatzanalyse wichtiger Formparameter länglicher Gemüsesorten:

3.3 Nicht-invasive Analyse der Samenanlage von Leguminosen

3.4.2 Farbklassifikation und Helligkeitsanalyse unter kontrollierten Bedingungen zur Sortenbeurteilung

3.2 Präzise Hochdurchsatzanalyse wichtiger Formparameter länglicher Gemüsesorten

3.2.1 Forschungsstand

Um die Fruchtkörper länglicher Gemüsesorten mittels 2D RGB-Verfahren analysieren zu können, müssen diese zunächst auf den 2-dimensionalen Bildern identifiziert und lokalisiert werden. Zwar gibt es beispielsweise für die Sojabohne einige Studien, welche die Detektion der Hülsen an den Pflanzen mittels 2D RGB-Verfahren in Kombination mit Methoden des maschinellen Sehens untersuchten, jedoch gibt es bei dieser Herangehensweise enorme Hürden, da die Blätter der buschigen Pflanzen die Hülsen oftmals verdecken. Eine Forschungsgruppe ging das Problem an, indem sie je 4 Sojabohnenpflanzen in einem Topf heranzogen und diese dann bei Vollreife der Hülsen aus 4 Richtungen mit einem Abstand von 90° fotografierten. Sie nutzten schließlich ein modifiziertes YOLOv3 neuronales Netzwerk, um die Hülsen zu detektieren und nach Anzahl der vorhandenen Samen in der Hülse zu klassifizieren (Lu et al., 2022). Diese Klassifizierung ist an Sojabohnen gut durchführbar, da diese sehr stark markieren. Eine andere Forschungsgruppe führte einen vergleichbaren Ansatz unter Feldbedingungen durch. In diesem Versuch wurde ein spezielles einreihiges Anbauverfahren genutzt, sodass zwischen den einzelnen Pflanzreihen genug Platz war, damit ein Feldroboter, ausgerüstet mit 2 Videokameras, welche orthogonal in beide Richtungen zur Fahrtrichtung filmten, dazwischenfahren konnten. Die Pflanzen wurden mit dem System von beiden Seiten abgefahren und gefilmt. Sie trainierten schließlich das neuronale Netzwerk RetinaNet, um zunächst die Hülsen in Einzelbildern des Videostreams zu detektieren. Es wurden dabei 1-3 Bilder pro Seite aus dem Videostream extrahiert und im zweiten Schritt wurde das neuronale Netzwerk darauf trainiert, eine Vorhersage über die tatsächliche Anzahl der Hülsen pro Pflanzreihe zu machen. Sie erreichten einen Korrelationskoeffizienten von

$R=0.65$ bei je 3 Bildern pro Blickrichtung (Riera et al., 2020). Wie anhand der beschriebenen Studien dargestellt, benötigt es für die Detektion länglicher Früchte mit Hilfe von 2D RGB-Verfahren angepasste Anbau Methoden, um die Pflanzen von mehreren Blickrichtungen fotografieren zu können. Hülsenfrüchtler wie die Sojabohne stellen im Vergleich zu anderen länglichen Gemüsesorten (wie beispielsweise Gurke, Zucchini, oder Aubergine) die herausforderndsten Kulturarten in Bezug auf die digitale Pflanzenphänotypisierung dar, da die Pflanzen bei einigen Arten sehr buschig wachsen können und die vergleichsweise kleinen Hülsen stark überdeckt werden. Zudem besitzen die Hülsen bei vielen Vertretern die gleiche Farbe wie das Blattwerk, was eine Detektion auf 2D RGB-Bildern weiter erschwert. Um beim Beispiel der Sojabohne zu bleiben, stellt es zudem einen erheblichen Unterschied dar, ob man die Hülsen in den Bildern lediglich detektiert, um diese zu zählen und eine Ertragsabschätzung zu machen (wie in den 2 Literaturbeispielen von zuvor dargestellt), oder ob man vielfältige strukturelle Qualitätsparameter einzelner Früchte evaluieren möchte.

Aus diesem Grund fokussiert sich die Arbeit auf die Evaluierung struktureller Qualitätsparameter der geernteten Früchte mit Hilfe von 2D RGB-Verfahren. Diese Herangehensweise ist deutlich erfolgsversprechender, um auch solche Qualitätsunterschiede feststellen zu können, welche mit etablierten manuellen Verfahren gar nicht oder nur erheblichem zeitlichen Aufwand bewertet werden können. Im Folgenden wird daher auf Studien eingegangen, in denen verschiedene Verfahren zur Erfassung von Formparametern und die notwendigen Voraussetzungen, wie die Detektion der Fruchttenden, beschrieben werden.

In einem konkreten Beispiel durften die Enden einer Bohnenhülse nicht in die Längenanalyse mit einfließen. Dies wurde durch eine automatische Modifikation (in der Quelle nicht näher spezifiziert) der Bilder erreicht (Torres et al., 2012). Eine andere Forschungsgruppe entwickelten das Software-Plugin *CultivarJ* für die Bildverarbeitungs-Plattform ImageJ, um Qualitätsparameter unter anderem von Bohnen- (*Phaseolus vulgaris*) und ErbsenHülsen (*Pisum sativum*) abzuleiten. Die Bilder der Hülsen wurden aus der Aufsicht vor weißem Hintergrund gemacht und es wurde darauf geachtet, dass sich einzelne Hülsen im Bild nicht berühren. Aus der binären Bildmaske, sowie den Skeletten der segmentierten Hülsen wurden dann die Länge, die Breite, die Krümmung, sowie die Größe der Spitze abgeleitet. Ebenfalls bei hier wurde die Länge wieder ohne die Enden der Hülse berechnet (wie auf Bildern der Veröffentlichung zu sehen ist), jedoch ist die algorithmische Umsetzung, wie die Spitze und der Blütenansatz ausgeschlossen wurden, nicht näher beschrieben worden. Cubero et al. (2014) entwickelten eine Methode für Weintrauben aufgenommen unter kontrollierten Bedingungen, um die Stiele an einzelnen Weintrauben für akkurate Gewichts- und Größenabschätzungen zu erkennen und aus dem Bild zu entfernen. Diese Technik basierte auf der Analyse der Radius-Richtungssignatur der Kontur der segmentierten Frucht und funktionierte gut für Früchte wie beispielsweise Trauben (100% Detektionsgenauigkeit), Mandarinen (100% Detektionsgenauigkeit) und Äpfel/Pfirsich (96,7% Detektionsgenauigkeit), welche eine sehr runde Form besitzen und bei denen sich der Stiel deutlich vom Rest der Frucht absetzt. Die Radius-Richtungssignatur gibt dabei an, in welche Richtung sich eine Kontur krümmt. Am Ansatz des Stiels der genannten runden Früchte ändert sich diese Richtung deutlicher. In Vorbereitung zu dieser Arbeit ist ein Verfahren für Bohnenhülsen entwickelt worden, welches sich eines RF-Klassifikators und Methoden der klassischen Bildverarbeitung bediente, um Spitzen und Blütenansätze im Bild zu klassifizieren und lokalisieren. Die Validierung erfolgte an 140 grünen und gelben Bohnenhülsen mit einem runden Querschnitt, für die eine durchschnittliche Klassifikationsgenauigkeit von 88,7% erreicht wurde. Nach Entfernung dieser Strukturen konnte die Länge, sowie das Kaliber mit Korrelationskoeffizienten von 0,96 zu manuellen Messungen berechnet werden (Jollet et al., 2021). In einer anderen Studie wurden die Enden von Gurken, welche einzeln unter

kontrollierten Bedingungen aufgenommen wurden, lokalisiert, indem die Krümmungen der Gurkenkontur mit lokalen Ableitungen über B-Spline-Regressionsmodelle approximiert wurden. Durch die Annahme, dass die Kontur an den Enden die stärkste Krümmung aufweist, konnten diese erfolgreich lokalisiert werden (Soleimanipour & Chegini, 2019). Eine andere Arbeitsgruppe berechnete die Krümmung der Zentrallinie von Gurken, indem sie diese mit Hilfe eines Krümmungskreises approximierten. Bei diesem Ansatz stellten die Autoren die Annahme, dass sich die Krümmung der Zentrallinie nicht signifikant ändert und dadurch, dass sie bei S-Förmigen Gurken zwei Krümmungskreise definierten, erreichten sie bei einer Stichprobengröße von 148 Gurken einen Determinationskoeffizient von $R^2 = 0,996$ zwischen der Zentrallinie und dem approximierten Bogen (Jin & Cai, 2020). Bezüglich der Berechnung der Breite länglicher Früchte wird oftmals ein Verfahren verwendet, welches die orthogonale euklidische Distanz jedes Punktes der Zentrallinie zur Kontur misst (Jollet et al., 2021; Van Eck et al., 1998). Während Van Eck et al. (1998) die Orthogonalität der Messung sicherstellten, indem sie die Gurken händisch vertikal im Bild ausrichteten und gekrümmte Gurken in gerade Gurken extrapolierten, nutzen der Autor dieser Arbeit in Jollet et al., 2021 die umliegenden 10 Pixel eines jeden Pixels der Zentrallinie, um orthogonale Linien zu approximieren.

3.2.2 Material und Methoden

3.2.2.1 Modellpflanze

Zur Beantwortung der Forschungsfrage wird die Gartenbohne (*Phaseolus vulgaris*) als Modellpflanze genutzt, da sie Eigenschaften aufweist, die eine vollumfängliche Beantwortung der Forschungsfragen zulässt. Je nach Wuchsform werden die einjährig wachsenden Pflanzen entweder als Buschbohnen oder Stangenbohnen bezeichnet. Stangenbohnen besitzen eine hohe und längliche Wuchsform und benötigen eine Kletterhilfe für ihr Wachstum, während Buschbohnen frei wachsen und eine buschige Wuchsform besitzen. In dieser Arbeit wurden ausschließlich Hülsen verschiedenen Buschbohnenarten untersucht. So besitzen die Hülsen Blütenansätze und Spitze, welche fließend am Rest der Hülsen ansetzen. Diese anatomischen Eigenschaften stellen die Methoden des maschinellen Sehens vor besondere Herausforderungen, da die Ansatzpunkte der Spitze und des Blütenansatzes an der Hülse definiert werden müssen. Weiterhin weisen Bohnenhülsen unterschiedlicher Sorten eine hohe Variabilität in ihren strukturellen und chromatischen Eigenschaften auf, was dienlich ist, um die Robustheit der Methoden des maschinellen Sehens (MS) zu testen. Dabei lassen sich die untersuchten Sorten je nach Erscheinungsbild der Hülse unterschiedlich klassifizieren. So gibt es Sorten mit grüner, gelber und zweifarbiger Hülsenfarbe, wobei letztere gelbe Hülsen mit leichtem Grünstich darstellen. Diese Eigenschaften werden für das folgende Kapitel über die Farbanalysen wichtig. Der Querschnitt der Hülsen kann flach, oval oder rund ausgeprägt sein, was ein 2D RGB-Verfahren natürlicherweise vor besondere Herausforderungen stellt, da die Hülsen lediglich aus einer Perspektive fotografiert werden können. In dieser Arbeit sollte daher auch untersucht werden, wie groß dieser perspektivisch bedingte Messfehler ist. Die Breite der Hülse wird durch das Kaliber definiert. Das Kaliber stellt die breiteste Stelle der Hülse dar. Dieses lässt sich einerseits gut manuell messen, für die Bildverarbeitung bedeutet dies jedoch, dass man den Breitenverlauf der Hülse über die gesamte Länge kennen muss, um die breiteste Stelle zu identifizieren. Das Kaliber der Hülsen wurde in extrem fein (<6,5mm), sehr fein (6,5-8,0mm), fein (8,0-9,0mm), mittelfein (9,0-10,5mm) und grob (>10,5mm) eingeteilt (Lindner & Billmann, 2006). Die Länge wurde in kurz (<120mm), mittel ($\geq 120\text{mm}$, $\leq 150\text{mm}$) und lang (>150mm) unterteilt (Jusoh, 2017). Weiterhin können Bohnenhülsen äußerst variable Krümmungen aufweisen: Von gleichmäßig gekrümmt, über abgeknickte Enden, S-Formen oder 3-dimensionale Verdrehungen kann alles vorkommen. Es gibt keine Vorgaben in der Literatur zur Klassifizierung der Krümmung und Züchtungsunternehmen verwenden ihre eigenen Skalen, welche sich ebenfalls weiter voneinander unterscheiden können.

3.2.2.2 Pflanzenanzucht

Der Standort Rosdorf, an dem die unterschiedlichen Bohnensorten für diese Arbeit angebaut wurden, liegt im Übergangsbereich zwischen maritimem und kontinentalem Klima der gemäßigten Breiten. Der Standort liegt 150 Meter über dem Meeresspiegel. Zwischen den Jahren 2001 und 2019 herrschte dort eine Jahresdurchschnittstemperatur von 9,7°C. Der Niederschlag war im Deutschlandvergleich mit einer Niederschlagssumme von 667mm/Jahr eher gering. Im Schnitt schien die Sonne 1422 Stunden im Jahr. Während der Anbauperiode zwischen Mai und August wurden Tagesdurchschnittstemperaturen von 13,4-18,1°C und durchschnittliche Niederschlagsmengen zwischen 60,8mm und 78,1mm pro Monat gemessen, wobei der Juli der regenreichste Monat ist (Quelle: Wetterstation Göttingen, Deutschland; <https://www.wetterstation-goettingen.de/>; 09.03.2023).

Die Pflanzen wurden im Jahr 2021 im konventionellen Feldanbau angepflanzt, um eine optimale phänologische Entwicklung der unterschiedlichen Genotypen zu gewährleisten. Effekte wie Krankheiten, Pathogene, und abiotische Stressfaktoren sollten so gering wie möglich gehalten werden, um die genotypisch veranlagte Entwicklung der Bohnenhülsen zu gewährleisten.

Im Sommer 2021 wurden insgesamt 40 verschiedene Buschbohnsorten unter Feldbedingungen am Standort Rosdorf nach Standardpraktiken der *Van Waveren Saaten GmbH* angebaut. Das Feld wurde am 10. April 2021 mit 500kg/ha Kalkstickstoff behandelt, um die Erde zu düngen und die Larven der Bohnenfliege (*Delia platura*) zu bekämpfen. Eine Herbizidbehandlung mit 0,5l/ha *Spectrum*[®] (BASF, Ludwigshafen, Deutschland) + 0,15l/ha *Centium 36CS* (FMC Agricultural Solutions, Philadelphia, USA) erfolgte 4 Tage vor Aussaat. Die Aussaat der 40 Bohnensorten erfolgte am 14. Mai 2021 mit Hilfe einer Saatmaschine in Reihen, wobei ein Reihenabstand von 75cm und ein Pflanzabstand von 8cm gewählt wurden. Eine Behandlung mit 0,25kg/ha *Pirimor*-Granulat (*Syngenta Agro GmbH*, Frankfurt, Deutschland) und 0,075l/ha *Karate Zeon* (*Syngenta Agro GmbH*, Frankfurt, Deutschland) wurde am 21. Juli durchgeführt, um herbivore Insekten zu bekämpfen. Die Ernte, Vermessung und Bildgebung der Hülsen geschah jeweils am selben Tag und wurde 2021 zweitägig am 28. und 29. Juli durchgeführt. Eine Zwischenlagerung der Hülsen war somit nicht nötig. Die Hülsen wurden zwischen den BBCH Stadien 75-79 händisch geerntet. Es wurde darauf geachtet, dass die geernteten Hülsen jenes Wachstumsstadium erreicht hatten, welches auch bei einer Frischvermarktung angestrebt wird. Die Hülsen sollten also die sortentypische Länge erreicht haben, die Samen aber noch weich sein. Es wurde weiterhin darauf geachtet, dass Hülsen sowohl mit wie auch ohne Blütenstiel geerntet wurden, um eine ausreichende Variabilität zu gewährleisten und somit die Robustheit der entwickelten MS-Methoden gewährleisten zu können. Bei der Auswahl der Sorten wurde darauf geachtet, dass die Hülsen sich sowohl in Form als auch Farbe deutlich unterschieden. Insgesamt gab es 31 Sorten mit grünen Hülsen und 9 Sorten mit Gelben Hülsen, von denen jedoch 6 Sorten mehr oder weniger dazu tendierten grünstichige Hülsen zu entwickeln. Unter den grünen Sorten hatten 19 einen runden, 6 einen flachen und 6 einen ovalen Querschnitt. Letzterer stellt eine Zwischenform des runden und flachen Querschnitts dar. Unter den gelben Sorten waren 3 vom Typ *Rund*, 2 vom Typ *Oval* und 4 vom Typ *Flach*. Auch in Bezug auf die Krümmung wurde Wert auf eine große Variabilität gelegt. Lediglich 3-dimensional verdrehte Hülsen waren nicht vertreten, da es sich hierbei um eine sehr seltene Ausprägung handelt, welche vor allem durch insektenbefall und dadurch verbundene Einstichstellen hervorgerufen wird. Dem wurde beim Anbau jedoch entgegengewirkt.

Es wurden 40 Hülsen pro Sorte geerntet, aus denen jeweils 10 Hülsen pro Sorte zufällig für den *Validierungsdatensatz* ausgewählt wurden. Die Bohnenhülsen wurden direkt nach der Vermessung und Inspektion durch den Zuchtexperten in die Aufnahmebox gelegt und schließlich als Batch aufgenommen (1 Bild mit 10 Hülsen pro Sorte). Es wurde bei der

Aufnahme darauf geachtet, dass die Hülsen sich im Bild nicht berühren und dass die Reihenfolge der manuellen Evaluation eingehalten wurde. Dies ermöglichte eine genaue Zuordnung der MS-Daten zu den jeweiligen manuellen Referenzdaten. Ein *Entwicklungsdatensatz* wurde zusätzlich erstellt, welcher 27 der insgesamt 40 Sorten durch ein weiteres Bild mit den übrigen 30 Hülsen pro Sorte abbildete. Diese Hülsen wurden nicht von Experten vermessen und besitzen keine Überschneidung mit dem Validierungsdatensatz.

3.2.2.3 Erhebung der Referenzdaten

Experten der *Van Waveren Saaten GmbH* erhoben die Referenzdaten für den Validierungsdatensatz nach ihren Standardpraktiken, um die entwickelten MS-Methoden zu validieren. Die Bewertung erfolgte dabei Hülsenweise. Es wurde zunächst annotiert, ob die Hülse eine Spitze wie auch einen Blütenansatz besitzt. Die Spitze ist die Verjüngung der Hülse am dessen apikalen Ende. Der Blütenansatz befindet sich am basalen Ende, welches die Hülse mit dem Rest der Pflanze verbindet und oftmals noch den Blütenstiel besitzt (Abbildung 10). Die Länge der Hülse wurde mit einem Lineal gemessen, wobei man die Spitze und den Blütenansatz von der Längenmessung ausschloss und die Hülse, falls diese gekrümmt war, so weit wie möglich aufbog, ohne sie zu beschädigen. Das Kaliber wurde mit Hilfe einer Lochlehre gemessen, welche eine Auflösung von 0,5mm besaß. Der Durchmesser des kleinsten Lochs, durch welches die Hülse ohne Beschädigungen hindurchgeführt werden konnte, gab das Kaliber für diese Hülse an. Die Krümmung der Hülse wurde visuell bonitiert. Da nicht-gekrümmte Bohnen in Züchtungsprogrammen favorisiert werden³, unterscheidet die Züchterskala gerade bis sehr schwach gekrümmte Hülsen (Klasse 1-4), welche für den Verkauf akzeptabel sind, und deutlicher gekrümmte Hülsen (Klasse 5-9). Laut der Skala bekommt eine komplett gerade Hülse eine Bewertung von 1 und eine Hülse mit starker Krümmung eine Bewertung von 9 (Abbildung 10). Dabei ist die Krümmung nicht zwangsläufig über die gesamte Bohnenlänge gleich, da, wie in Abbildung 10 dargestellt, auch die Hülseenden stark eingerollt sein können. Weiterhin sind auch S-förmigen Krümmungen möglich, welche, unabhängig der Ausprägung der Krümmung, immer mit dem Züchterwert 9 beurteilt wurden

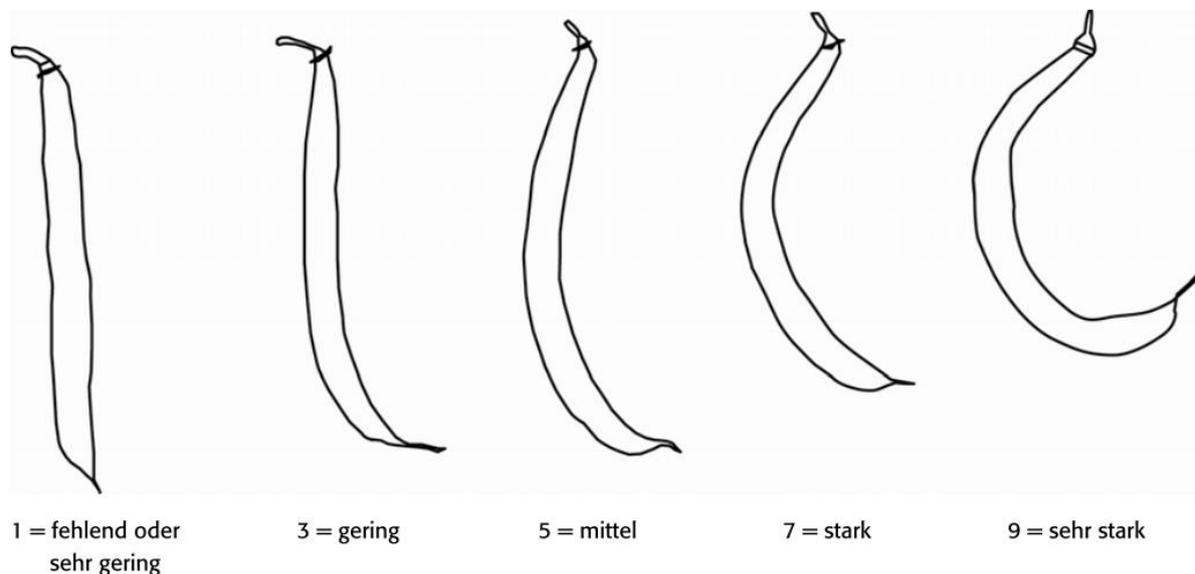


Abbildung 10: Referenzskala zur Bewertung der Krümmung (Lindner & Billmann, 2006). Der Blütenansatz der Hülsen liegt jeweils am oberen Bildrand.

³ Latauschke, 2021 - Mittelfeine Buschbohnen bei sehr guten Bedingungen mit Spitzenerträgen https://www.gartenbau.sachsen.de/download/Buschbohnen_mittelfein_2021.pdf (20.02.2024)

3.2.2.4 Erstellung der Bildaufnahmen

Für die Bildgebung der Bohnenhülsen wurde eine maßgeschneiderte Fotobox mit diffus reflektierenden Wänden aus Aluminium verwendet, um eine möglichst ausgeglichene Ausleuchtung der Grundplatte zu ermöglichen (Abbildung 11). Die Oberseite der Fotobox bildet ein LED-Panel mit integrierter Kameraöffnung in der Mitte (*Effilux*, Hürth, Deutschland). Das Panel besitzt eine Beleuchtungsstärke von 40000lx (gemessen bei 0cm Abstand). Die Grundplatte bedeckte eine Fläche von 50 x 38cm und besteht aus einer robusten blauen Polypropylenplatte (*Item*, Solingen, Deutschland), deren Farbe und geringe Reflexionseigenschaften die Erkennung der Hülsen im Bild erleichterten. Der Abstand zwischen Kameraobjektiv und der Grundplatte beträgt 80cm. Die Bilder wurden mit einer industriellen RGB-Kamera Manta G-507 (*Allied Vision*, Stadtroda, Deutschland, 5MP) aufgenommen, die mit einem 12mm FL-CC1218-5MX Objektiv (*Ricoh*, Tokyo, Japan) ausgestattet war. Die Kamera wurde nach Z. Zhang (2000) kalibriert, um die Auswirkungen von radialen Verzerrungen zu prüfen und die Pixelmetrik ins metrische System umzurechnen. Bei einem Abstand von 80 cm wurden keine nennenswerten Verzerrungen festgestellt. Die Bedienung der Kamera erfolgte über die *Vimba Viewer 5.1* Software (*Allied Vision*, Stadtroda, Deutschland). Die Belichtungszeit wurde auf 1,4ms mit einem Gain von 0 und einem Gamma-Wert von 1 eingestellt. Diese Einstellungen minimierten das Bildrauschen und sorgten für ausgeglichene helle Bilder (Abbildung 11). Der Weißabgleich wurde mit dem Farbkalibrierungs-Target *ColorChecker Passport Photo 2* (*Xrite*, Planegg-Martinsried, Deutschland) durchgeführt, wobei für den roten und den blauen Kanal Werte von 2,4 und 2,0 festgelegt wurden. Der *ColorChecker* wurde außerdem eingesetzt, um mögliche Beleuchtungsänderungen während der Datenaufzeichnung zu erkennen, welche jedoch nicht auftraten. Die Kamera wurde weiterhin so eingestellt, dass sie 8-Bit tiefe unkomprimierte TIFF-Bilder erstellte.

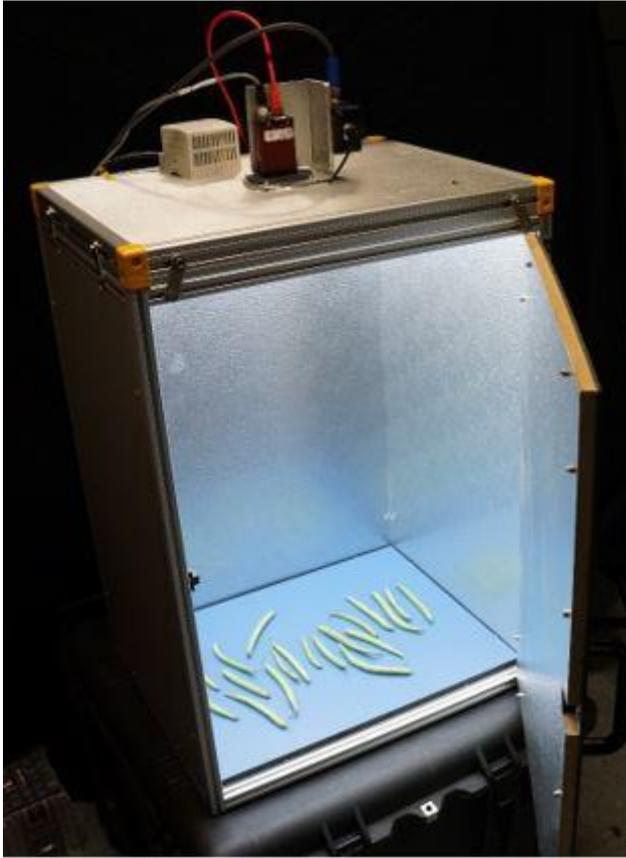


Abbildung 11: Diffus beleuchtete Aufnahmebox mit einer im LED-Panel verbauten Industriekamera, welche über einen externen Rechner gesteuert wird; Bei geschlossener Tür können bis zu 30 Bohnenhülsen in einer Aufnahme erfasst werden (Jollet et al., 2021).

3.2.2.5 Vorverarbeitung der 2D RGB-Bilder

Sämtliche Methoden der Bildverarbeitung und Analysen in dieser Arbeit wurden unter Windows 10 mit *Python 3.9*⁴ unter Nutzung von *Numpy 1.20.2* (Harris et al., 2020), *OpenCV 4.5.1.48* (Bradski, 2000), *scikit-learn 1.0.1* (Pedregosa et al., 2012), *PlantCV 3.13.2* (Gehan et al., 2017), *SciPy 1.8.0* (Virtanen et al., 2020), *Pandas 1.3.5* (McKinney, 2010) und *Matplotlib 3.5.1* (Hunter, 2007) durchgeführt. Für Ansätze des tiefen Lernens nutze man das *Detectron2 0.2.1* Framework (Wu et al., 2019) von *Meta AI* (früher *Facebook AI*, Menlo Park, USA), sowie *Pytorch 1.8.0* (Paszke et al., 2019).

Zunächst wurden die Bilder, welche die Kamera montagebedingt generierte, um 1° mit dem Uhrzeigersinn gedreht, um die Kanten der blauen Grundplatte parallel zur Bildkante auszurichten. Dies erleichterte das anschließende Zurechtschneiden des Bildes. Hierbei wurden alle Bereiche, welche außerhalb der Grundplatte lagen, entfernt, da diese Bildinformationen irrelevant für die weitere Verarbeitung waren. Da trotz optimiertem Aufnahme-Setting geringe Helligkeitsverläufe noch vorhanden waren, wurden die Beleuchtungsgradienten im Bild mit einer 2D Gauß-Verteilungsfunktion zusätzlich korrigiert. Für eine robuste Hülsensegmentierung wurde, falls im Bild vorhanden, zunächst der *ColorChecker* ausmaskiert, da ansonsten dessen grüne oder gelbe Referenzfarben als Hülsen interpretiert würden (Abbildung 12). Dies funktionierte in diesem Fall durch Ausmaskierung der größten Kontur im Bild. Zur Segmentierung des Bildes wurde das RGB-Bild in den HSV-Farbraum umgewandelt. Blaue Hintergrundpixel wurden mittels Schwellenwertsegmentierung im Farbtonwert (H-Kanal) getrennt, wobei ein Schwellenwert von 160 (bezogen auf die 8bit

⁴ Python; Python Software Foundation; Python Language Reference; Version 3.9; <https://www.python.org/>

Bildskala mit Werten zwischen 0 und 255; gilt auch für nachfolgende Schwellenwert-Angaben) gewählt wurde, da dieser in allen Fällen zu einer robusten Trennung von Vordergrund und Hintergrund führte. Pixel mit Werten unterhalb des Schwellenwerts gehörten zum Vordergrund, Pixel mit einem Wert darüber zum Hintergrund.

Die resultierenden Konturen wurden der Größe nach sortiert, wobei deren Fläche wie in Abschnitt 2.3.6 beschrieben bestimmt wurde (Riemann, 2013). Da der ColorChecker die größte Kontur im Bild darstellte, konnte dieser sehr einfach extrahiert werden. Durch Berechnung der maximalen Ausdehnung der Kontur in x- und y-Richtung, sowie der Berechnung des geometrischen Schwerpunkts der Kontur, wurde diese mit einem blauen Rechteck ausmaskiert, welches die Mittlere Farbe des Hintergrunds aufwies. Die Konturen der Hülsen wurden anschließend automatisch in Bildausschnitte aufgeteilt, die jeweils eine einzelne vollständige Bohnenhülse enthielten. Waren in diesen Bildausschnitten noch Fragmente benachbarter Hülsen enthalten, wurden diese ausmaskiert. Diese Fragmente definierten sich dadurch, dass sie im Gegensatz zur zentralen und damit vollständigen Hülse den Rand des Bildausschnittes berührten. Jede Hülse konnte nun einzeln analysiert werden.

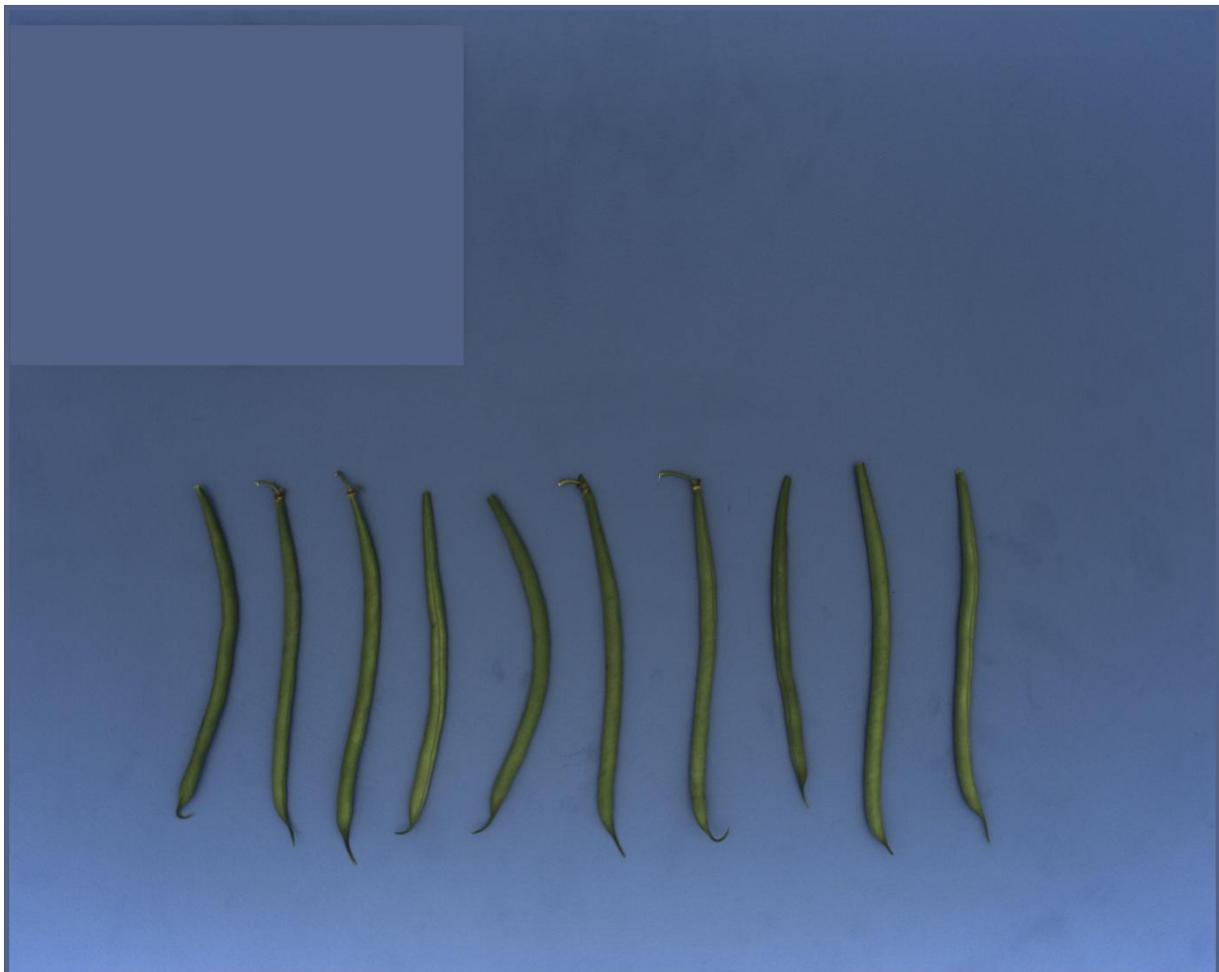


Abbildung 12: Vorverarbeitung der Bohnenbilder, welche in einer für diesen Zweck entwickelten Fotobox aufgenommen werden. Zunächst werden die Bilder gerade ausgerichtet und zugeschnitten und, falls vorhanden, der ColorChecker aus dem Bild entfernt und die Helligkeitsverteilung mittels einer Gaussfunktion optimiert.

3.2.2.6 Formparameteranalyse durch maschinelles Sehen

3.2.2.6.1 Detektion von Spitzen und Blütenansätzen

Da Blütenansätze und Spitzen der Hülsen nicht in weitere Analysen, wie beispielsweise die Länge oder Krümmung, einfließen sollten, mussten diese Strukturen in den Bildausschnitten erkannt und ausmaskiert werden. Hierfür wurde ein Deep Learning Ansatz gewählt, und zwar

ein *mask region based convolutional neural network (Mask R-CNN)* aus dem *Detectron2*-Framework. Dieses neuronale Faltungsnetzwerk gibt es mittlerweile in vielen unterschiedlichen Versionen und stellt seit seiner Veröffentlichung im Jahr 2019 den Stand der Technik im Bereich der Instanz-Segmentierung von Bildern dar (Hassan et al., 2022). Die genutzte Version bestand aus einem *Feature-Pyramid ResNet-Backbone* mit 101 Schichten und einer 3x Lernrate. Letzteres bedeutet, dass während der Aufwärmphase des Trainings eine 3-fach höhere Lernrate im Vergleich zur Standardlernrate genutzt wird, um das Training zu beschleunigen. Für das Training wurde in dieser Arbeit die Methode des Transferlernens genutzt, indem eine Version des Netzwerks genutzt wurde, welche auf dem Image-Net Datensatz (Deng et al., 2010) vortrainiert wurde. Das Netzwerk wurde somit mit den *Model Zoo* Gewichten des *Detectron2*-Frameworks initialisiert⁵. Das Transferlernen reduziert vor Allem die Menge der benötigten Trainingsdaten und somit auch die Trainingszeit sowie die benötigte Rechenleistung für das Training.

Training: Für das Training des *Mask R-CNN* wurden insgesamt 380 Bildausschnitte von Einzelhülsen aus dem oben beschriebenen *Entwicklungsdatensatz*, sowie zu kleinen Teilen auch aus anderer Quelle genutzt. So ergab sich im Zuge dieser Arbeit die Möglichkeit das *Sächsische Landesamt für Umwelt, Landwirtschaft und Geologie Abteilung Gartenbau* in Dresden-Pillnitz (Ostdeutschland) zu besuchen, wo Bildmaterial von Bohnenhülsen gesammelt werden konnte. Während des Besuchs fand eine Versuchsauswertung für mittelfeine (Kaliber von 9-10,5mm) Buschbohnen für die Tiefkühlindustrie (runder Querschnitt und grüne Farbe) statt. Bei der Auswahl der Trainingsbilder (Bildausschnitte einzelner Hülsen) aus den genannten Quellen wurde besonders darauf geachtet, dass ein breites Spektrum an Phänotypen enthalten war, was sich sowohl auf Eigenschaften der Hülsen bezüglich Form und Farbe bezieht, aber auch auf das Vorhandensein und Erscheinungsbild der Spitzen und Blütenansätze. Bei der Zusammenstellung der Trainingsbilder mit Blütenansätzen wurde besonders auf vielfältige Ausprägungen geachtet, da ein Blütenansatz mit Blütenstiel und Blütenblättern, oder nur mit einem von beidem oder mit keinem von beiden auftreten kann. In jedem der 380 Bilder wurden, falls vorhanden, die Spitze sowie der Blütenansatz mit Hilfe des *VGG Image Annotator* (Dutta & Zisserman, 2019) maskiert und klassifiziert. Die Bilder wurden in 295 Trainings- und 85 Testbilder aufgeteilt, was einen Trainingsdatensatz mit 293 annotierten Spitzen und 164 Blütenansätzen und einen Testdatensatz mit 83 annotierten Spitzen und 47 Blütenansätzen ergab. Für das Training wurden die Standardeinstellung des *Detectron2*-Frameworks für die meisten Hyperparameter genutzt, außer dass die Lernrate auf 0,00025 angepasst wurde, um Over-Fitting zu vermeiden. Um die Vorhersagen des trainierten Netzwerks robuster zumachen, wurden Augmentationen an den Trainingsdaten vorgenommen. Hierzu wurde zunächst die kürzeste Kante der Bildausschnitte auf 100 Pixel festgelegt, wobei das Seitenverhältnis der Bilder unverändert blieb. Durch diesen Schnitt wurden die unregelmäßigen Bildausschnitte aus der Vorverarbeitung für das Netzwerk genormt. Die Bilder wurden zufällig gespiegelt und zufällig um 1 bis 360 Grad gedreht, um alle möglichen Ausrichtungen der Bohnenhülsen zu simulieren. Sättigung, Helligkeit und Kontrast wurden nach dem Zufallsprinzip auf einen Faktor zwischen 0,75 und 1,25 eingestellt. Eine Evaluierung des Trainingserfolgs mit dem Testdatensatz wurde alle 1000 Epochen durchgeführt, wobei die Gewichte des Netzwerks gespeichert wurden, wenn die Verlustfunktion für den Testdatensatz ein neues Minimum erreichte. Das absolute Minimum wurde nach 35000 Epochen erreicht.

Validierung: Zur Validierung der Vorhersagen des trainierten *Mask R-CNN* wurden die 400 Bilder von Einzelbohnen des oben beschriebenen *Validierungsdatensatz* genutzt, für welche auch züchterische Referenzmessungen vorlagen. Für die Vorhersagen des Netzwerks wurde

⁵ Detectron2 – Model Zoo:

https://github.com/facebookresearch/detectron2/tree/main/detectron2/model_zoo (20.02.2024)

eine Konfidenzschwelle von 0,5 verwendet, was der Standardeinstellung entspricht. Grundsätzlich macht das *Mask R-CNN* 3 Vorhersagen: Klasse, Instanz-Segmentierung und umschließendes Rechteck. Die vorhergesagten Klassen wurden dazu genutzt, um festzustellen, ob die untersuchte Hülse eine Spitze und/oder einen Blütenansatz besaß. Für jede vorhergesagte Klasse gibt das Netzwerk weiterhin einen Vertrauenswert zwischen 0 und 1 an, also eine Wahrscheinlichkeit, mit der diese Klasse zutrifft. Die vorhergesagten umschließenden Rechtecke fanden keine weitere Verwendung. Die vorhergesagten Masken wurden dazu genutzt, die gegebenenfalls vorhandene Spitze oder den Blütenansatz zu maskieren und für die weitere Analyse auszuschließen.

3.2.2.6.2 Länge der Hülsen

Um die Länge der Hülse abzuschätzen, berechnete man die Länge der zentralen Linie der Hülse im Bild. Um diese zu erhalten, wurden folgende Verarbeitungsschritte an den maskierten binären Bildausschnitten der Einzelhülsen (Abbildung 13) vorgenommen: Zunächst wurde die Maske der Hülse nach Zhang & Suen (1984) skelettonisiert (Abbildung 13), was die zentrale Linie der Hülse inklusive einiger abzweigender kürzerer Linien (auch als Äste bezeichnet) generierte. Diese Äste wurden durch einen iterativen Zuschnitt entfernt. Dazu wurden zunächst alle Verästelungspunkte des Skelettes, sowie dessen Spitzen ausfindig gemacht (Gehan et al., 2017). Mit Hilfe dieser Informationen wurde für jeden Ast des Skeletts die Länge durch Zählen der zugehörigen Pixel bestimmt und in einem iterativen Prozess so lange jeweils der kleinste Ast entfernt, bis keine Verästelungspunkte mehr vorhanden waren (Abbildung 13). Der iterative Ansatz hat den Vorteil, dass bei verzweigten kurzen Enden der längste Ast schließlich dem zentralen Skelett zugeordnet wird und dadurch erhalten bleibt. Zuletzt wurden jene Pixel der Zentrallinie, welche der Spitze oder dem Blütenansatz der Hülse zugeordnet waren, mittels Subtraktion entfernt (Abbildung 13). Die Länge der resultierenden Zentrallinie wurde schließlich mit folgender Berechnung angenähert: Die Länge ergibt sich aus der Summe der Abstände benachbarter Pixel, wobei horizontale und vertikale Pixelnachbarn einen Abstand von 1 und schräg zueinander liegende Pixel nach dem Satz des Pythagoras einen Abstand von $\sqrt{2}$ besitzen (Bradski, 2000).

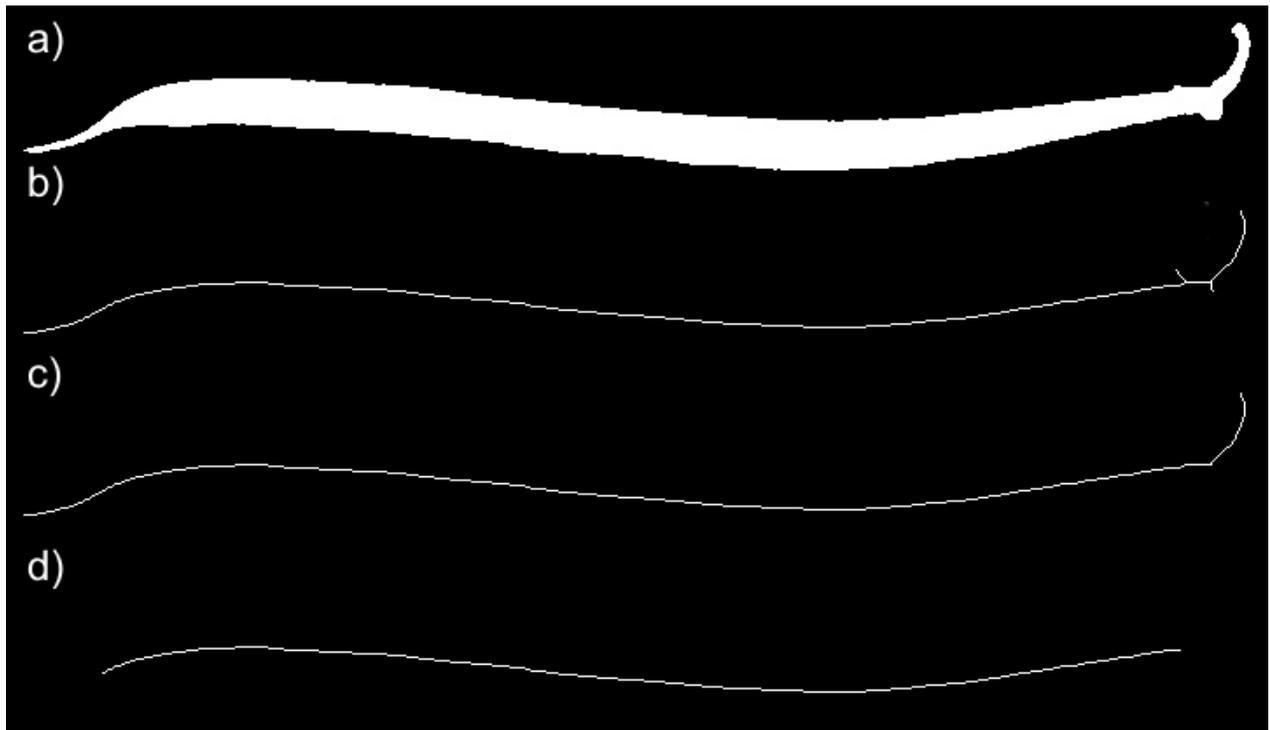


Abbildung 13: Generierung der Zentrallinie aus der a) Bildmaske einer Hülse; b) Skelett nach Zhang & Suen (1984); c) Zentrallinie nach iterativer Entfernung von Ästen; d) Zentrallinie nach Subtraktion der Spitze und des Blütenansatzes.

3.2.2.6.3 Kaliber der Hülsen

Um das Kaliber der Hülse zu bestimmen, galt es deren dickste Stelle zu lokalisieren, was mit Hilfe der Zentrallinie sowie der Maske der Hülse realisiert wurde. Dazu wurde an jedem Pixel der Zentrallinie und unter Einbezug der 20 Nachbarpixel (10 zu jeder Seite) eine orthogonale Gerade gespannt, welche ihre beiden Endpunkte an den Konturpunkten der Maske hatte, auf welche die virtuelle Gerade stieß. Die euklidische Distanz zwischen diesen zwei gegenüberliegenden Endpunkten entsprach der Breite der Hülse am entsprechenden Punkt der Zentrallinie. So konnte für die gesamte Hülse ein Breitenprofil erstellt werden, aus dem sich das Kaliber als Maximum dieses Profils ableiten ließ. Für die ersten und letzten 10 Pixel der Zentrallinie wurde keine orthogonale Gerade modelliert, da diese Stellen für die Kaliber Messung irrelevant waren, da sich die Hülsen zu den beiden Enden hin stets verjüngen.

3.2.2.6.4 Krümmung der Hülsen

Zur Berechnung der Krümmung nutzte man die Zentrallinie der Hülse und machte sich die Eigenschaft zu Nutze, dass deren Verlauf in fast allen Fällen als mathematische Funktion approximierbar ist, wobei die Krümmung aus der 2. Ableitung dieser Funktion berechnet werden kann. Hierfür müssen aber nachfolgende Voraussetzungen geschaffen werden: Zunächst wurde die Zentrallinie so weit rotiert, bis sie eine maximale Ausdehnung in X-Richtung erreicht. Dadurch sollte sichergestellt werden, dass es zu jeder X-Koordinate der Zentrallinie nur eine Y-Koordinate gab, was für fast alle Hülsen mit Ausnahme der Krümmungsklasse 9 galt, da sie sehr stark gekrümmt waren. Anschließend wurde der Verlauf der Zentrallinie mit Hilfe einer polynomischen Funktion 5. Grades in der Form

$$f(x) = a * x^5 + b * x^4 + c * x^3 + d * x^2 + e * x + f \quad [13]$$

modelliert. Die 2. Ableitung dieser Funktion ergab nun die Krümmung für jeden Pixel der Zentrallinie. Da die züchterische Bewertung Krümmungen in der Mitte der Hülse stärker gewichtet wie an dessen Enden (welche manchmal eingerollt sind), wurde die Krümmungsfunktion mit einer doppel-sigmoidalen Funktion multipliziert (Abbildung 14). Die Gewichtungsfunktion besaß somit in der Mitte einen Wert von 1, an den Enden Werte von 0, und die Wendepunkte befanden sich etwa bei 1/12 und 11/12 der X-Achsenausdehnung der Funktion. Diese Positionen wurden empirisch ermittelt, d.h. dass die resultierende Gewichtungsfunktion zu einer besseren Übereinstimmung mit den Referenzwerten führte. Die Krümmung der gesamten Hülse ergab sich nun aus dem Mittel aller gewichteten Krümmungswerte der Zentrallinie.

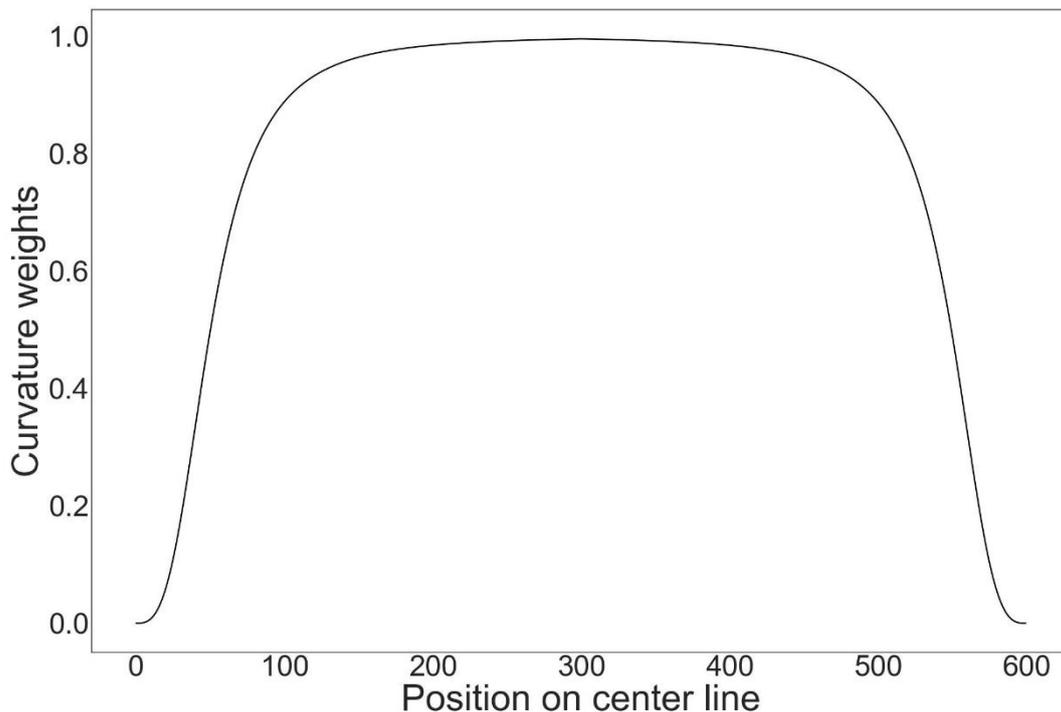


Abbildung 14: Gewichtungsfunktion zur Berechnung der Krümmung für eine hypothetische Bohnenhülse mit einer Länge der Zentrallinie von 600 Pixeln. Sie dient dazu die Computerberechnungen der Krümmung an die Züchterbewertung anzugleichen, welche Krümmungen an den Enden der Hülse weniger stark bewerten

3.2.3 Ergebnisse

3.2.3.1 Erkennung der Spitzen und Blütenansätze

Ein trainiertes *Mask R-CNN* Modell wurde zur Erkennung von Spitzen und Blütenansätzen an Bohnenhülsen verwendet, um diese Komponenten von der weiteren Analyse der Länge, des Kalibers und der Krümmung auszuschließen. Das Netzwerk erkannte die meisten Blütenansätze korrekt mit einer Genauigkeit von 95,5% (Abbildung 15a, 2c I-VI). 15 Blütenansätze wurden nicht erkannt (Abbildung 15c VII-IX). In 3 Fällen klassifizierte das Netzwerk Spitzen fälschlicherweise als Blütenansätze (Abbildung 15c X, XI). Spitzen wurden ebenfalls mit einer Genauigkeit von 95,5% erkannt (Abbildung 15b, 15c XII-XVII). In 17 Fällen war das Modell nicht in der Lage, die Spitze zu erkennen (Abbildung 15c XVIII,XX). In zwei Fällen wurden 2 Spitzen an einer einzigen Bohnenhülse detektiert. Entweder, weil ein Blütenansatz als Spitze erkannt wurde (Abbildung 15c XXI) oder ein abstehender Faden in der Mitte der Bohnenhülse (Abbildung 15c XXII). In einem Fall erkannte das Netzwerk einen stark markierenden Samen in der Mitte der Hülse als Spitze (Abbildung 15c XXIII). Diese Hülse

wurde von der weiteren Analyse ausgeschlossen. Für den Testdatensatz wurde während des Trainings ein mAP_{50} von 92,3% erzielt.

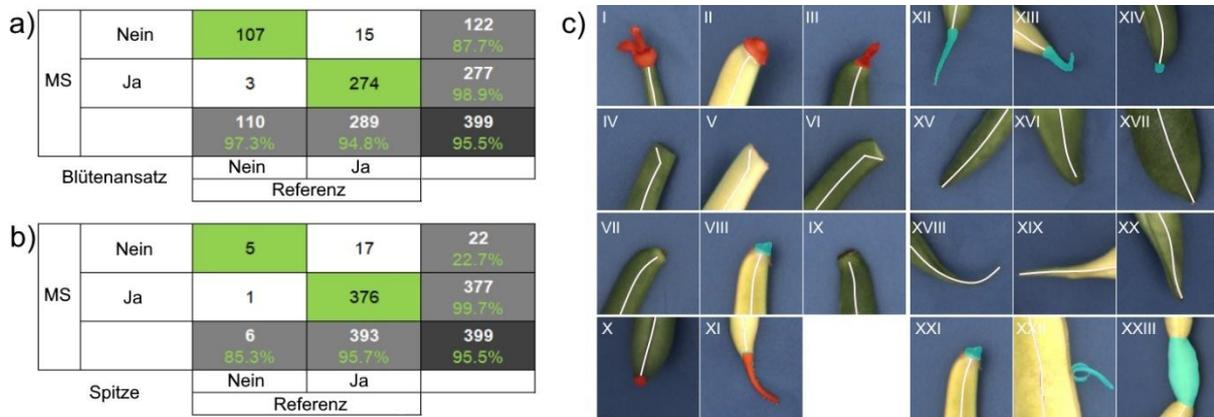


Abbildung 15: Wahrheitsmatrizen zur a) Erkennung der Blütenansätze und b) Erkennung der Spitzen durch ein trainiertes Mask R-CNN; c) Beispiele der Detektion bzw. Segmentierung durch das Mask R-CNNs für I-III: Richtig-Positive Blütenansatzerkennung (rote Maske), IV-VI: Richtig-Negative Blütenansatzerkennung, VII-IX: Falsch-Negative Blütenansatzerkennung, X-XI: Falsch-Positive Blütenansatzerkennung, XII-XIV: Richtig-Positive Spitzenerkennung (cyane Maske), XV-XVII: Richtig-Negative Spitzenerkennung, XVIII-XX: Falsch-Negative Spitzenerkennung, XXI-XXIII: Falsch-Positive Spitzenerkennung (Jollet et al., 2023).

3.2.3.2 Berechnung der Länge, des Kalibers und der Krümmung

Für die Analyse der Länge und des Kalibers wurden bestimmte Ausreißer (wie in den nachfolgenden Abschnitten beschrieben) manuell ausgeschlossen (Abbildung 16a,b graue Punkte), welche aus Limitierungen des 2-dimensionalen bildgebenen Verfahrens oder aus Fehlern vorausgegangener Bildanalyseverfahren resultierten. Die Funktionalität der genutzten Verfahren wurde durch einen Vergleich der Absolutwerte überprüft, also wie nahe der Fit an der Identitätslinie liegt. Da die Ausreißer sehr systematisch verteilt sind, würden sie den Fit verschieben und dadurch die Bewertung der Funktionalität der genutzten Verfahren erschweren.

Länge: Die Hüslenlängen, welche mittels maschinellen Sehens (MS) berechnet wurden, zeigen einen stark linearen Zusammenhang mit den mit einem Lineal durchgeführten Referenzmessungen (Abbildung 16a). Insgesamt wurden 17 Bohnenhülsen von 399, bei denen die Spitze nicht erkannt wurde, von der Längenanalyse ausgeschlossen (Abbildung 16a graue Punkte). Die nicht erkannten Blütenansätze wurden nicht von der Längenanalyse ausgeschlossen, da es sich hierbei ausschließlich um Blütenansätze ohne Stiel und Blütenblätter handelte. Dadurch beeinflussten diese die Längenmessung kaum. Die Regression zeigt einen MAPE von 1,7 % und einen Korrelationskoeffizienten von $R=0,99$ (Abbildung 16a). Um potenzielle systematische Größeneffekte auf die MS-Längenanalyse zu bewerten, wurden die Residuen der linearen Anpassung von drei Gruppen von Bohnenhülsen mittels einfaktorieller ANOVA verglichen. Die Gruppen wurden gemäß Jusoh (2017) in kurz ($<120\text{mm}$), mittel ($\geq 120\text{mm}$, $\leq 150\text{mm}$) und lang ($>150\text{mm}$) unterteilt. Ein p-Wert von 1 zeigte keine statistisch signifikanten Unterschiede zwischen den Längengruppen und damit keinen Einfluss der Länge auf die Berechnung an.

Kaliber: Das Kaliber wurde aus 399 Einzelhülsenbildern berechnet und mit den Referenzmessungen verglichen, die mit einer Lochlehre vorgenommen wurden (Abbildung 16b). Bohnenhülsen, welche der Kamera ihre schmale gekrümmte Seite zeigten (6,6 % aller Bohnenhülsen, nur ovale und flache), wurden manuell von der Regressionsanalyse ausgeschlossen. Die lineare Regression ergab einen MAPE von 3,3% und einen

Korrelationskoeffizienten von $R=0,99$. Statistisch signifikante Unterschiede wurden für die Residuen der Regression zwischen allen Bohnenhülsestypen (*Rund*, *Oval*, *Flach*) gemessen ($p=0,0008$). Der höchste Cohen's D – Wert von 0,44 wurde zwischen den Residuen der Bohnentypen *Rund* und *Flach* gemessen, was immer noch auf eine geringe Effektgröße hinweist.

Krümmung: Die MS-Schätzungen der Krümmung wurden mit den visuellen Bewertungen (nach Züchterstandard) verglichen. Ein Merkmal der Züchterskala war, dass mit zunehmender Krümmung die Unterschiede in den Krümmungen, also die Abstände zwischen benachbarten Boniturwerten immer größer wurden. Die Züchterskala ist also nicht-linear. Dies ist aus Züchtersicht erforderlich, da insbesondere bei den eher geraden Bohnen der Bewertungsklassen 1-4 noch eine gute Differenzierbarkeit über den Boniturwert gegeben sein muss, da neue Bohnensorten diesen Klassen entsprechen sollten. Daher wurde der Vergleich zwischen den MS-Messungen und den Ergebnissen der Referenzmessungen auf der Basis der log-transformierten MS-Ergebnisse durchgeführt, da sich hierdurch ein linearer Zusammenhang ergab. Für die Bewertungsklassen 1-4 wurde eine größere Variation beobachtet (Abbildung 16c). Dies spiegelt sich auch in einem MAPE von 13,8 % ($R=0,81$) wider. Eine gute Übereinstimmung mit der Regression ist auch bei der Betrachtung der einzelnen Bohnenhülsekkategorien zu beobachten (Abbildung 16c) und es wurden keine statistisch signifikanten Unterschiede zwischen den Residuen der Typen *Rund*, *Oval* und *Flach* gemessen (p -Wert = 2,09). Lediglich flache Bohnenhülsen der Wertungsklasse 4 wurden leicht unterschätzt.

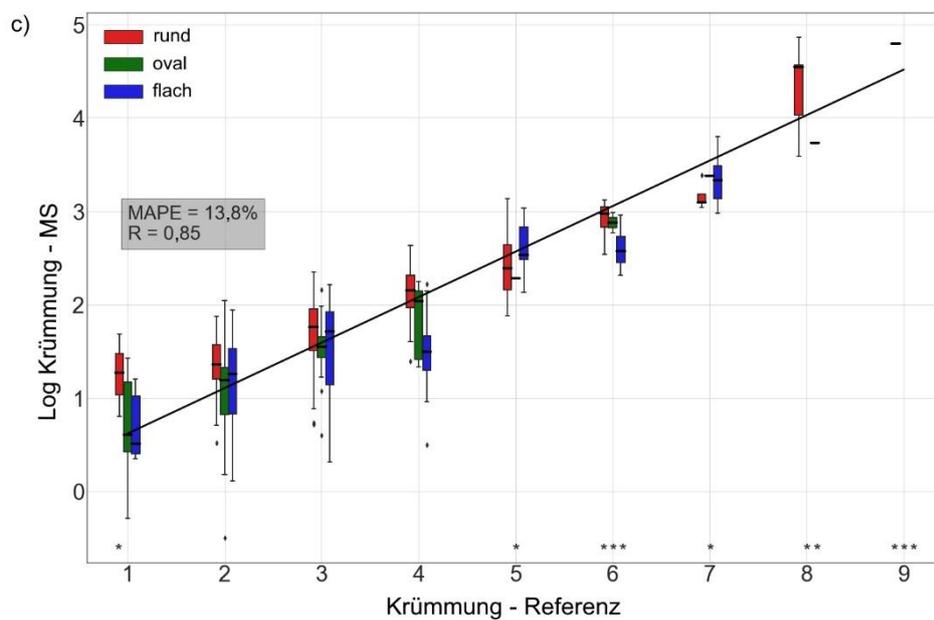
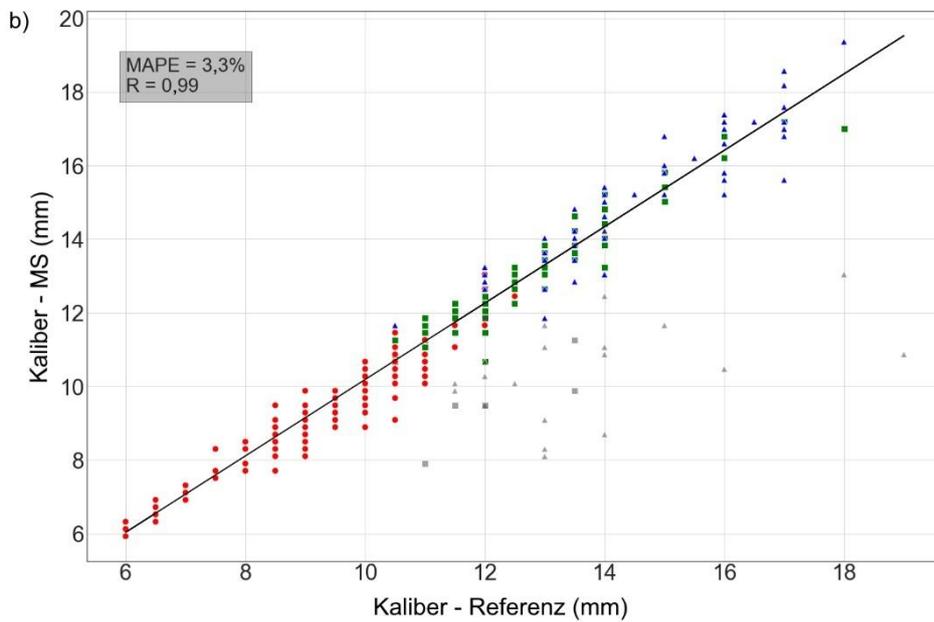
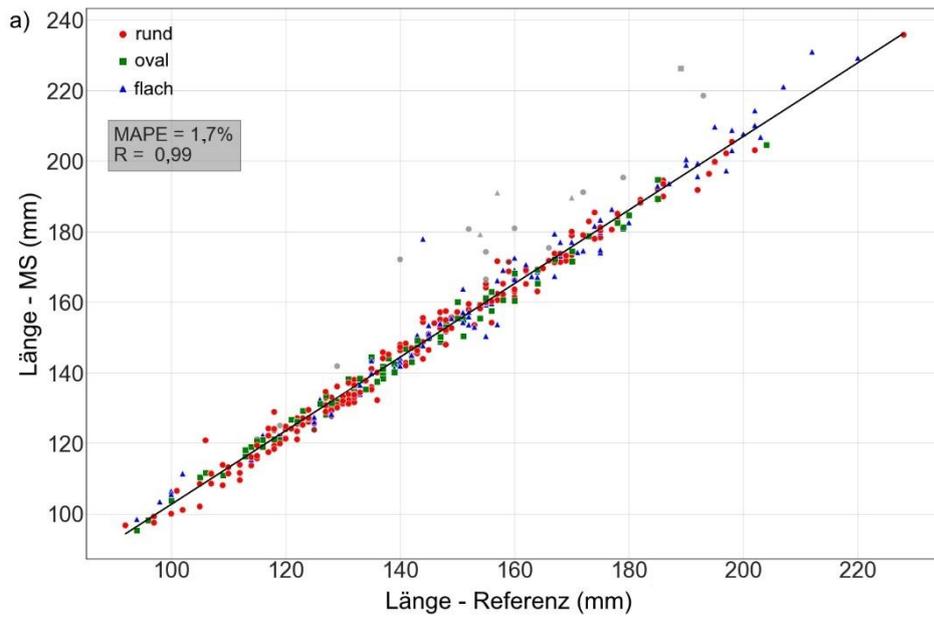


Abbildung 16: Vergleich zwischen den Referenzmessungen und den Messungen der MS-Methoden für a) die Hülsenlänge, b) das Kaliber und c) die Hülsenkrümmung (n=399). Der mittlere absolute Fehler in Prozent (MAPE) sowie der Korrelationskoeffizient (R) sind für die drei Merkmale angegeben. Ausreißer (graue Punkte) zeigen für a) Falsch-Negative Spitzenerkennungen des Mask R-CNN (n=17) und für b) Hülsen der Bohnentypen Oval und Flach, welche auf Grund der Krümmung ihre flache Seite zur Kamera zeigen (n=26). Ausreißer wurden von der Berechnung der Regression ausgeschlossen. Sterne in c) zeigen Box-Plots mit geringer Stückzahl (n<5) (Jollet et al., 2023).

3.2.4 Diskussion

Die Ergebnisse der MS-Methoden wurden für alle untersuchten Merkmale mit den Bewertungen von Zuchtexterten verglichen. Um die Grenzen der entwickelten MS-Methoden zu testen, wurde eine breite Palette von Bohnensorten mit einer großen Variation in den untersuchten Merkmalen analysiert. Ein zentrales Element der entwickelten Methodik ist die abgeschlossene Aufnahmekammer. Für vergleichbare Kulturarten sind zwar auch Methoden entwickelt worden, bei denen die Datenaufnahme unter den natürlichen Feld- und Lichtbedingungen erfolgte, z.B. für die Quantifizierung der Hülsenzahl und die Erstellung von Ertragsprognosen von Sojabohnen (X. Fu et al., 2022; Riera et al., 2020), allerdings ermöglicht eine vollständig steuerbare Aufnahmebox eine präzise Bewertung von deutlich mehr züchtungsrelevanten Merkmalen.

Eine Herausforderung, die sich für die Forschungsfrage aus dem Kamera-basierten Ansatz ergibt, stellt die Limitierung auf einen Aufnahme-Blickwinkel dar, durch die manche Hülsen in der 2-dimensionalen Darstellung nicht ausreichend für alle Analysen repräsentiert sind. Zwar können Bohnenhülsen bei der manuellen Beurteilung von allen Seiten betrachtet werden, aber nur durch automatisierbare Verfahren wie das hier vorgestellte bildgebende System ist eine ausreichende Durchsatzsteigerung möglich, welche im züchterischen Kontext von großer Relevanz ist. Siehe z.B. Baek et al. (2020), welche ein bildgebendes Hochdurchsatzsystem zur Bewertung züchtungsrelevanter Formparameter an Sojabohnensamen entwickelten.

Die Erkennung von Blütenansätzen und Spitzen an den Hülsen ist in zweierlei Hinsicht von Bedeutung: Zum einen ist die Erkennung der Blütenansätze ein wichtiger ernterelevanter Parameter, da dadurch eine Aussage über die industrielle Verarbeitbarkeit einer Sorte gemacht werden kann. Hülsen, welche von einer Erntemaschine ohne den Blütenansatz geerntet werden können, sind in der industriellen Verarbeitung leichter zu prozessieren. Weiterhin ist die Erkennung der Blütenansätze und Spitzen im Bild ein wichtiger Vorverarbeitungsschritt, um korrekte Messungen an der Hülse vornehmen zu können. Verfahren zur automatischen Erkennung des Blütenansatzes gibt es im Bereich der Gartenbohne - nach aktuellem Kenntnisstand des Autors zum Zeitpunkt der Anfertigung dieser Arbeit - nicht. Das im Vorfeld dieser Arbeit entwickelte Random-Forest (RF)-Verfahren (Jollet et al., 2021) wurde ausschließlich auf runden Hülsen trainiert, sodass es bei breiteren Hülsen vom Typ *Oval* und *Flach* häufiger zu Fehlklassifikationen kam. Weiterhin wurde der Ansatz der Blüten häufig nicht richtig lokalisiert, da die in dieser Veröffentlichung genutzten klassischen Bildverarbeitungsmethoden nicht in der Lage waren die vielfältigen Erscheinungsformen der Blütenansätze abzudecken (kurzer/langer/kein Stiel, ein/zwei/kein Blütenblatt, verschiedenste Lagebeziehungen zwischen Stiel und Blütenblätter). Das in dieser Arbeit trainierte CNN-Modell konnte die Erkennungsgenauigkeit trotz der größeren Datenvariation verbessern. Die Ergebnisse zeigten, dass Spitzen und Blütenansätze mittels eines trainierten *Mask R-CNNs* besser klassifiziert und lokalisiert werden konnten. Nur in einigen wenigen Fällen erkannte das neuronale Netz die Blütenansätze vor allem aufgrund ihrer Größe nicht (5,2%), d. h., es gab in diesen Fällen nur sehr kurze Stiele und es waren keine oder kaum Blütenblätter sichtbar. Wurden solch kurze Blütenansätze nicht erkannt, beeinflusste dies die Längenschätzung jedoch kaum. Bei 4 Hülsen wurden zwei Spitzen erkannt. In diesen Fällen kann eine korrekte Zuordnung auf der Grundlage der Vertrauenswerte der Vorhersagen genutzt werden. In den

genannten 4 Fällen wäre diese erfolgreich gewesen. Erkannte Spitzen und Blütenansätze wurden sehr genau maskiert, was sich auch in genauen Längenschätzungen widerspiegelte.

Im Gegensatz zu nicht detektierten Blütenansätzen beeinflussten nicht detektierte Spitzen die Längenmessung deutlich stärker. Die Ausreißer, die durch diesen Fehler erzeugt wurden, wurden jedoch ausgeklammert, um konkrete Aussagen über die Funktionalität der Längenmessung, sowie über die Maskierungsgenauigkeit des *Mask R-CNNs* machen zu können. Kılıç et al. (2007) bestimmten die Länge und Breite von Bohnensamen, indem sie ein umschließendes Rechteck um die Kontur des Samens legten und anschließend die Länge und Breite dieses Rechtecks bestimmten. Mit dieser Technik erzielten sie für die Längenbestimmung einen hohen Korrelationskoeffizienten von 0,98. Da es sich bei den Hülsen jedoch um längliche Strukturen handelt, welche nicht nur alle Variationen zwischen gerade gestreckt und sichelförmig aufweisen, sondern auch s-förmig gekrümmt sein können, erfolgte die Längenmessung in dieser Studie anhand der Zentrallinie der Hülsen. Van Eck et al. (1998) führten ebenfalls bereits erfolgreich Längenmessungen anhand der Zentrallinie an leicht gekrümmten Gurken (*Cucumis sativus*) durch und erzielten einen Korrelationskoeffizienten von 0,99 zwischen manuellen Referenzmessungen und ihrer MS-Methode. Mit der hier vorgestellten Methode konnte, trotz der teils starken Krümmungen der Hülsen, ebenfalls ein Korrelationskoeffizient von 0,99, sowie ein sehr geringer MAPE von 1,7% erzielt werden. Bei einer 15cm langen Hülse entspräche dies einer Abweichung im Mittel von 2,6mm. Es konnten somit sehr präzise und robuste Ergebnisse erzielt werden, was zusätzlich auf eine hohe Maskierungsgenauigkeit des trainierten *mask R-CNNs* hindeutet.

Die Funktionalität MS-Kalibermessung nach Jollet et al. (2021) wurde in dieser Arbeit, durch die deutliche höhere Variation der Phänotypen der untersuchten Hülsen, erneut auf den Prüfstand gestellt. Die Kalibermessung funktionierte für runde Hülsen, wie bei Jollet et al. (2021) sehr präzise (MAPE = 3,3%, bei einer Hülse mit einem Kaliber von 9mm entspricht dies einer Abweichung von 0,3mm), war jedoch für flache und ovale Hülsen leicht eingeschränkt. Hier zeigte sich eine Limitierung des genutzten Systems, insbesondere bei Betrachtung der Ausreißer beider Typen in den Validierungsergebnissen. Insgesamt 26 der 180 Hülsen von den Typen *Oval* und *Flach* zeigten ihre schmale Seite zur Kamera, sodass für diese Hülsen das Kaliber nicht genau gemessen werden konnte. Trotz der Empfindlichkeit des Verfahrens gegenüber der Lagebeziehung zwischen Hülse und Kamera und dem damit verbundenen Messfehler (in diesen Untersuchungen betraf das ~14% der Hülsen) können durch die Möglichkeit der automatisierten Messung genügend Hülsen einer Sorte analysiert werden, dass verlässliche Aussagen über eine Sorte auf Grundlage der Verteilungsstatistiken getroffen werden können. Durch den Vergleich der einzelnen Messwerte mit dem medianen Kaliber aller Messungen (der Mittelwert ist nicht geeignet, da er zu empfindlich auf Ausreißer reagiert) könnten Ausreißer nachträglich gefiltert werden. Die Ergebnisse dieser Studie deuten darauf hin, dass sich die größten Herausforderungen aus der Limitierung durch nur einen Kamera-Blickwinkel ergeben. Wenn die Bohnehülse abgeflacht (Typ *Oval* oder *Rund*) und über die schmale Seite gekrümmt ist, hängt es von der Ausrichtung zur Kamera ab, ob die Krümmung oder das Kaliber präzise gemessen werden kann. Weiterhin können mit dem 2D RGB-Verfahren keine verdrehten Bohnehülsen mit einer komplexen 3D Krümmung akkurat vermessen werden. Solche Hülsen sind jedoch sehr selten und in Züchtungsprogrammen kaum mehr vorhanden.

In dieser Studie waren wir in der Lage die Vermessung der Krümmung im Vergleich zur Standardpraxis objektiver zu gestalten, was vor allem durch die mathematische Abstraktion der Krümmung als Funktion realisiert wurde. Zum ersten Mal wurde hier eine Methode, welche auf lokalen Ableitungen polynomischer Funktionen basiert, zur Berechnung der Krümmung genutzt. Sie berücksichtigt auch eine heterogene Verteilung der Krümmungen entlang der Hülse bis hin zu Sonderformen wie S-Formen. Diese Entwicklung stellt eine erhebliche

Verbesserung gegenüber früheren Krümmungsmaßen dar, die beispielsweise auf dem Verhältnis der minimalen und maximalen Breite der konvexen Hülle der Mittellinie (Polder et al., 2012) oder dem Längenverhältnis der Zentrallinie zur euklidischen Distanz zwischen Start- und Endpunkt der zentralen Linie (Haque et al., 2021) basierten. Die oben erwähnten perspektivischen Probleme traten auch bei den flachen Bohnenhülsen der Bewertungsklasse 4 auf. Einige Proben wurden so positioniert, dass die Messungen des Kalibers genauer waren als die der Krümmung. Durch eine korrekte Positionierung der Bohnenhülsen kann dieser Verzerrungseffekt jedoch minimiert werden. Weiterhin wurde ein vergleichsweise hoher *MAPE* für das Krümmungsregressionsmodell festgestellt. Dies ist aber nicht weiter verwunderlich, da die *MAPE*-Metrik sehr empfindlich auf Werte nahe 0 reagiert, was z. B. in den Bewertungsklassen 1-4 vorkam. Darüber hinaus könnten diese Abweichungen auch auf Ungenauigkeiten bei der Referenzbewertung zurückzuführen sein.

3.3 Nicht-invasive Analyse der Samenanlage von Leguminosen

3.3.1 Forschungsstand

Nach Kenntnisstand des Autors gibt es aktuell lediglich im Bereich der Sojabohne Forschungsarbeiten, welche die Anzahl der Samen einer Sojabohnenhülse nicht-invasiv zur Ertragsabschätzung untersuchten. Wie zuvor bereits beschrieben nutzten Lu et al. (2022) ein einfaches bildgebendes RGB-Verfahren und kombinierten dieses mit Methoden des tiefen Lernens, um die Hülsen entsprechend ihrer Samenzahl in eine der 5 Kategorien mit Einem, Zwei, Drei, Vier oder Fünf Samen zu klassifizieren. Dieser Ansatz ist jedoch aus zweierlei Gründen nicht auf alle Leguminosen übertragbar: Zunächst markieren Sojabohnen deutlich stärker wie viele andere Leguminosen, wie beispielsweise Gartenbohnen, bei denen die Hülsen oftmals nicht oder nur schwach markieren. Weiterhin gibt es bei Sojabohnen die Eigenschaft, dass sich die Anatomie der Hülse der Anzahl der angelegten Samen anpasst. Eine Sojabohnenhülse mit 2 Samen ist grundsätzlich kürzer wie eine Hülse mit 4 Samen, wobei an einer Pflanze alle Ausprägungen ausgebildet werden können (X. Fu et al., 2022). Gartenbohnen hingegen entwickeln uniforme Hülsen, wobei die Anlage der Samen von genetischen Faktoren und Umwelteinflüssen abhängt. Möchte man also ein Verfahren entwickeln, welches für eine größere Bandbreite an Leguminosen geeignet ist, stellen transmittierende Verfahren, welche die Samen in der Hülse sichtbar machen, eine vielversprechende Lösung dar. Jedoch gibt es bezüglich der Identifikation von Samen innerhalb der Hülse aktuell keine entsprechenden Forschungsarbeiten, weshalb sich an dieser Stelle auf analoge Studien aus anderen Forschungsbereichen bezogen wird. So gibt es einige Studien im Bereich der tomographischen Verfahren, z.B. mittels Magnetresonanz-Bildgebung, um nicht-invasiv in das Innere von Fruchtkörpern ausgewählter Gemüsesorten wie beispielsweise Tomaten (Pagés et al., 2021; S. X. Wang et al., 2017) oder Gurken (Jakusch et al., 2018) zu schauen. Zwar handelt es sich hierbei um ein sehr hochauflösendes verfahren, jedoch ist ein solches Verfahren auch sehr teuer, durch die hochenergetischen Strahlen äußerst aufwendig in der Handhabung und jede einzelne Aufnahme benötigt sehr viel Zeit. Da die Fruchtkörper von Leguminosen im Vergleich zu anderen Kulturarten sehr durchscheinend sind, soll in dieser Arbeit ein Verfahren untersucht werden, welches unkompliziert in der Handhabung ist, einen erhöhten Durchsatz erlaubt und dazu auch noch kostengünstig ist. Laut Kenntnisstand des Autors gibt es zudem bisher noch keine Forschungsarbeiten, welche die Möglichkeiten von simplen LED-Panels in Kombination mit einfachen RGB-Verfahren untersuchten, um die Samenanlagen von Leguminosen durch transmittierendes Licht in einem 2D Bild sichtbar zu machen. Ist dieser Ansatz erfolgreich, kann die Samenanlage mit Hilfe von neuronalen Netzen bewertet werden. Hier hat sich z.B. das *Faster R-CNN* bereits in verschiedenen landwirtschaftlichen Kontexten als sehr effektiv erwiesen hat (L. Fu et al., 2018; Z. Li et al., 2021; Q. Wang & Qi, 2019).

3.3.2 Material und Methoden

3.3.2.1 Modellpflanze

Zur Beantwortung der Forschungsfrage wird die Gartenbohne (*Phaseolus vulgaris*) als Modellpflanze genutzt, da sie Besonderheiten aufweisen, die eine detaillierte Beantwortung zulässt. So passt die Gartenbohne, im Vergleich zur Sojabohne, ihre Hülsenform nicht der Anzahl entwickelter Samen an. Form und Farbe der Hülsen sind sortentypisch ausgeprägt und markieren, besonders bei vermarktbareren Sorten, sehr gering. Diese generellen Eigenschaften, aber auch die Variabilität des Phänotyps unterschiedlicher Gartenbohnenarten sind somit sehr geeignet, um die Funktionalität und Genauigkeit des entwickelten Verfahrens bewerten zu können. Im Zuge dieser Arbeit wurden 2 unterschiedliche Sorten getestet. Eine Sorte bildete gelbe Hülsen, welche stark markierten, dadurch dass die Hülsen dieser Sorte länger reiften. Anhand dieser Sorte ließ sich die generelle Tauglichkeit des Verfahrens bewerten, da gelbe Hülsen besonders durchscheinend sind und große Samen die Detektion im Bild erleichtern. Die zweite Sorte wies grüne Hülsen mit variabler Helligkeit auf und wurde nicht so lange reifen gelassen, sodass diese Sorte genutzt werden kann, um mögliche Limitierungen des Verfahrens zu verdeutlichen.

3.3.2.2 Pflanzenanzucht

Am Forschungszentrum Jülich wurden 10 Pflanzen einer Bohnensorte mit grünen Hülsen und 10 Pflanzen einer Sorte mit gelben Hülsen unter Gewächshausbedingungen angebaut. Dazu wurde jeweils eine Pflanze in einem 3 Liter Plastiktopf mit ED73 Substrat (*Einheitserde Werkverband e.V.*, Uetersen, Deutschland) aufgezogen. Die Aussaat geschah am 13. September 2021 und die Temperatur wurde auf 25°C/18°C Tag/Nacht bei einer Luftfeuchtigkeit von 50% eingestellt. Mit Hilfe künstlicher Beleuchtung wurde ein Tag/Nacht Rhythmus von 16/8 Stunden realisiert, wozu man MGR-K 400 Lampen (*DH Licht GmbH*, Wülfrath, Deutschland) nutzte. Die Pflanzen wurden einmal täglich automatisch durch Tropfbewässerung bewässert. 107 Bohnenhülsen der grünen Sorte wurden am 28. Oktober 2021 geerntet und am selben Tag fotografiert. Die geernteten Hülsen befanden sich in der Vermarktungsreife. Sie hatten also ihre sortentypische Länge erreicht, die Samen waren aber noch in der Anfangsphase ihrer Entwicklung und somit relativ klein. Die Ernte und Aufnahme der 220 gelben Bohnenhülsen geschahen am 10. November 2021. Die Hülsen waren zur Ernte gelb und frisch und die Samen hatten zum größten Teil zwar schon ihre volle Größe erreicht, waren aber immer noch weich.

3.3.2.3 Erhebung der Referenzdaten

Die Entwicklung der Samenanlage wurde visuell anhand der Bildaufnahmen bewertet, um eine bestmögliche Vergleichbarkeit zwischen den Referenzdaten und den MS-Daten zu erhalten. Zur Validierung wurden 6 Durchlichtbilder von grünen Hülsen und 6 Durchlichtbilder von gelben Hülsen genutzt, wobei jedes Bild 9-10 Hülsen enthielt, sodass der *Validierungsdatensatz* aus insgesamt 117 Hülsen bestand (60 gelbe, 57 grüne Hülsen). Wurde ein Samen, welcher zwischen anderen Samen oder subapikal lag, nicht entwickelt, spricht man von sogenannten Fehlstellen. Sind keine Samen im oberen Viertel des basalen Endes oder darüber hinaus nicht entwickelt worden, spricht man von Schwanenhälsen. In der Referenzmessung wurde somit das Vorhandensein eines Schwanenhalses (als binäre ja/nein-Feststellung), sowie die Anzahl der Fehlstellen und Samen pro Hülse bestimmt.

3.3.2.4 Erstellung der Bildaufnahmen

Für diese Forschungsfrage wurde die gleiche Fotobox genutzt, wie im Abschnitt zuvor beschrieben, jedoch wurde die blaue Grundplatte durch eine Hintergrund-LED-Flächenleuchte ersetzt (Abbildung 17). Hierdurch wurde die Erkennung der Samenanlagen in Bohnenhülsen erleichtert, da die Samen aufgrund ihrer geringeren Lichtdurchlässigkeit sich von der Hülse abheben. Für diesen Modus wurde eine LEDP260C LED-Flächenleuchte (*GODOX Photo Equipment Co. Ltd*, Shenzhen, China) an Stelle der Grundplatte genutzt, während das Decken LED-Panel ausgeschaltet war. Die Lichtintensität der LED-Hintergrundbeleuchtung wurde an

die Helligkeit der Bohnenhülsen angepasst, um den Kontrast zwischen Samen und Hülsen zu verstärken. Bilder von grünen Bohnenhülsen wurden mit einer Beleuchtungsstärke von 14000lx (gemessen bei 0cm) und einer Belichtungszeit von 8ms aufgenommen. Gelbe Bohnenhülsen, die durchscheinender sind, wurden mit einer Beleuchtungsstärke von 4000lx (gemessen bei 0cm) und einer Belichtungszeit von 15ms aufgenommen.



Abbildung 17: Transmissionsaufnahme von gekrümmten Bohnenhülsen unter Nutzung einer LED-Flächenleuchte

3.3.2.5 Vorverarbeitung der Bilder

Da der verwendete Datensatz mit 12 Bildern klein war, wurde der Zuschnitt der Bilder manuell unter der Nutzung von *GIMP 2.8* durchgeführt. Die Bilder wurden nach Bedarf gedreht, sodass die Kanten des LED-Flächenpanels parallel zur Bildkante lagen. Anschließend wurde alles, was außerhalb der beleuchteten Fläche lag, weggeschnitten. Im nächsten Schritt wurden die Bilder von RGB in Graubilder umgewandelt. Man entschied sich dafür Graubilder zu verwenden, da die Erkennung der Samen vor allem auf Helligkeitsunterschieden im Bild beruht. Das Vereinzeln der Hülsen geschah mit dem Snipping-Tool von *GIMP 2.8*. Die Einzelbilder wurden schließlich als .png Dateien abgespeichert.

3.3.2.6 Detektions- und Klassifikationsmethode

Zur Erkennung der Samen, Fehlstellen und Schwanenhälse in den Graubildern wurde ein *Faster R-CNN* genutzt. Dieses weist grundsätzlich dieselbe Struktur wie das *Mask R-CNN* auf, nur dass dort ein Element (der sogenannte Segmentierungs-Kopf zur Erstellung der binären Bildmasken) fehlt. Das *Mask R-CNN* baut somit auf dem *Faster R-CNN* auf (Ren et al., 2017).

Training: Das hier genutzte *Faster R-CNN* zeichnet sich grundsätzlich durch die gleichen Modellmerkmale wie das zuvor in Abschnitt 3.2.2.6.1 vorgestellte *Mask R-CNN* aus. Die Eingabeschicht wurde lediglich dahingehend geändert, dass diese Graubilder anstatt RGB-Bilder akzeptiert. Es wurde ebenso wieder eine vortrainierte Version des Netzwerks genutzt (Transferlernen). Die Klassenzugehörigkeit sowie das umschließende Rechteck wurden für die jeweiligen Objekte (mit den drei Klassen Samen, Fehlstellen und Schwanenhäse) in 210 Graubildern mit einzelnen Bohnenhülsen mit dem VGG Image Annotator (Dutta & Zisserman, 2019) annotiert. Für das Training wurde ein separater *Entwicklungsdatensatz* mit 210 Einzelgraubildern von Hülsen erstellt, welcher keine Überschneidung mit dem *Validierungsdatensatz* hatte. In den 210 Bildern waren 50 Bilder mit Hülsen der grünen Sorte und 160 Bilder der gelben Sorte vorhanden. Diese wurden anteilig in 170 Trainings- und 40 Testbilder aufgeteilt. Die Trainingsdaten enthielten 713 annotierte Samen, 90 Fehlstellen und 75 Schwanenhäse. Die Testdaten enthielten 160 annotierte Samen, 22 Fehlstellen und 15 Schwanenhäse. Die Hyperparameter, die Augmentationen, sowie das Stoppkriterium für das Training wurden wie in Abschnitt 3.2.2.6.1 beschrieben festgelegt. Das Training wurde spätestens nach 15000 Epochen beendet.

Validierung: Für die Validierung wurden Einzelgraubilder von 60 gelben und 57 grünen Hülsen des *Validierungsdatensatzes* genutzt, welche nicht im *Entwicklungsdatensatz* enthalten waren. Die Konfidenzschwelle für die Modellvorhersagen wurde auf 0,5 festgelegt, was der Standardeinstellung entspricht. Die Klassenvorhersagen wurden dazu genutzt um zu schauen, wie viele Samen und Fehlstellen das Netzwerk in den Bildern erkannte, und ob es Schwanenhäse identifizierte. Die umschließenden Rechtecke gaben Aufschluss über die Lokalisation der entsprechenden Elemente, sodass eine positionsweise Validierung durchgeführt werden konnte.

3.3.3 Ergebnisse

Die Samenbildung in Bohnenhülsen wurde anhand von Bildern bewertet, die im *Durchlicht-Verfahren* aufgenommen wurden. Bei der gelben Sorte (Abbildung 18a,b) wurden 93,3 % der Samen und Fehlstellen durch das *Faster R-CNN* korrekt erkannt (Abbildung 18a). Dieses Ergebnis teilt sich auf in 96,9 % richtig erkannte Samen und 76,7 % richtig erkannte Fehlstellen. Die Spezifität für die Samenerkennung lag bei 96,5% und für die Erkennung von Fehlstellen bei 86,8%. Hülsen mit Schwanenhäsen wurden in 90 % aller Fälle erkannt (Abbildung 18b). Im Vergleich dazu betrug die Korrektklassifikationsrate bei der grünen Sorte (Abbildung 18c,d) 79,8 %, nämlich 82,9 % für Samen und 42,9 % für Fehlstellen (Abbildung 18c). Die Spezifität für die Samenerkennung lag bei 100%, die für die Erkennung von Fehlstellen bei 75%. Schwanenhäse wurden mit einer Korrektklassifikationsrate von 75,4 % erkannt (Abbildung 18d). Abbildung 19 zeigt Beispiele für die Detektion von Samen, Fehlstellen und Schwanenhäsen für die gelbe (Abbildung 19a,b) und grüne Sorte (Abbildung 19c,d).

a)	MS	Samen	190	7		197	96.5%
		Fehlstelle	5	33		38	86.8%
		nicht erkannt	1	3		4	0%
			196	43	0	239	93.3%
		Samen	Fehlstelle	nicht erkannt			
		Referenz					

b)	MS	Nein	9	4	13	69.2%
		Ja	2	45	47	95.7%
			11	49	60	90%
		81.8%	91.8%			
		Schwanenhals	Nein	Ja		
			Referenz			

c)	MS	Samen	285			285	100%
		Fehlstelle	4	12		16	75%
		nicht erkannt	55	16		71	0%
			344	28	0	372	79.8%
		82.9%	42.9%	0%			
		Samen	Fehlstelle	nicht erkannt			
		Referenz					

d)	MS	Nein	40	9	49	81.6%
		Ja	5	3	8	37.5%
			45	12	57	75.4%
		88.9%	25%			
		Schwanenhals	Nein	Ja		
			Referenz			

Abbildung 18: Evaluierung der Analyse der Samenanlage mittels Faster R-CNN. Die Ergebnisse sind für die gelbe Bohnensorte (a, b, n=60) und die grüne Bohnensorte (c, d, n=57) aufgeteilt, wobei a) und c) die Ergebnisse für die Erkennung der Samen und Fehlstellen und b) und d) die Ergebnisse für die Erkennung von Schwanenhälsen zeigen (Jollet et al., 2023).

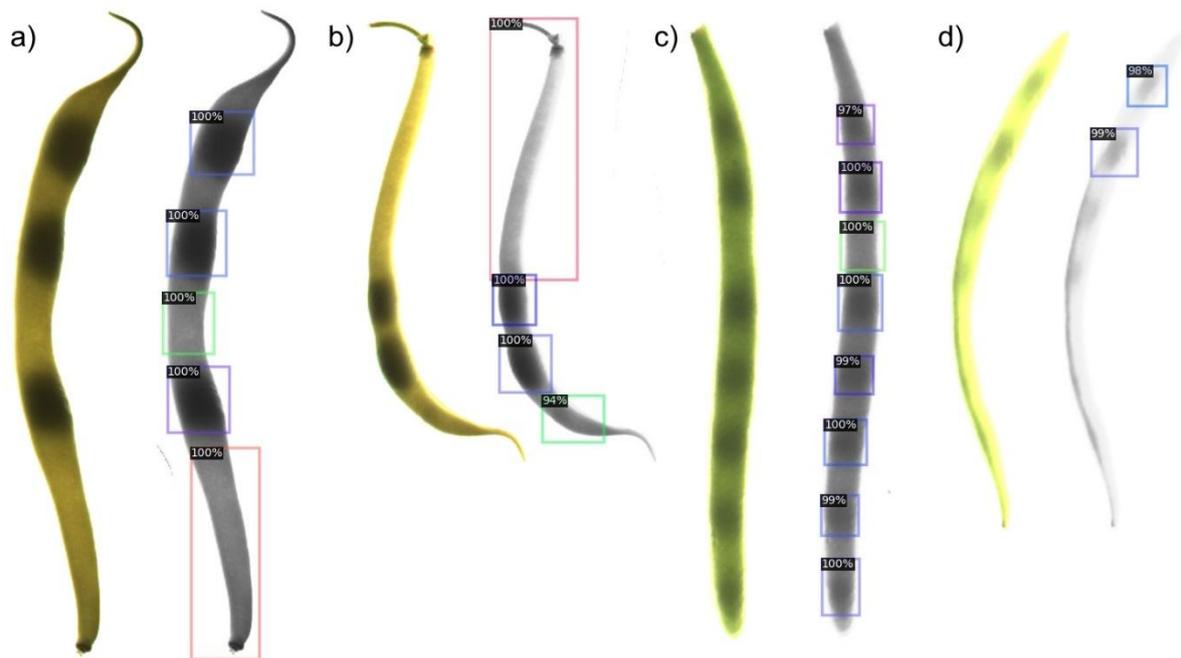


Abbildung 19: Faster R-CNN-Vorhersagen für Samen (blaue Rechtecke), Fehlstellen (grüne Rechtecke) und Schwanenhäule (rote Rechtecke). Die Werte in den Kästchen geben die Konfidenzwerte an. a) und b) zeigen gelbe Bohnenhülsen, c) eine dunklere grüne Bohnenhülse mit größeren Samen und d) eine hellere grüne Bohnenhülse mit kleineren Samen (Jollet et al., 2023).

3.3.4 Diskussion

Um Rückschlüsse auf die Anzahl der Samen, Fehlstellen und Schwanenhäule zu ziehen, verlassen sich die Züchter auf ihre haptischen Beurteilungen. Die Oberfläche der Hülse wird mit den Fingern abgetastet, um die Samen zu zählen und Fehlstellen und Schwanenhäule zu finden. Gerade bei kleinen Samen oder einer fehlenden Samenmarkierung (wie es bei den meisten züchtungsrelevanten Sorten erwünscht ist) führt dieses Verfahren zu schlechten Ergebnissen. Das hier vorgestellte Verfahren zeigt sehr gut, bei welchen Sorten und Entwicklungsstadien bereits nutzbare Ergebnisse erzielt werden können und wo Limitierungen

bestehen: So konnte für die durchscheinende gelbe Sorte, durch entsprechende Abstimmung der Durchlichtstärke und Kameraeinstellungen, ein guter Kontrast zwischen den Samen und dem Rest der Hülse erzielt werden konnte (Abbildung 19a,b). Dies ermöglichte die Erkennung der Samen, Fehlstellen und Schwanenhälse durch das neuronale Netz *Faster R-CNN*. Bei gut entwickelten Hülsen wurden Samen, Fehlstellen und Schwanenhälse durch das *Faster R-CNN* zuverlässig erkannt, was sich an einer Korrekturklassifikationsrate von 96,9% für Samen und 76,7% von Fehlstellen bei der gelben Sorte widerspiegelt. Die Grenzen des Verfahrens wurden anhand einer grünen Sorte mit variabler, aber vergleichsweise geringer Transmission, kleinen Samen, und Hülsen in verschiedenen Entwicklungsstadien getestet. Die Ergebnisse zeigten, dass die Genauigkeit der MS-Methode hauptsächlich von der Samengröße und der Transmission der Bohnenhülsen als Folge der Sorteneigenschaften und des Alters der Hülse beeinflusst wird. Zwar konnten bei der grünen Sorte immer noch 82,9% der Samen erkannt werden, aber es konnten lediglich 12 der insgesamt 28 Fehlstellen richtig identifiziert werden. Um einen ausreichend starken Kontrast zwischen den Samen und der Hülse im Bild zu erzeugen, müssen die Beleuchtungsstärke des Gegenlichts sowie die Belichtungszeit der Kamera gut auf die Transmissionseigenschaften der Hülse abgestimmt werden, was durch eine entsprechende Vorauswahl bzw. Vortests ermöglicht werden kann. Waren hingegen hell- wie auch dunkelgrüne Hülsen in einer Aufnahme enthalten, führten gute Einstellungen für dunkle Hülsen zu überbelichteten hellen und umgekehrt und damit zu kontrastarmen Hülsen (Abbildung 19b,c).

3.4 Chromatische Analyse von RGB-Bilddaten zur Evaluierung von Sorteneigenschaften, sowie Stresssymptomen in der Gemüsezücht

Um das Forschungsthema detailliert zu bearbeiten, wird es im Folgenden in 3 Themenkomplexe eingeteilt:

Farbanalyse unter kontrollierten Bedingungen: Der erste Themenkomplex behandelt die farbliche Analyse von Buschbohnenhülsen unter kontrollierten Bedingungen. Diese Kulturart ist dahingehend interessant, da es nicht nur grüne und gelbe Hülsen gibt, sondern auch eine Variante der gelben Sorte, welche zwar grundsätzlich gelbe Hülsen ausbilden, jedoch auf Grund von Genotyp-Umwelt Interaktionen einen leichten Grünstich aufweisen können und als Ausschuss gelten. Es geht hier somit um sehr feine Farbunterschiede, die es zu erfassen gilt. Im Folgenden soll untersucht werden, ob RGB-Verfahren in der Lage sind, Bohnenhülsen mit diesem Grünstich zuverlässig zu identifizieren. Weiterhin wird eine Helligkeitsanalyse grüner Hülsen erfolgen.

Infektionsevaluierung unter kontrollierten Bedingungen: Der zweite Themenkomplex beschäftigt sich mit der Bewertung von Infektionsstadien von Möhrenblättern (*Daucus carota*) mit dem Erreger *Alternaria dauci*. Die Anatomie von Möhrenblättern ist äußerst feingliedrig, was die Erkennung von leichten Krankheitssymptomen an der Pflanze besonders schwierig macht. Die Möhre ist als Modellpflanze also insofern interessant, da sie die technischen Grenzen eines RGB-Verfahrens, aber auch der Analyseverfahren zur Klassifikation des Infektionsstadiums, auf eine besondere Probe stellt.

Farbanalyse unter Feldbedingungen: Der dritte Themenkomplex befasst sich mit der präzisen Farbanalyse von Bildern, welche unter herausfordernden Feldbedingungen aufgenommen wurden. Es sollen hier vier Interpolationsmethoden miteinander verglichen werden, und schließlich ein Bezug zu einer praktischen Anwendung hergestellt werden. Als Modellpflanze wird hier der Weißkohl (*Brassica oleracea* convar. *capitata* var. *alba*) genutzt, da es sich hierbei um eine sehr empfindliche Kultursorte handelt, welche besondere Ansprüche an den Boden

und das Klima stellt und in Stresssituationen vielerlei Verfärbungen wie beispielsweise Chlorosen und Nekrosen, aber auch lila Anthocyan Verfärbungen ausprägt. Des Weiteren bildet der Weißkohl je nach Genotyp und Umwelteinfluss verschiedenste Grundfarben des Blattwerks wie grün, grünlich-weißlich, oder grünlich-bläulich aus, was für Farbanalysen mittels RGB-Bildgebung besonders interessant ist.

3.4.1 Forschungsstand

Der Erfolg einer Farbanalyse bei Gemüsepflanzen mittels maschinellen Sehens hängt von vielerlei Faktoren ab, von denen hier die Bildqualität zu nennen ist. Im wissenschaftlichen Kontext bedeutet eine hohe Bildqualität eine möglichst genaue Darstellung der Wirklichkeit, wobei in 2-dimensionalen Darstellungen vor allem eine ausreichende Auflösung, minimale Bildverzerrungen und eine möglichst genaue Farbwiedergabe anzustreben sind. Dies kann einerseits dadurch erreicht werden, dass man möglichst kontrollierbare Aufnahmebedingungen schafft (zum Beispiel durch Nutzung einer Fotobox) und die Kamera entsprechend der Umgebung einstellt (um in einem fixen Setting zu arbeiten), oder indem man das Bild nachträglich korrigiert. Vor allem unter Feldbedingungen ist es sehr schwierig gute Aufnahmebedingungen zu schaffen, was vor allem durch die natürlich fluktuierenden Lichtbedingungen sowie Schattenwurf bedingt ist. Feldbedingungen stellen somit eine erhebliche Herausforderung für Farbanalysen dar. Dabei ist jedoch zu unterscheiden, ob man farbliche Eigenschaften lediglich lokalisieren und klassifizieren möchte, oder ob Farben quantifiziert werden müssen.

Es gibt zahlreiche Studien die zeigen, wie Methoden des maschinellen Lernens helfen, bestimmte Stresssymptome auch in komplexen Bildern zuverlässig zu lokalisieren und klassifizieren. Bei der Klassifikation kann jedoch noch weiter differenziert werden. So gibt es binäre Klassifikationen (z. B. für gesunde und kranke Blätter) und Klassifikationen für mehr als zwei Gruppen (z. B. für die Identifikation verschiedener Krankheitserreger). Eine Arbeitsgruppe untersuchte beispielsweise den Gesundheitszustand von Bohnenblättern und klassifizierte diese zunächst binär mittels unterschiedlicher neuronaler Faltungsnetzwerke (Densenet121, ResNet34, ResNet50, VGG-16, and VGG-19), wobei Densenet121 eine Korrektklassifikationsrate von 99,03% erreichte. Im nächsten Schritt klassifizierten sie die Blätter nach ihrem Krankheitsbild, also gesunden Blätter, Blättern infiziert mit *Pseudocercospora griseola* (engl. Angular leaf spot) und Blättern infiziert mit *Uromyces appendiculatus* F. Strauss (engl. Bean rust). Hier erreichten sie eine Korrektklassifikationsrate von 91,01% (Abed et al., 2021). Die Klassifizierung der zuvor genannten Krankheiten in der Bohne wurde mit einem ähnlichen Ansatz auch von einer anderen Arbeitsgruppe untersucht, wobei hier ein MobileNet-Netzwerk genutzt wurde und eine Korrektklassifikationsrate von 92% erzielt wurde (Elfatimi et al., 2022). Harakannanavar et al. (2022) untersuchten 5 unterschiedliche Krankheiten (plus eine gesunde Klasse) an Tomatenblättern und segmentierten einzelne Blätter zunächst mittels K-Means Clustering um anschließend mittels Discrete-Walvet-Transforms (DWT) und einer Grey-Level Co-Occurrence Matrix (GLCM) Merkmale wie die Homogenität, Autokorrelation, Ungleichheit, Entropie, Quadrierte Summe, das Mittel und die Varianz zu extrahieren. Diese Merkmale wurden schließlich in verschiedene ML-Algorithmen (SVM, KNN, CNN) gespeist um den Krankheitstyp zu klassifizieren, wobei mit dem CNN die höchste Korrektklassifikationsrate mit 99,6% erreicht werden konnte. Eine weitere Arbeitsgruppe stellte eine Webapplikation vor, welche gesunde Möhrenblätter von mit *Alternaria dauci* infizierten Blättern und gesunde Möhrenwurzeln von solchen mit der Cavity Spot Krankheit (Erreger *Pythium spp.*) unterschied. Für die einzelnen Klassen gab es 148-456 Trainingsbilder, welche aus selbsterhobenen und aus dem Web entnommenen Bildern bestanden. Für das Training skalierten sie die Größe aller Bilder auf 300*300 Pixel, führten diverse Augmentierungen durch und teilten die Daten in Trainings- und Testbilder auf. Die Autoren testeten insgesamt 5 unterschiedliche selbstdesignte Faltungsnetzwerke (CNNs) zur

Lösung der Klassifikationsaufgabe und erreichten mit ihrem besten Modell eine Korrektklassifikationsrate für den Testdatensatz von 99,81% (Ray et al., 2022). Eine weitere Forschungsgruppe untersuchte die Klassifizierung von 5 Krankheiten an Möhrenwurzeln (+ eine Klasse für gesunde Wurzeln) mittels verschiedener neuronaler Faltungsnetzwerken. Die Autoren nutzen in dieser Studie einen natürlichen Datensatz mit 80 aufgenommenen Bildern pro Klasse und einen synthetischen Datensatz mit ca. 275 Bildern pro Klasse, in denen der Hintergrund geändert wurde. Weiterhin wurden je 5 unterschiedliche Augmentierungen für jede Klasse durchgeführt, sodass insgesamt 1775 Trainingsbilder für jede der 6 Klassen vorhanden waren. Es wurden insgesamt 4 unterschiedliche neuronale Faltungsnetzwerke zur Klassifikation getestet, wobei das *Inception V3* Netzwerk mit einer Korrektklassifikationsrate von 97,4% das beste Ergebnis erreichte (Methun et al., 2021). Q. Wang & Qi (2019) nutzten ein *Faster R-CNN* um Tomaten unter Gewächshausbedingungen an der Pflanze zu lokalisieren und gleichzeitig 6 unterschiedliche Stresssymptome/Krankheiten (plus die gesunde Klasse) an den Früchten zu klassifizieren. Sie erreichten eine durchschnittliche Korrektklassifikationsrate von 90,87% (über alle Klassen gemittelt).

Neben der Klassifikation der Krankheitsart ist es oft auch wichtig das Infektionsstadium beziehungsweise den Grad des Stresses einer Pflanze zu beurteilen. So kann man beispielsweise das Infektionsstadium diskreten Klassen zuordnen und ein ML-Verfahren entsprechend dieser Einteilung trainieren. Für diese Klassifikation werden häufig Ansätze des tiefen Lernens verwendet: Prabhakar et al., (2020) nutzten ein auf ResNet101-basiertes CNN um den Infektionsgrad von einzelnen Tomatenblättern mit *Alternaria solani* Befall in 4 Klassen zu unterteilen (gesund, leicht, mittel, schwer). Hier wurde eine Korrektklassifikationsrate von 94,6% erreicht. Eine andere Forschungsgruppe untersuchte die Funktionalität verschiedener CNNs zur Klassifikation von 4 Infektionsstadien von Apfelblättern mit dem Erreger *Botryosphaeria obtusa*, wobei die besten Ergebnisse (90,4%) mit einem VGG16 Netzwerk erreicht wurden, welches mittels Transferlernens trainiert wurde. Des Weiteren kann man das Infektionsstadium bzw. den Grad des Stresses auch als kontinuierlichen Wert bestimmen, indem man den Anteil der Farbe an einer Verteilung über das Blatt oder die Frucht im Bild bestimmt. So konzentrierte sich folgende Studie darauf, drei unterschiedliche Krankheiten des Weißkohls (*Brassica oleracea* convar. *capitata* var. *alba*) zu klassifizieren und das Infektionsstadium zu bestimmen. Zur Lösung der Aufgabenstellung wurden Farbmerkmale aus dem RGB-Farbraum extrahiert und eine Farbsegmentierung mit Hilfe des K-Means Clustering Algorithmus durchgeführt. Der prozentuale Anteil an beeinträchtigter Blattfläche, welcher aus dem RGB-Farbraum extrahiert wurde, diente zur Bestimmung des Infektionsstadiums, während die Farbcluster (des K-Means Algorithmus) Rot, Grün, Gelb und Weiß Aufschluss über die Art der Krankheit gaben (Divyashree et al., 2015).

Um die Anteile verschiedener Farben an einer Gesamtstruktur jedoch verlässlich anhand eines Bildes zu bestimmen (und eventuell auch mit anderen Messungen durchgeführt unter anderen Bedingungen zu vergleichen), ist es vorteilhaft, wenn die dargestellten Farben möglichst nah an der Realität, also den echten Farbwerten, liegen. Dies ist in der Regel nicht durch den Messaufbau zu erreichen. Man sollte daher eine geeignete Referenz nutzen, um die Farbdarstellung im Bild korrigieren zu können. Im Bereich der Fernerkundung mittels Drohnen wird häufig ein Ansatz der semantischen Korrespondenz für eine Farbkalibrierung genutzt. Im Grunde nutzt man bei diesem Ansatz Referenzbilder, um die Farben des Ausgangsbildes in die Farben des Referenzbildes zu überführen (H. Huang et al., 2021, 2022). Ein anderer Ansatz besteht darin, eine Farbtabelle mit genau definierten Farben in das Bild zu integrieren, um mit einer geeigneten Interpolationsmethode die Farben im Ausgangsbild anzupassen.

Sunoj et al. (2018) nutzten einen linearen Interpolationsansatz und testeten verschiedene Zusammensetzungen (Reihenfolgen) von Referenzfarben und sie testeten ebenfalls die Genauigkeit der Kalibrierung in Abhängigkeit von der Anzahl der Referenzfarben und unter

unterschiedlichen Aufnahmebedingungen im Labor und Feld. Je nach Farbzusammensetzung reichten bereits 6 Referenzfarben aus, um den Fehler zwischen den Referenzfarben und den Bildfarben auf unter 5% zu senken. Ab 10 bis zu den maximal 24 getesteten Referenzfarben hatte die Zusammensetzung, unabhängig der Aufnahmebedingungen, keine nennenswerten Auswirkungen mehr. Eine andere Forschungsgruppe untersuchte die Farbe unterschiedlicher grüner Gemüsesorten. Da hier sehr feine Graduierungen von Grüntönen zu bewerten waren, kalibrierten sie ihre Bilder zuvor mit 140 Referenzfarben, welche vor allem den Farbraum der zu untersuchenden Gemüsesorten gut abdeckten, und einem polynomischen Interpolationsansatz 3. Grades (Manninen et al., 2015). Eine weitere Arbeit verglich eine PLS-Interpolation mit einer 3D TPS-Interpolation zur Farbkalibrierung anhand verschiedener Bilder, welche unter kontrollierten Bedingungen aufgenommen wurden. Sie fanden heraus, dass die 3D TPS-Interpolation präzisere Ergebnisse liefern kann und auch lokale Unterschiede im Bild besser kompensiert, jedoch ist der Ansatz auch auf eine gute Abdeckung des Farbraums durch Referenzfarben angewiesen. Der Ansatz funktioniert demnach besser, wenn mehr Referenzfarben vorhanden sind und ist im Vergleich zu PLS anfälliger, wenn nur wenige Referenzfarben zur Verfügung stehen (Menesatti et al., 2012).

3.4.2 Farbklassifikation und Helligkeitsanalyse unter kontrollierten Bedingungen zur Sortenbeurteilung

3.4.2.1 *Material und Methoden*

3.4.2.1.1 Modellpflanzen

Als Modellpflanzen zur Untersuchung der Forschungsfrage soll hier die Buschbohne dienen, da die Hülsen je nach Sorte unterschiedliche Farben und Helligkeiten ausprägen. Diese Kulturart ist weitergehend interessant, da es nicht nur grüne und gelbe Hülsen gibt, sondern auch Varianten, welche grundsätzlich gelbe Hülsen ausbilden, jedoch auf Grund von Sorteneigenschaften einen leichten Grünstich aufweisen können. Es geht hier somit um sehr feine Farb- und Helligkeitsunterschiede, die es zu erfassen gilt. Im Folgenden soll untersucht werden, ob RGB-Verfahren in der Lage sind, Bohnenhülsen mit unterschiedlichsten Farbausprägungen zuverlässig zu identifizieren.

3.4.2.1.2 Pflanzenanzucht

Die Anzucht der Pflanzen ist in Abschnitt 3.2.2.2 beschrieben.

3.4.2.1.3 Erhebung der Referenzdaten

Die Helligkeit der Hülsen wurde nur für grüne Hülsen bestimmt, da nur für diese Referenzdaten vorlagen. Der Züchter bewertet die Helligkeit von gelben Hülsen nicht, da dieser Parameter für diese Hülsen nicht marktrelevant ist. Die Helligkeit grüner Hülsen wurde visuell mit Hilfe einer ordinalen Skala mit Werten zwischen 1 und 9 bestimmt. Eine Bewertung von 1 steht dabei für eine weiße Bohne und die Bewertung 9 für eine sehr dunkelgrüne Hülse. Weiße Bohnenhülsen sind im Zuge dieser Arbeit jedoch nicht vorgekommen (Abbildung 20a). Für gelbe Bohnensorten ist es hingegen wichtig zu bewerten, ob eine Sorte dazu neigt Hülsen mit grüner Verfärbung (im Folgenden zweifarbig bezeichnet) zu entwickeln und somit die Qualität einer Sorte verringert. Die Farbklassifikation in grün, gelb, oder zweifarbig wurde ebenfalls visuell durchgeführt (Abbildung 20b).



Abbildung 20: a) Referenzskala zur visuellen Bewertung der Helligkeit grüner Bohnenhülsen; b) Referenzskala zur Beurteilung der Farbe von Bohnenhülsen

3.4.2.1.4 Erstellung der Bildaufnahmen

Die Erhebung der Daten ist in Abschnitt 3.2.2.4 beschrieben.

3.4.2.1.5 Vorverarbeitung der Bilder

Die Vorverarbeitung der Bilder ist in Abschnitt 3.2.2.5 beschrieben.

3.4.2.1.6 Farb- und Helligkeitsanalyse mittels 2D RGB-Verfahren

Insgesamt galt es für die untersuchten Hülsen 3 Farbklassen zu unterscheiden: Grüne, Gelbe und Zweifarbig. Da sich grüne Hülsen sehr stark von Gelben und Zweifarbig unterscheiden, konnten diese mit einer einfachen Schwellenwertanalyse des H-Kanals im HSV-Farbraum bestimmt werden. Hatte ein Pixel, welches zur Hülse gehörte, einen H-Wert zwischen 70 und 180 wurde es als grün und mit Werten zwischen 15 und 35 als gelb klassifiziert. Überwog die Anzahl an grünen Pixeln, wurde die gesamte Hülse als grün klassifiziert. War dies nicht der Fall, wurde also der überwiegende Teil der Pixel als Gelb klassifiziert, dann wurde für diese ein trainiertes Random Forest Modell zur weiteren Unterscheidung von gelben und zweifarbig Hülsen genutzt. Dies war nötig, da der Grünstich bei zweifarbig Hülsen in den Bildern teils als grüner, teils als grauer und teils als dunkelgelber Schleier zu erkennen war. Da zweifarbig Hülsen sehr selten sind und es für diese Klasse nur wenige Daten (inklusive der Züchterklassen) gab, musste sowohl das Training wie auch die Validierung des RF-Modells mit den 90 gelben oder zweifarbig Hülsen durchgeführt werden, welche im Validierungsdatensatz enthalten waren, welcher in Abschnitt 3.2.2.2 beschrieben ist. Das RF-Modell wurde anhand von 10 Merkmalen trainiert: dem prozentualen Anteil und den mittleren Farbtonwerten der grünen und gelben Pixel sowie dem Mittelwert und der Standardabweichung der Farbtonwerte, der Sättigungswerte und der Helligkeitswerte aller Hülsenpixel. Das RF-Modell wurde wie folgt parametrisiert (siehe *RandomForestClassifier* (*scikit-learn*) Funktion für Details): Anzahl der Entscheidungsbäume 100, als Qualitätskriterium wurde "Entropie" eingestellt, welches die Aufteilung an jedem Knoten im Baum steuert und die Gesamtstruktur der Entscheidungsbäume beeinflusst. Alle anderen Parameter wurden auf den Standardeinstellungen belassen. Aufgrund der geringeren Datengrundlage wurde eine 10-fache Kreuzvalidierung durchgeführt, wofür die Daten jeweils nach dem Zufallsprinzip in 72 Trainings- und 18 Validierungsbilder von einzelnen Bohnenhülsen aufgeteilt wurden.

Wurde eine Hülse als grün klassifiziert, wurde dessen Helligkeit durch Berechnung des mittleren Helligkeitswerts (V-Kanal, HSV-Farbraum) aller Hülsenpixel bestimmt. Da jedoch vor allem runde Hülsentypen auf Grund ihrer Form zum Rand hin abdunkeln, wurde die Maske der Hülse vor der Berechnung der mittleren Helligkeit der Hülse um 15 Pixel erodiert.

3.4.2.2 Ergebnisse

Grüne Hülsen wurden mittels Schwellenwertsegmentierung des H-Kanals (HSV-Farbraum) korrekt von gelben und zweifarbigen Bohnenhülsen getrennt (100 % Korrektklassifikationsrate). Grüne Hülsen wurden anschließend im Hinblick auf ihre Helligkeit analysiert. Die Ergebnisse zeigen eine klare lineare Beziehung zwischen den Züchter-Referenzwerten und den berechneten Log-transformierten Helligkeitswerten (Abbildung 21). Da der Züchter für Hülsen der Typen Flach und Oval immer eine der breiten Seiten bewertete, wurden jene Hülsen, welche mit der flachen Seite zur Kamera ausgerichtet waren, in Abbildung 21 mit orangen Sternen markiert. Diese wurden jedoch nicht von der Analyse ausgeschlossen, da sie in der Mehrheit nicht als Ausreißer zu bewerten sind und, im Gegensatz zur Analyse der Länge und des Kalibers, keine Absoluten Werte miteinander verglichen wurden. Es wurde ein MAPE von 2,1% und ein Korrelationskoeffizient von $R=-0,87$ erreicht. Darüber hinaus ergab die Residualanalyse keine statistisch signifikanten Unterschiede ($p=0,39$) zwischen den Bohnenhülsenkategorien Rund, Oval und Flach. Bei nicht-grünen Hülsen konnte das RFC-Modell zwischen gelben und zweifarbigen Hülsen mit einer Korrektklassifikationsrate von $83\pm 7\%$ (getestet mit einer 10-fachen Kreuzvalidierung) unterscheiden.

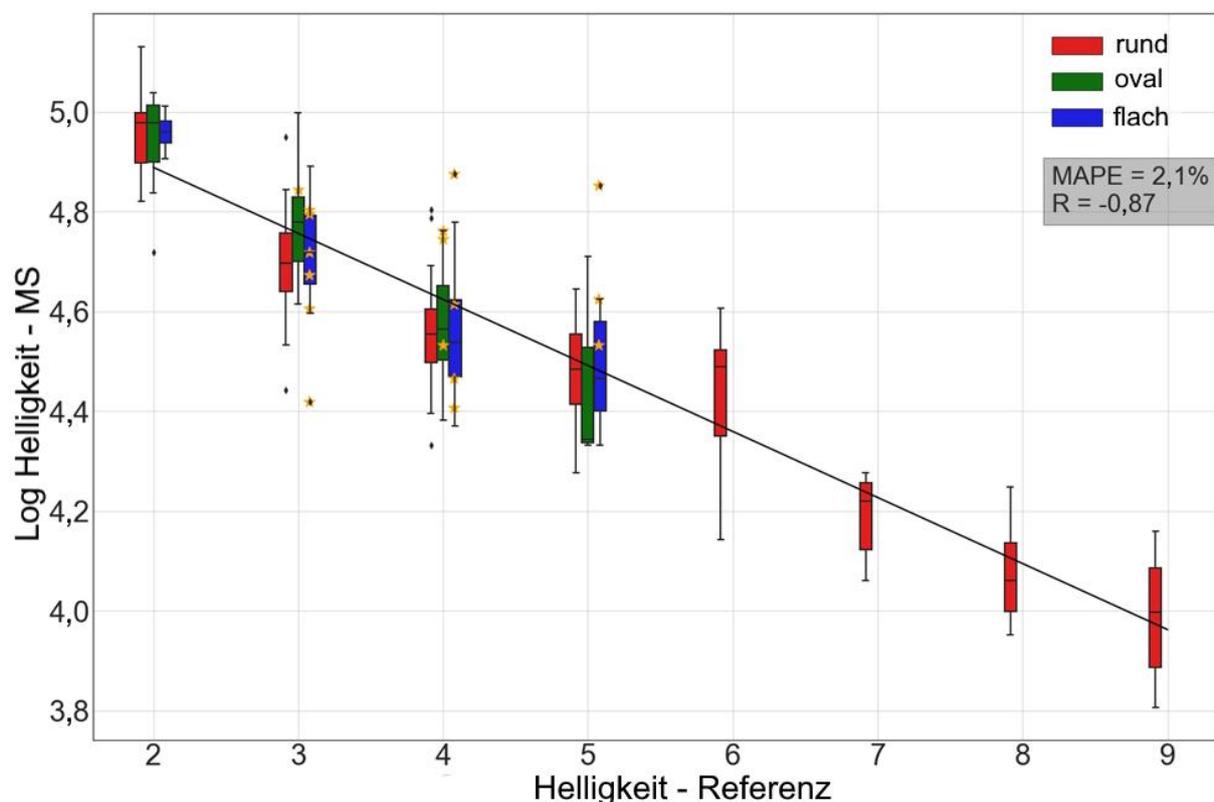


Abbildung 21: Zusammenhang zwischen den Züchterbewertungen und den Messungen der MS-Methode für die Helligkeit von grünen Bohnenhülsen. Die orangen Sterne deuten ovale oder flache Hülsen an, welche mit der Schmalen Seite zur Kamera ausgerichtet waren, während die visuellen Bewertungen an einer der breiten Seite durchgeführt wurden (Jollet et al., 2023).

3.4.2.3 Diskussion

Mit der RF-Methode konnten wir die Helligkeit von grünen Bohnenhülsen präzise (MAPE=2,1% und $R=-0,87$) bestimmen. Weiterhin ist eine systematische Abweichung von der Identitätslinie

über alle Hülsenklassen hinweg zu beobachten, was wahrscheinlich auf die nicht-linearität der Züchterskala zurückzuführen ist, welche nicht durch eine einfache Funktion beschrieben werden kann. Das Verfahren zur Helligkeitsabschätzung ist jedoch ohne weiteres auf andere Hülsenfarben übertragbar, wobei beachtet werden muss, dass eine Bewertung immer nur innerhalb einer Farbklasse (Grün, Gelb, Lila) erfolgen kann, da die Werte nicht normalisiert sind. Leichte Abweichungen zwischen Referenz- und MS-Messungen wurden jedoch bei einigen wenigen Bohnenhülsen beobachtet. Dies ist wahrscheinlich darauf zurückzuführen, dass in diesen Fällen unterschiedliche Hülsenseiten für die Referenz- und die MS-Messung ausgewertet wurden. Die Helligkeiten der beiden Hülsenseiten können sich, je nach sonnen zu- und abgewandter Seite, jedoch deutlich unterscheiden.

3.4.3 Bewertung komplexer biotischer Stresssymptome unter kontrollierten Bedingungen

3.4.3.1 Material und Methoden

3.4.3.1.1 Modellpflanzen

Als Modellpflanze dienen hier mit *Alternaria dauci* infizierte Möhrenblätter, was die Bildverarbeitung vor eine doppelte Herausforderung stellt, da die Infektion mit mikroskopisch kleinen Läsionen beginnt und die Möhrenblätter sehr feingliedrig sind. Die Möhre ist als Modellpflanze also insofern interessant, da sie die technischen Grenzen eines RGB-Verfahrens, aber auch der Analyseverfahren zur Klassifikation des Infektionsstadiums, auf eine besondere Probe stellt.

3.4.3.1.2 Pflanzenanzucht

Die Möhrenpflanzen wurden am Standort Nunhem angebaut. Der Experimentstandort Nunhem (NL) liegt 22 Meter über dem Meeresspiegel im Tal des Flusses Haelensebeek. Zwischen 1990 und 2022 betrug die Jahresdurchschnittstemperatur 10,8°C und es fielen im Schnitt 685mm Regen im Jahr. Außerdem schien die Sonne durchschnittlich 1789 Stunden im Jahr (Quelle: Wetterstation Eil, Niederlande; <https://www.knmi.nl>; 09,03,2023). Die Pflanzen unterschiedlicher Genotypen wurden in den Jahren 2021 und 2022 unter Feldbedingungen angebaut und künstlich mit dem Erreger der Möhrenschräge (*Alternaria dauci*) infiziert. Die Symptome der Krankheit werden ca. 8-10 Tage nach der Infektion an den älteren Blättern sichtbar (Que et al., 2019). Die Infektion erfolgt unter moderaten Temperaturen zwischen 16 und 25°C und längerer Blattnässe von 8 bis 12 Stunden durch beispielsweise Regen, Nebel oder künstliche Bewässerung (Farrar et al., 2004; Gugino et al., 2004). Zunächst entstehen kleine, grün-braune, mit Wasser angereicherte Läsionen, welche sich im Laufe der Zeit vergrößern, dunkelbraun bis schwarz werden und teilweise von chlorotischen Ringen umschlossen sind (Farrar et al., 2004; Lecomte et al., 2014). Unter günstigen Bedingungen verbinden sich die Läsionen großflächig und führen zum Absterben des infizierten Blattes. Durch die reduzierte photosynthetisch aktive Blattfläche wurden bereits Ertragsminderungen von bis zu 40-60% berichtet (Ben-Noon et al., 2001; Que et al., 2019). Im Rahmen dieser Arbeit war es nicht notwendig, gesonderte Feldexperimente zur Generierung von Daten durchzuführen, da von den Versuchsflächen des Projektpartners *Nunhems* genug mit *Alternaria dauci* infiziertes Blattmaterial gesammelt werden konnte. Die Pflanzung erfolgte somit nach den Standardpraktiken des Projektpartners. Die Aussaat der Pflanzen fand jeweils Anfang August in den Jahren 2021 und 2022 statt und die Pflanzen wurden, ausgenommen der Behandlung mit Fungiziden, nach konventionellen Methoden angebaut. Die Inokulation der unterschiedlich resistenten Sorten mit Sporen von *Alternaria dauci* fand in den beiden Jahren jeweils Mitte September statt. Die Zeitpunkte wurden jeweils so gewählt, dass feuchtes Klima die Infektion begünstigte. Ende September bzw. Anfang Oktober zwischen den BBCH-Stadien 45 und 49 (je nach Sorte) wurden Blätter mit unterschiedlich starken Infektionsgraden gesammelt und von Zuchtexterten von *Nunhems* nach deren Standardpraxis bewertet.

3.4.3.1.3 Erhebung der Referenzdaten

Die Zuchtexterten klassifizieren das Infektionsstadium anhand des Blattmaterials (nicht der Bilder) in 9 diskrete Klassen. Die Züchterskala reichte von 1, komplett infiziertes Blatt, bis 9, gesundes Blatt (Abbildung 22), wobei in die visuelle Bewertung laut Experten neben dem chromatischen Eindruck des Blattes auch der geschätzte Anteil der infizierten Fläche, die größte verbundene Infektionsfläche, sowie die Welke des Blattes einfließen. Um die getesteten Methoden zur Bestimmung des Infektionsstadiums gut vergleichen zu können, wurde zur Validierung immer derselbe *Validierungsdatensatz* genutzt. Der *Validierungsdatensatz* bestand insgesamt aus 192 Bildern, wobei 27 Bilder dem Infektionsstadium 1, 47 dem Stadium 2, 31 dem Stadium 3, 32 dem Stadium 4, 6 dem Stadium 5, 18 dem Stadium 6, 10 dem Stadium 7, 13 dem Stadium 8 und 8 dem Stadium 9 zuzuordnen waren. Die Bilder des *Validierungsdatensatzes* wurden nicht für Entwicklungszwecke genutzt. Die Züchterbewertungen wurden im Kontext des maschinellen Lernens nicht als nominalskalierte Klassen, wie beispielsweise bei Objektbezeichnungen, sondern als ordinalskalierte Werte betrachtet. Diese Skala berücksichtigt, dass die Abfolge nicht aber die Abstände zwischen den Klassen bekannt ist.

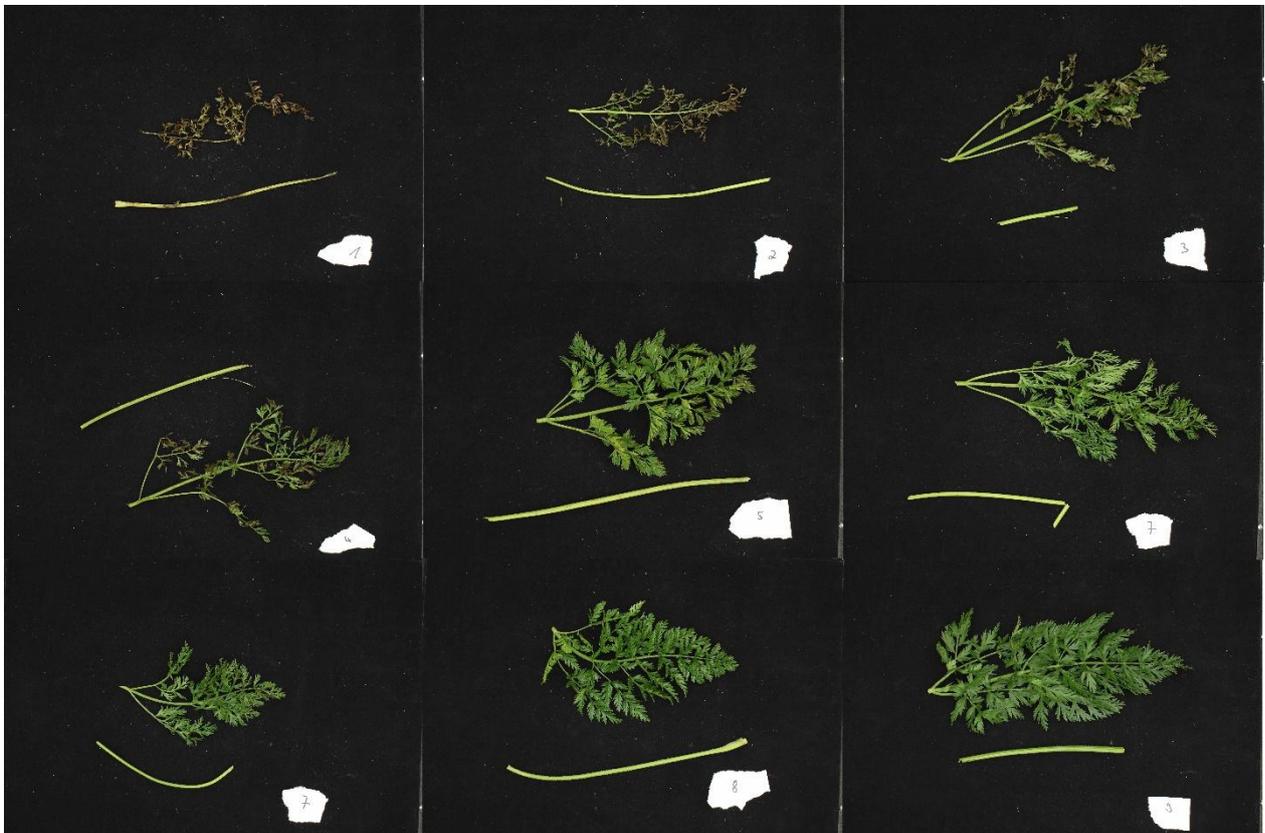


Abbildung 22: Infektionslevel einzelner Möhrenblätter mit *Alternaria dauci* nach Züchterskala. Diese reicht von 1, abgestorbene Blätter, bis 9, vollkommen gesunde Blätter ohne jegliche Läsionen. Diese Skala beruht auf Erfahrungswerten sehr erfahrener Züchter.

3.4.3.1.4 Erstellung der Bildaufnahmen

Es wurde eine fast baugleiche Fotobox genutzt wie im Abschnitt 3.2.2.4 beschrieben. Hier wurde jedoch eine *EOS 6D* (Canon AG, Tokio, Japan) Digitalkamera mit 20,6MP (und damit deutlich besserer) Auflösung verwendet, um auch kleine chlorotische oder nekrotische Infektionsstellen erkennbar zu machen. Das Kameraloch im LED-Panel war entsprechend größer. Man nutzte die in Abschnitt 3.2.2.4 beschriebene blaue Grundplatte, um eine klare Segmentierung der Blätter zu ermöglichen. Um die Grundplatte optimal im Bild einzufangen, wurde ein *EF 17-40mm f/4L USM* (Canon AG, Tokio, Japan) Objektiv genutzt, welches auf

40mm Brennweite adjustiert wurde. Eine mittlere Blendenzahl von F/9,93 wurde genutzt, damit keine Tiefenunschärfe im Bild entstand, aber dennoch möglichst viel Licht das Objektiv passieren konnte. Weiterhin wurde ein ISO-Wert von 100 eingestellt, um das Bildrauschen möglichst gering zu halten. Für optimale Belichtungsverhältnisse im Bild wurde eine Belichtungszeit von 1/125 Sekunden (8ms) angenommen. Da das genutzte LED-Panel laut Herstellerangaben eine Farbtemperatur von $5700^{\circ}\text{K} \pm 700^{\circ}\text{K}$ besaß, wurde der Weißabgleich der Kamera auf 5700°K eingestellt. Die Methode nach Z. Zhang (2000) wurde wieder genutzt, um die Auswirkungen von radialen Verzerrungen zu prüfen, welche jedoch auch hier nicht auftraten. Die Bilder wurden im unkomprimierten .PNG-Format abgespeichert. Die Fotobox wurde außerdem dahingehend erweitert, dass das LED-Panel mit einem Yeti 500X Lithium-Akku (GROFA Action Sports GmbH, 65520 Bad Camberg, Deutschland) betrieben werden konnte, was einen Betrieb der Fotobox im Feld ermöglichte (Abbildung 23). Insgesamt wurden 227 Bilder von Möhrenblätter vor der blauen Grundplatte aufgenommen, wobei 42 zur Klasse 1, 62 zur Klasse 2, 46 zur Klasse 3, 47 zur Klasse 4, 21 zur Klasse 5, 33 zur Klasse 6, 25 zur Klasse 7, 28 zur Klasse 8, und 23 zur Klasse 9 gehörten (hier sind die Validierungsdaten enthalten). Zusätzlich wurden für Entwicklungszwecke 782 zusätzliche Möhrenblätter gesammelt, welche sich wie folgt auf die 9 Infektionsklassen aufteilten: 98 Bilder der Klasse 1, 92 der Klasse 2, 91 der Klasse 3, 84 der Klasse 4, 70 der Klasse 5, 77 der Klasse 6, 102 der Klasse 7, 85 der Klasse 8 und 83 der Klasse 9.



Abbildung 23: Aufnahmebox für den Teilbereich Möhre. Das LED-Panel der Aufnahmebox wurde mit einem Camping Akku betrieben, wobei dessen Laufzeit mit einem Solarmodul verlängert werden kann. Die Aufnahmen wurden mit einer Digitalkamera gemacht, welche durch ein Loch im LED-Panel auf die Grundfläche der Aufnahmebox gerichtet war.

3.4.3.1.5 Vorverarbeitung der Bilder

Für die Analyse der Möhrenblätter wurde keine automatisierte Verarbeitungs-Pipeline erstellt. Alle Vorverarbeitungsschritte wurden in *GIMP 2.8* durchgeführt. Die Bilder wurden lediglich mit dem *Zuschneiden*-Werkzeug auf die Größe des Blattes zugeschnitten, um überflüssige Informationen für die nachfolgenden Analysen auszuschließen.

3.4.3.1.6 Bewertung biotischer Stresssymptome mittels 2D RGB-Verfahren

Zunächst wurde der Anteil der nicht-befallenen Blattfläche durch Schwellenwert-Segmentierung bestimmt, um auf das Infektionsstadium zurückzuschließen. Da eine *Alternaria*-Infektion zum Absterben der Blattzellen führt, zeichnet sich die nicht befallene Blattfläche durch eine grüne Färbung aus. Der Anteil der grünen Blattfläche hat dabei einen wesentlichen Einfluss auf den Gesamteindruck des Blattes, welcher in die Züchterbewertung einfließt. Zur Bestimmung des Anteils der grünen Blattfläche wurde das vorprozessierte Bild zunächst in den HSV-Farbraum transferiert, um die Farbeigenschaften des Blattes über den Farbtonwert-Kanal (H) zu klassifizieren. Die Schwellenwerte wurden dabei auf Basis der Übergangsbereichen zwischen den Grundfarben (Rot, Orange, Gelb, Grün, Cyan, Blau, Magenta) festgelegt, wie in Abschnitt 2.2.2 dargestellt. Mittels Schwellenwertsegmentierung im H-Kanal ($H \leq 200$ & $H > 240$) und V-Kanal ($V > 1$ & $V \leq 240$) wurde zunächst das gesamte Blatt vom Hintergrund getrennt. Die Identifikation der gesunden Blattpixel des so segmentierten Blattes basierte ebenfalls auf einer Schwellenwertsegmentierung im H-Kanal ($H \geq 70$ & $H < 150$), sodass der prozentuale Anteil der gesunden Blattpixel an der gesamten projizierten Blattfläche ermittelt werden konnte. Dieser wurde schließlich mit den Züchterbewertungen der Infektionsstadien verglichen.

Ein anderer Ansatz beruhte auf der Klassifikation der Bilder mit Hilfe eines *Random Forest* (RF) Modells. Insgesamt galt es 9 Infektionsstadien zu unterscheiden, und man extrahierte insgesamt 10 Merkmale aus den Pixelwerten der maskierten Möhrenblätter, um mit diesen ein *Random Forest* Modell zu trainieren. Zunächst nutzte man Texturmerkmale auf Basis des Graubildes, da sich die grundlegende Textur des Blattes im Verlauf der Krankheit verändert. Aus den segmentierten Pixeln des Graubildes wurden zunächst der Variationskoeffizient, die Entropie, die Asymmetrie, die Kurtosis sowie das Verhältnis von hellen und dunklen Pixeln bestimmt. Zur Einteilung der Pixel in hell und dunkel nutzte man das Histogramm der Helligkeitswerte, welches man in 2 *bins* (Gruppen) mit jeweils gleicher Breite einteilte. Die Annahme bestand darin, dass der Anteil der dunklen Pixel mit steigendem Befall zunimmt. Aus der binären Bildmaske berechnete man weiterhin das Verhältnis zwischen Fläche und Umfang der Maske, da auf Grund des sinkenden Turgor Drucks mit fortschreitendem Befall eine Veränderung des Verhältnisses zu erwarten ist. Aus selbigem Grund wurde das Verhältnis zwischen Länge und Breite des kleinsten umschließenden Rechtecks der Maske berechnet. Des Weiteren nutze man wie im vorherigen Abschnitt beschrieben den prozentualen Anteil der gesunden Blattfläche, aber auch den prozentualen Anteil der nicht-gesunden Blattpixel. Letztere definierte sich ebenfalls über den H-Kanal des HSV-Bildes H-Kanals ($H < 70$ & $H \geq 345$). Es wurde außerdem die Fläche der Kontur, sowie die Größe der größten nicht-gesunden Fläche bestimmt. Ein weiteres Merkmal bildete der Variationskoeffizient aller als gesund klassifizierten Pixel. Das *Random Forest* Modell wurde mit 135 Bildern aus dem *Entwicklungsdatensatz* trainiert (keine Überschneidung mit dem *Validierungsdatensatz*), sodass für jedes der 9 Infektionsstadien 15 Trainingsbilder vorhanden waren. Für das RF-Modell wurde als Qualitätskriterium ‚Entropie‘ festgelegt und die Anzahl der Entscheidungsbäume auf 1000 gesetzt. Alle anderen Parameter wurden auf den Standardeinstellungen belassen (siehe *RandomForestClassifier* (*scikit-learn*) Funktion für Details).

Schließlich untersuchte man noch die Klassifizierung der Infektionsstadien mittels neuronalem Faltungsnetzwerk. *EfficientNet* stellt eine Familie von neuronalen Faltungsnetzwerken mit

optimierter Skalierung dar, bei denen die Tiefe, Breite und Auflösung der Netzwerke systematisch aufeinander abgestimmt wurden. Dadurch ergeben sich kleinere und effizientere Faltungsnetzwerke (Tan & Le, 2019)(Tan & Le, 2019). *EfficientNet-V2* stellt eine Weiterentwicklung der *EfficientNet* Familie dar, wobei die Entwicklung dieser Netzwerkfamilie auf einer Kombination aus trainingspezifischer neuronaler Architektursuche und Skalierung beruhte. *EfficientNet-V2* erzielte bessere Ergebnisse als andere aktuelle Netzwerke, wie beispielsweise Vision Transformers, und trainierten gleichzeitig 5-11-mal schneller (Tan & Le, 2019). Für diese Untersuchung nutzte man ein auf *ImageNet* (Deng et al., 2010) vortrainiertes *EfficientNet-V2-L* (Transferlernen) Netzwerk. Das L (Large) im Namen gibt die Größe des Netzwerks in Bezug auf die Anzahl der trainierbaren Parameter (118515272) an. Das Netzwerk wurde folgendermaßen modifiziert: Zunächst froz man alle Gewichte des Netzwerks ein, um das Training effizienter zu gestalten. Durch das Einfrieren behielten sämtliche Gewichte den Wert, welche sie durch das Training auf dem *ImageNet*-Datensatz erhalten hatten. Um das Netzwerk auf die Klassifikation von Infektionsgraden an Möhrenblättern zu spezialisieren, fügte man an das Ende des Netzwerks eine lineare Schicht mit 1280 Eingangsdaten und 556 Ausgangsdaten, eine Dropout-Schicht mit einer Dropout-Rate von 0,2 und noch eine lineare Schicht mit 556 Eingangsdaten und 9 Ausgangsdaten (eine für jedes Infektionsstadium) hinzu. Im folgenden Training wurden somit ausschließlich die Gewichte der manuell hinzugefügten Schichten optimiert werden, was enorme Rechenleistung einsparte und die Trainingszeit reduzierte. Die Dropout-Schicht nullt dabei bei jeder Trainingsiteration zufällig 20% der Neuronen (schaltet sie ab), was Co-Adaption zwischen den Neuronen vorbeugt und somit *Overfitting* verringert. Für das Training des *EfficientNet-V2-L* wurden je 30 Bilder (vorverarbeitet wie oben beschrieben) aus dem *Entwicklungsdatensatz* von infizierten Möhrenblättern pro Infektionsstadium genutzt. Weiterhin wurden die Daten in 70% Trainings- und 30% Testbilder aufgeteilt, sodass für jedes Infektionsstadium 21 Trainings- und 9 Testbilder vorhanden waren. Um ein Training auf dem gegebenen System (*Geforce RTX 3070* Grafikkarte, *Nvidia*) durchführen zu können wurde die Größe der Bilder auf eine Größe von 224x224 Pixeln skaliert, wodurch kleine Läsionen für das Netzwerk nicht mehr erkennbar waren. Das Aufteilen des Bildes in mehrere kleine Segmente war nicht möglich, da die Referenzbewertungen immer für ein gesamtes Blatt galten, und nicht auf Teilstücke des Blattes übertragbar sind. Es wäre somit für jedes Bildsegment eine separate Referenzbewertung nötig, welche nicht vorhanden war. Als Verlustfunktion wurde *Cross Entropy* (CrossEntropyLoss-Funktion, Pytorch) und als Optimierungsalgorithmus für die *Backpropagation* der *Stochastic Gradient Descent* (SGD-Funktion, Pytorch) mit einer initialen Lernrate von 0,001 und einem Momentum von 0,9 genutzt. Alle 5 Epochen wurde die Lernrate im Training mit einem Gamma von 0,1 angepasst. Alle 5 Epochen verringerte sich die Lernrate somit um 1/10. Das Training dauerte maximal 20 Epochen und es wurden jene Gewichte gespeichert, welche innerhalb des Trainings die höchste Genauigkeit an den Testdaten erzielten. Hierfür erfolgte eine Evaluierung des Trainingserfolgs mit den Testdaten nach jeder Epoche. Zur Validierung nutzte man den oben beschriebenen *Validierungsdatensatz*. Bevor man die Bilder in das trainierte Netzwerk einspeiste, skalierte man auch diese auf 224x224 Pixel. Das Netzwerk generierte für jedes analysierte Bild eine Liste von 9 Konfidenzwerten (für jedes Infektionsstadium), sodass sich die finale Vorhersage aus jenem Infektionsstadium mit dem höchsten Konfidenzwert ergab.

3.4.3.2 Ergebnisse

Für die Bewertung des Infektionsstadiums durch *Alternaria dauci* wurden 3 verschiedene MS-Verfahren getestet. Um diese besser vergleichen zu können, wurde eine einheitliche Darstellungsweise gewählt. Hierfür wurden die Vorhersagen der verschiedenen Modelle auf die Skala der Züchterbonituren transformiert. Damit handelt es sich bei den folgenden Analysen um ordinale Regressionen, welche sowohl Charakteristika der Regression, aber

auch der Klassifikation aufweisen. Um die Leistung der unterschiedlichen Verfahren zu bewerten, wurden sowohl Klassifikations- wie auch Regressionsstatistiken verwendet.

Für das Schwellenwertverfahren wurde der prozentuale Anteil der grünen Blattpixel an der gesamten projizierten Blattfläche mittels eines polynomischen Regressionsmodells 2. Grades in die Züchterskala transformiert. Die entstandenen Gleitkommazahlen wurden dabei durch Runden in ganze Zahlen umgewandelt. Wie in Abbildung 24a dargestellt, lag nach dieser Umwandlung die Korrektklassifikationsrate bei 46,4%, der mittlere absolute Fehler bei 0,67 Infektionsstadien und der Korrelationskoeffizient zwischen vorhergesagtem und züchterischem Infektionsstadium bei 0,93. Tendenziell liegen die Vorhersagen mittels Grünpixelanalyse eher unter den Bewertungen des Züchters, es wird also im Vergleich eine stärkere Infektion vorhergesagt. Die mittleren absoluten Fehler der einzelnen Infektionsstadium reichen von 0,19 für das Infektionsstadium 1 bis 1,22 für das Infektionsstadium 6, Unter nicht Beachtung des Infektionsstadium 9, welches einen MAE von 1 besitzt, nimmt der MAE beginnend von Infektionsstadium 6 hin zu den Rändern (Infektionsstadium 1 und 8) kontinuierlich ab.

Für die Vorhersagen des RF-Modells wurden eine Korrektklassifikationsrate von 49,5%, ein mittlerer absoluter Fehler von 0,64 Infektionsstadien und ein Korrelationskoeffizient zwischen vorhergesagtem und züchterischem Infektionsstadium von 0,91 erreicht (Abbildung 24b). Eine Tendenz des Modells, das Infektionsstadium stärker oder schwächer zu bewerten im Vergleich zu der Züchterbewertung ist nicht vorhanden. 56 Vorhersagen des Infektionsstadiums fallen im Vergleich zur Züchterbewertung schwächer aus (die Punkte liegen über der Identitätslinie), 41 Vorhersagen fallen stärker aus (die Punkte liegen unter der Identitätslinie). Die MAEs der einzelnen Infektionsstadien liegen zwischen 0,12 für Stadium 9 und 1,06 für Stadium 4. Tendenziell sind sie für die mittleren Infektionsstadien höher wie für die Extreme.

Für die Vorhersagen des *EfficientNet-V2* -Modells wurden eine Korrektklassifikationsrate von 43,8%, ein mittlerer absoluter Fehler von 0,73 Infektionsstadien und ein Korrelationskoeffizient zwischen vorhergesagtem und züchterischem Infektionsstadium von 0,9 erreicht (Abbildung 24c). 73 der insgesamt 192 Vorhersagen des Infektionsstadiums fallen im Vergleich zur Züchterbewertung schwächer aus (die Punkte liegen über der Identitätslinie), 35 Vorhersagen fallen stärker aus (die Punkte liegen unter der Identitätslinie). Die MAEs der einzelnen Infektionsstadien liegen zwischen 0,12 für Stadium 9 und 1,09 für Stadium 4. Tendenziell sind sie für die Mittleren Infektionsstadium höher wie für die Extreme.

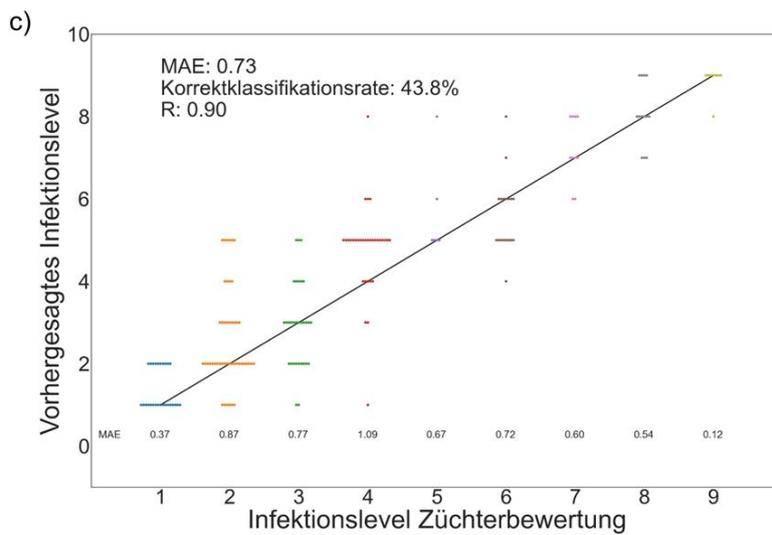
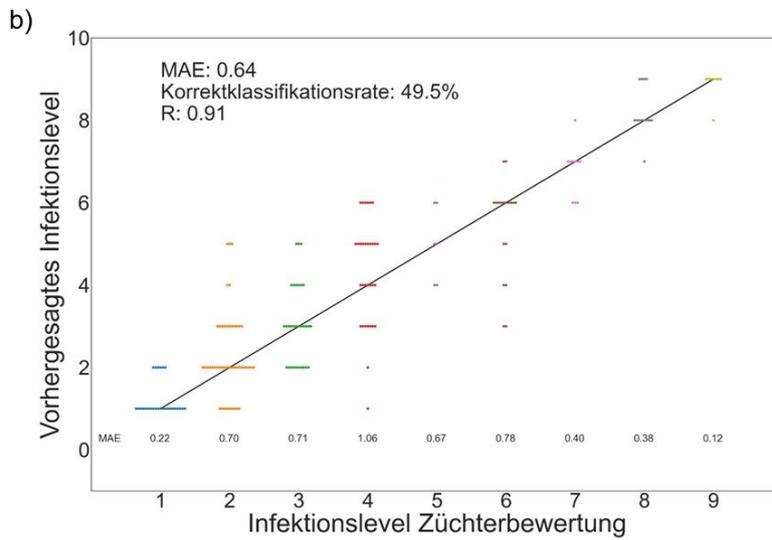
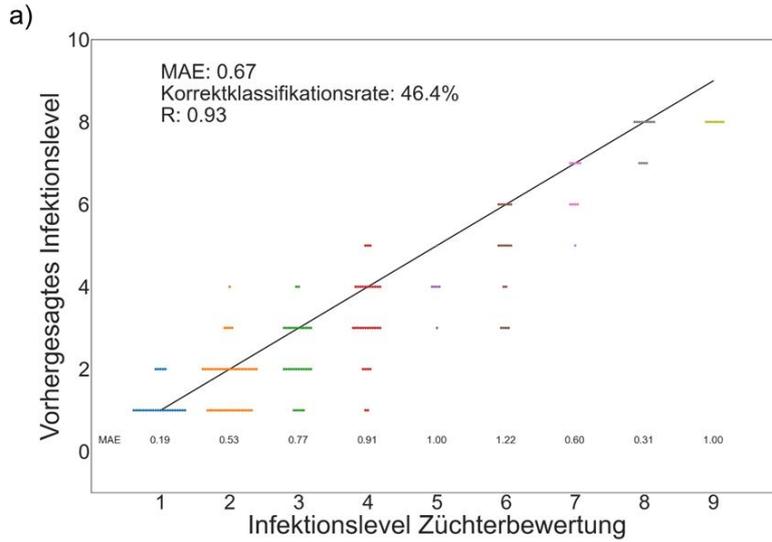


Abbildung 24: Vergleich der visuellen Bewertung des Infektionsstadiums einzelner Möhrenblätter mit der Vorhersage durch eine MS-Segmentierungsmethode im HSV-Farbraum (a), einer Random Forest Klassifikation (b) und der Vorhersage eines EfficientNet V2 (c). Angegeben sind jeweils der mittlere absolute Fehler (MAE) für die einzelnen Infektionsstadien, sowie für die gesamten Daten, der Korrelationskoeffizient R und die Korrektklassifikationsrate.

3.4.3.3 Diskussion

In dieser Studie sollten verschiedene MS-Methoden zur Quantifizierung des Befallsgrads von Möhrenblättern mit *Alternaria daucii* miteinander verglichen werden, um das Potenzial einer kostengünstigen Methode basierend auf RGB-Bildern zu evaluieren. Als Referenz dienten Bonituren des Züchterexperten. Hierfür wurde ein Schwellenwert-Segmentierungsverfahren im HSV-Farbraum mit Verfahren des maschinellen und tiefen Lernens verglichen. Ersteres bildet nur einen Aspekt der Züchterbewertung (Anteil infizierten Pixel) ab, während die anderen beiden Verfahren das Potential besitzen auch komplexere Bewertungs-Merkmale (z.B. Cluster von infizierten Pixeln) zu erkennen. Von der Nutzung von neuronalen Netzwerken zur semantischen Segmentierung der infizierten Blattareale, wie sie beispielsweise in Gao et al. (2021) zur Bewertung des Krankheitsverlaufs von späten - und damit einfach zu annotierenden - Infektionsstadien der Kraut- und Knollenfäule (*Phytophthora infestans*) an Kartoffelpflanzen zur Anwendung kamen, wurde bewusst abgesehen. In unserer Studie wurden sämtliche Infektionslevel an den Möhrenblättern betrachtet. Vor allem der Beginn der Möhrenschwärze zeichnet sich diese durch winzig kleine Läsionen aus, die im Bild nur wenige Pixel groß sind. Diese breiten sich erst im Verlauf der Krankheit aus und schließen sich flächig zusammen. Eine Annotation der infizierten Flächen im Bild stellt somit eine kaum zu bewältigende Aufgabe dar, da an einem einzelnen infizierten Blatt bereits hunderte Läsionen auftreten können. Weiterhin ist die Zuordnung der Pixel in gesund oder infiziert sehr schwierig, da die Mischpixel⁶ anteilmäßig sehr zahlreich sind. Aus diesem Grund nutzte man für das Verfahren des tiefen Lernens ein Klassifikationsnetzwerk, um das Infektionslevel einzelner infizierter Möhrenblätter vorherzusagen.

Wie in Abschnitt 3.4.2.1.3 beschrieben, fallen neben den Anteil der infizierter Blattfläche auch noch weitere Parameter in die Züchterbewertung des Infektionsgrades mit ein, deren Einfluss mit Hilfe der RGB-Verfahren teilweise erfasst werden kann: Zunächst haben die 3 Verfahren gute Korrelationswerte, jedoch schwächere Korrektklassifikationsraten. Tendenziell schneidet jedoch das RF-Modell mit einem MAE von 0,64 Klassen am besten, und das neuronale Faltungsnetzwerk mit einem MAE von 0,73 Klassen am schwächsten ab. Zunächst legen die Ergebnisse nahe, dass der Anteil der infizierten Blattfläche den größten Faktor bei der Evaluierung der Infektionsstadien durch den Züchters ausmacht, wie der hohe Korrelationskoeffizient von 0,93 für die Segmentierungsmethode zeigt. Segmentierungsmethoden stellen eine gängige Praxis in der Evaluierung der Infektionsstadien von Blättern dar und wurden beispielsweise auch in Goodwin & Hsiang (2010) genutzt, um Mehltauinfektionen an Blättern der Hohen Flammenblume (*Phlox peniculata*) zu quantifizieren. Jedoch war der Segmentierung-Ansatz in unserer Studie nicht in der Lage, die Infektionslevel 8 und 9 zu separieren. Gründe dafür könnten sein, dass Läsionen im Bild von Teilen des Blattes verdeckt wurden oder einfach zu klein waren, um im Bild aufgelöst werden zu können. Die Züchterskala ist nicht-linear und gerade für diesen Bereich besonders fein graduiert, sodass ein Blatt mit einer einzigen Läsion bereits dem Infektionslevel 8 zugeordnet wird. Die Segmentierungsmethode linearisiert die Züchterbewertungen über eine quadratische Regression, kann die Klasse 9 jedoch nicht gut approximieren, da diese keinen Toleranzbereich besitzt und sozusagen die untere Grenze der Klasse 8 darstellt. Dieses Phänomen spielt für das RF-Modell und das neuronale Klassifikationsnetzwerk jedoch keine Rolle, da diese die Nicht-Linearität der Züchterbewertungen nicht über eine Regressionsfunktion approximieren. So zeigen diese beiden Verfahren eine klarere Abgrenzung der Klassen 8 und 9. Weiterhin unterstützen die Ergebnisse des RF-Modells die Aussage des Züchters, dass nicht ausschließlich die Infektionsfläche in die Bewertung des

⁶ Mischpixel: Übergangspixel zwischen unterschiedlichen Farben oder Helligkeiten in einem Bild, welche durch die Bayer-Interpolation (demosaiicing) zustande kommt, mit der ein RAW-Bild in ein RGB-Bild umgewandelt wird.

Infektionslevels einfließt. Zwar war der Korrelationskoeffizient mit 0,91 etwas schwächer als für die Segmentierung, jedoch sank der MAE um 0,04 und die Korrektklassifikationsrate stieg um 3,1% bei dem RF-Modell. Dass das neuronale Netzwerk die schwächsten Resultate erzielte, könnte an der begrenzten Rechenleistung des genutzten Systems liegen. Die Referenzbewertungen des Zuchtexterten erfolgte auf der Grundlage des kompletten Blattmaterials, welches dann durch eine Aufnahme erfasst wurde. Durch die technischen Limitierungen musste für das Training die Auflösung der Bilder stark reduziert werden (224*224 Pixel), wodurch wichtige Details für das Netzwerk verloren gingen. Für alle 3 Methoden ließ sich der Effekt beobachten, dass die MAEs für die mittleren Infektionslevel tendenziell größer waren wie für die Randbereiche (Infektionslevel 9 beim Segmentierungs-Ansatz ausgelassen). Dem könnte die Natur der Züchterskala zugrunde liegen, welche an den Randbereichen (Infektionslevel 1 und 9) sehr klar definiert ist (1: kein grünes Blattmaterial mehr vorhanden; 9: Keine Läsionen vorhanden). Somit fällt eine eindeutige Bewertung mittlerer Infektionslevel immer schwerer, je mehr man sich von den klar definierten Randbereichen entfernt.

3.4.4 Bewertung chromatischer Pflanzeigenschaften unter Feldbedingungen

3.4.4.1 Material und Methoden

3.4.4.1.1 Modellpflanzen

Die Forschungsfrage befasst sich mit der präzisen Farbkalibrierung von Bildern, welche unter herausfordernden Feldbedingungen aufgenommen wurden. Es sollen hier vier Interpolationsmethoden miteinander verglichen werden, und schließlich ein Bezug zu einer praktischen Anwendung hergestellt werden. Als Modellpflanze wird hier der Weißkohl (*Brassica oleracea* convar. *capitata* var. *alba*) genutzt, da es sich hierbei um eine sehr empfindliche Kultursorte handelt, welche besondere Ansprüche an den Boden und das Klima stellt und Sortenbedingt, aber auch Stressbedingt vielerlei Verfärbungen wie beispielsweise Chlorosen und Nekrosen, aber auch lila Anthocyan Verfärbungen ausprägt. Des Weiteren bildet der Weißkohl je nach Genotyp und Umwelteinfluss verschiedenste Grundfarben des Blattwerks wie grün, grünlich-weißlich, oder grünlich-bläulich aus, was für Farbanalysen mittels RGB-Bildgebung besonders interessant ist.

3.4.4.1.2 Pflanzenanzucht

Die Kohlpflanzen wurden am Standort Marne angebaut. Der Standort Marne im Kreis Dithmarschen im Norden Deutschlands zeichnet sich durch seine Nähe zur Nordsee aus. Hier finden sich fruchtbare Marschböden und topographisch liegt das Gebiet auf der Höhe des Meeresspiegels. Während der Anbauperiode des Kohls zwischen Mai und Oktober herrschen Tagesdurchschnittstemperaturen zwischen 10°C und 17°C und es fallen im Schnitt zwischen 38,4 und 57,8mm Niederschlag in diesen Monaten, wobei der Oktober der regenreichste Monat im Jahr ist (Quelle: Wetterstation Elpersbüttel; <https://weatherspark.com/>, 09.03.2023). Die Kohlpflanzen wurden dort nach konventionellen Praktiken angebaut und bei Bedarf gewässert, um Stressfaktoren wie Krankheit, oder Mangelernährung vorzubeugen. Einige Pflanzen zeigten nichtsdestotrotz Stresssymptome wie Anthocyanverfärbung (Abbildung 25a). Zur Erhebung der Daten wurde kein gesondertes Experiment aufgestellt, da geeignete Pflanzen und Sorten aus den Beständen des Projektpartners *Rijk Zwaan* für die Vermessung und Bildgebung entnommen werden konnten. Insgesamt wurden hier 8 Sorten mit jeweils 10 Replikaten pro Sorte analysiert. Unter den untersuchten Sorten waren 4 mit runder Kopfform (R1-R4), unter denen eine Sorte dem Wirsing (*Brassica oleracea* convar. *capitata* var. *sabauda*) und 3 Sorten dem Weißkohl (*Brassica oleracea* convar. *capitata* var. *alba*) zuzuordnen waren. Des Weiteren gab es 2 Weißkohlsorten mit spitzer Kopfform (P1, P2) und 2 Sorten mit flacher Kopfform (F11, F12). Somit beinhaltete dieser Datensatz Bilder von 80 einzelnen Pflanzen.

3.4.4.1.3 Erhebung der Referenzdaten

Es mussten keine separaten Referenzdaten erhoben werden, da sich die Validierung aus den wahren Farbwerten des Referenzpanels ergeben.

3.4.4.1.4 Erstellung der Bildaufnahmen

In der Feldsaison 2022 wurden Aufnahmen von ausgewählten Kohlsorten aus den laufenden Feldexperimenten des Züchters *Rijk Zwaan Marne* aufgenommen. Man nutzte hierfür eine Digitalkamera, um die insgesamt 80 Kohlpflanzenbilder zu erstellen. Die Daten wurden mit einer *Sony Alpha 6400* (*Sony Group Corporation*, Tokyo, Japan) Digitalkamera in Kombination mit einem *Viltrox 23mm f1,4 E* (*Shenzhen Jueying Technology*, Shenzhen, China) Objektiv mit automatischem Fokus aufgenommen. Die Kombination aus dem Advanced Photo System type-C (APS-C) CMOS-Sensor der Kamera und dem Objektiv erzielte eine effektive Brennweite von 35mm. Die Kamera besitzt eine Auflösung von 20,3 Megapixeln bei einem Seitenverhältnis von 16:9, Die Belichtungszeit wurde auf 1,5ms (1/640s) eingestellt, die Blendenzahl auf F/5,6 und der ISO-Wert (das Pendant zum Gain) und der Weißabgleich auf automatisch belassen. Die Bilder wurden als unkomprimierte .PNG Dateien abgespeichert. Die Aufnahmen der Pflanzen wurden am 30. August 2022 auf den Versuchsflächen des Projektpartners in Marne, Norddeutschland durchgeführt. Es wurden jeweils 2 Aufnahmen pro Kohlpflanze angefertigt. Zunächst wurde eine Aufsicht Aufnahme gemacht, indem die Kamera mit etwas Abstand zum Körper auf Brusthöhe über die Kohlpflanze gehalten wurde (Abbildung 25a). Da durch dieses Verfahren die Aufnahmehöhe variieren konnte, nutzte man den *ColorChecker Passport Photo 2* als Referenzobjekt zu Umrechnung der Pixelwerte in das metrische System. Der *ColorChecker* besaß im geschlossenen Zustand eine Länge von 12,5cm und eine Breite von 9cm und wurde flach an einem Stab befestigt, um ihn genau auf der Höhe auszurichten, in der der Kohlkopf seine größte Ausdehnung hatte. Die zweite Aufnahme wurde nicht vom Kohl, sondern einem selbstentwickelten Farbpanel angefertigt (Abbildung 25b), welches als Referenz für die Farbanalyse in den Bildern diente.



Abbildung 25: Aufnahme einer Kohlpflanze (a) zusammen mit der Aufnahme eines selbstentwickelten RHS-Farbpanels (b), um nachträgliche eine präzise Farbkalibrierung der Aufnahmen durchführen zu können.

3.4.4.1.5 Vorverarbeitung der Bilder

Farbkalibrierung: Das Bild der Kohlpflanze wurde nicht vorverarbeitet. Das Bild mit dem Farbpanel wurde unter Nutzung von *GIMP 2.8* um 90° gedreht und wie in Abbildung 26 dargestellt zugeschnitten.

3.4.4.1.6 Evaluierung der Farbe unter Feldbedingungen mittels 2D RGB-Verfahren

Um eine Aussage darüber treffen zu können, ob verlässliche Farbanalysen der Bildern mit erhöhtem Durchsatz durchführbar sind, wurde die Farbanalyse mit chromatisch transformierten und nicht-transformierten Kohlbildern durchgeführt und verglichen. Mit dieser Methode galt es zu untersuchen, wie groß die qualitativen Unterschiede zwischen transformierten und nicht-transformierten Bildern waren.

Zur Kalibrierung der chromatischen Transformation wurden verschiedene Verfahren im LAB-Farbraum getestet (welche später erläutert werden), da dieser Farbraum wie in Abschnitt 2.2.3 beschrieben zu den wahrnehmbaren Farbräumen zählt und gemessene und wahrgenommene Abstände zwischen Farben im Vergleich zu anderen Farbräumen gut übereinstimmen (Hastings & Rubin, 2012)(Hastings & Rubin, 2012)(Hastings & Rubin, 2012). Es wurde daraus geschlussfolgert, dass dies für die Farbkorrektur der Bilder der geeignetste Farbraum ist. Es wurden eine lineare Interpolation, eine polynomische Interpolation 2. Grades, eine PLS-Interpolation sowie die zuvor vorgestellte 3D TPS-Interpolation angewandt (Menesatti et al., 2012). Alle Methoden wurden zunächst mit einem *ColorChecker Passport Photo 2*, welcher 18 Farbfelder und 6 Graufelder besitzt, getestet. Da hier insbesondere in Hinblick auf das Spline-Interpolationsverfahren keine zufriedenstellenden Ergebnisse erzielt werden konnten wurde ein erweitertes Farbpanel mit 144 Farbfeldern entwickelt, da diese Methode besonders von einer engmaschigen Abdeckung des Farbraums profitieren sollte. Dieses Panel setzte sich aus genau definierten Farben des *Colour Charts Edition V (Royal Horticultural Society, London, Vereinigtes Königreich)* zusammen und wird im folgenden *RHS-Farbpanel* genannt (Abbildung 26). Das *Colour Charts Edition V* ist eine Standardreferenz aus insgesamt 920 Referenzfarben, welche von Pflanzenzüchtern weltweit eingesetzt werden, um Farben von Blättern, Blüten, Früchten und anderen Pflanzenorganen genau zu bestimmen und einheitlich zu kommunizieren. Auf eine genaue Beschreibung, welche Farben für das *RHS-Farbpanel* ausgewählt wurden, wird an dieser Stelle verzichtet, im Detail können diese aber im folgenden GitHub nachgeschaut werden: https://github.com/djollet/RHS-Color-Correction-Python/blob/main/patch_data.json. Dort werden die Farben im Querformat reihenweise mit RHS-Farbbezeichnung und dem dazugehörigen HEX-Code in der *patch_data.json* Datei angegeben. Den Startpunkt für die reihenweise Aufzählung der Farbwerte bildet der große QR-Code auf dem *RHS-Farbpanel*. Bei der Farbauswahl wurde besonders darauf geachtet, dass ein sehr breites Farbspektrum an Referenzfarben abgebildet wurde. Der HEX-Code der *Referenzfarbe* wurde zunächst in den RGB- und schließlich in den LAB-Farbraum übertragen. Die Ermittlung der medianen Farbwerte der Farbfelder im Bild geschah durch einen semi-automatischen Prozess, welcher in *Python* umgesetzt wurde. Nach Laden der .PNG-Bilddatei, einer Transformation in den LAB-Farbraum und einem Zuschnitt auf den wesentlichen Bereich des RHS-Farbpanels wurden die Positionen der QR-Codes durch Anklicken bestimmt. Dabei galt es folgende Reihenfolge einzuhalten: oben Links – oben Rechts – unten Links – unten Rechts. Unter Nutzung der Positionsdaten der Eckpunkte wurde ein virtuelles Raster über das RHS-Farbpanel gelegt, welches die Positionsdaten für jedes einzelne Farbfeld beinhaltete. Zur Bestimmung der medianen Farbwerte nahm man einen Kreis mit einem Radius von 15 Pixeln um die entsprechenden Positionen an, wodurch ausreichend viele Bildpunkte erfasst wurden, aber genügend Abstand zu den Rändern jeder Farbfläche bestand. Für jedes Farbfeld gab es somit eine *Referenzfarbe* und die mediane *Bildfarbe*.

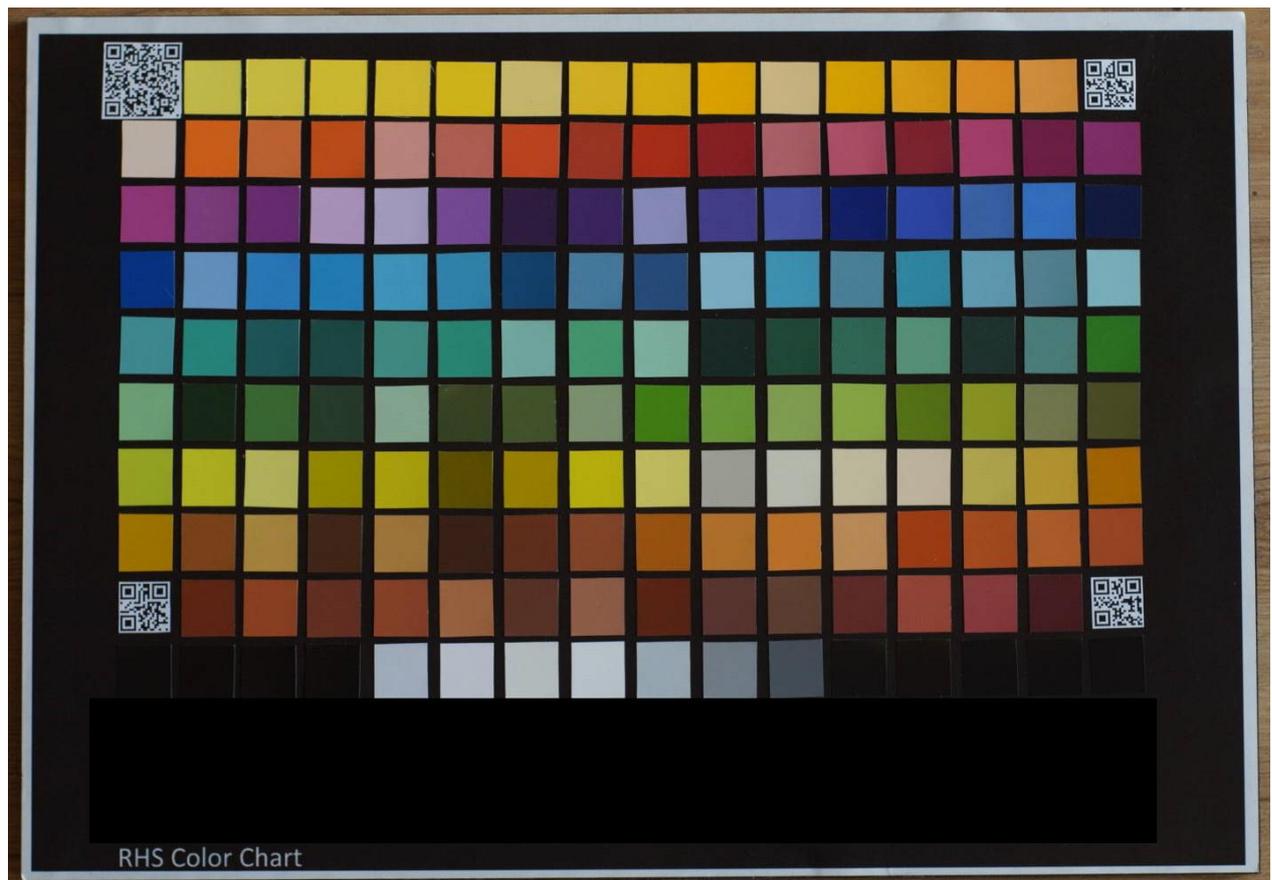


Abbildung 26: Eigenentwickeltes RHS-Farbpanel mit 144 Referenzfarben des Colour Charts Edition V der Royal Horticulture Society. Die wahren Farbwerte der einzelnen Farbfelder sind über einen Link aufrufbar, welche in dem großen QR-Code gespeichert ist.

Um die Funktionalität der Farbkalibrierung mittels *RHS-Farbpanel* und 3D TPS-Interpolation im LAB-Farbraum im Vergleich mit anderen Verfahren der Farbkalibrierung zu testen, wurde eine Kalibrierungsreihe aufgestellt. Für die Kalibrierungsreihe machte man zunächst verschiedene Aufnahmen des *RHS-Farbpanels* mit einer *Sony Alpha 6400* Digitalkamera in Kombination mit einem *Viltrox 23mm f1,4 E* Objektiv. Alle Aufnahmen wurden an einem sonnigen Tag im Freien zur Mittagszeit mit zusätzlicher Schattierung aufgenommen. Die Blendenzahl betrug F/16, der ISO-Wert 400 und die Belichtungszeit 20ms (1/50s), der Weißabgleich wurde jedoch in jedem Bild unterschiedlich eingestellt. Die Einstellungen der Farbtemperaturwerte der Kamera reichten zwischen 2500°K und 9500°K mit Schritten von 500°K, sodass insgesamt 15 Bilder mit variierendem Weißabgleich entstanden. Die Farbfelder des RHS-Farbpanels wurden im Folgenden in *Kalibrierungs-* und *Validierungsfarbfelder* eingeteilt, wobei jedes zweite Feld in der Zählung zu den *Validierungsfarbfeldern* gezählt wurde und alle verbleibenden zu den *Kalibrierungsfarbfeldern*. Mit Hilfe der *Referenz-* und *Bildfarben* der *Kalibrierungsfarbfelder* wurde nun für jedes der 15 Bilder Farbkalibrierungen mittels 3D TPS-Interpolation, sowie linearer Interpolation, polynomischer Interpolation 2. Grades und PLS-Interpolation durchgeführt. Im Vergleich zum Weißabgleich, der die Gewichtung des Blau- und Rotkanals während der Bayer-Interpolation (demosaiicing) bestimmt, gehören diese Verfahren zu den chromatischen Transformationen und verändern die Werte der Pixel im 3-dimensionalen Farbraum. Für die Kalibrierung wurden die Bilder und Referenzfarben zunächst in den LAB-Farbraum übertragen und normalisiert, sodass alle Werte zwischen 0 und 1 lagen. Anschließend wurden die entsprechenden Farbkalibrierungen durchgeführt und die Bilder anschließend für die weitere Verarbeitung wieder zurück transformiert und als .PNG Datei abgespeichert. Nach der Farbkalibrierung analysierte man die Bildfarben der *Validierungsfarbfelder* in den kalibrierten Bildern und berechnete die

euklidische Distanz der *Bildfarben* zu den *Referenzfarben* der *Validierungsfarbfelder* im 3-dimensionalen LAB-Farbraum.

Für die Farbanalyse der Kohlpflanzen wurden chromatisch transformierte Bilder mit nicht-transformierten Bildern verglichen. Als Transformationsverfahren nutzte man die 3D TPS-Interpolation im LAB-Farbraum, wie sie im vorherigen Abschnitt beschrieben wurde. Wie zuvor beschrieben, wurde direkt im Anschluss an jede Aufnahme einer Kohlpflanze eine Aufnahme des RHS-Farbpanels mit den gleichen Kameraeinstellungen gemacht, wodurch eine präzise Transformation für jedes Bild möglich gemacht werden sollte. Die Farbanalyse wurde wie folgt an den chromatisch transformierten und nicht-transformierten Bildern durchgeführt: Zunächst fertigte man in *GIMP 2.8* unter Nutzung des *Freie Auswahl-Werkzeugs* und des *Füllen-Werkzeugs* Bildmasken des Umblatts an, welche jedoch den Kohlkopf und den *ColorChecker* ausschlossen. Laut Züchter sei vor allem das Umblatt der Pflanzen für farbliche Bewertungen von Interesse. Die zu untersuchenden Bilder wurden in den HSV-Farbraum transformiert und mit Hilfe der Bildmaske wurden die Pixelwerte des Umblatts extrahiert. Bevor eine Zuordnung der Pixel in eine Farbklasse erfolgte, wurden jene Pixel von der Analyse ausgeschlossen, welche einen Helligkeitswert von kleiner als 50 oder größer als 240 besaßen, da solche Pixel keine präzise Farbwiedergabe, sondern ein starkes Rauschen in den Farbtonwerten (H-Kanal) besaßen. Die Zuordnung der Pixel in Farbklassen geschah über Schwellenwerte im H-Kanal. Pixel mit Werten ≥ 345 und < 15 erhielten die Zuordnung Rot, ≥ 15 bis < 45 Braun, ≥ 45 bis < 75 Gelb, ≥ 70 bis < 150 Grün, ≥ 150 bis < 195 Cyan, ≥ 195 bis < 255 Blau und ≥ 255 bis < 345 Magenta.

3.4.4.2 Ergebnisse

Wie die Kalibrierungsreihe in Abbildung 27a zeigt, konnte die Darstellung der Farben durch die Nutzung des *RHS-Farbpanels* im LAB-Farbraum mit allen genutzten Verfahren verbessert werden, was sich durch eine verminderte Distanz zwischen den medianen Bildfarben und den wahren Referenzfarben, sowie einer geringeren Streuung der Daten widerspiegelt. Dabei schnitt die *TPS-Interpolation* über das komplette Weißabgleichsspektrum am besten ab. Dies zeigt sich auch in der Detaildarstellung für die Kalibrierung bei 6500°K (Abbildung 27b). Bei dieser Weißabgleichseinstellung der Kamera konnte ohne weitere Farbkalibrierungen die beste Farbdarstellung mit einer medianen Distanz von 31 erreicht werden. Durch die Kalibrierungen konnte die mediane Distanz auf 7-10 weiter vermindert werden, wobei die *TPS-Interpolation* die geringste mediane Distanz aufwies.

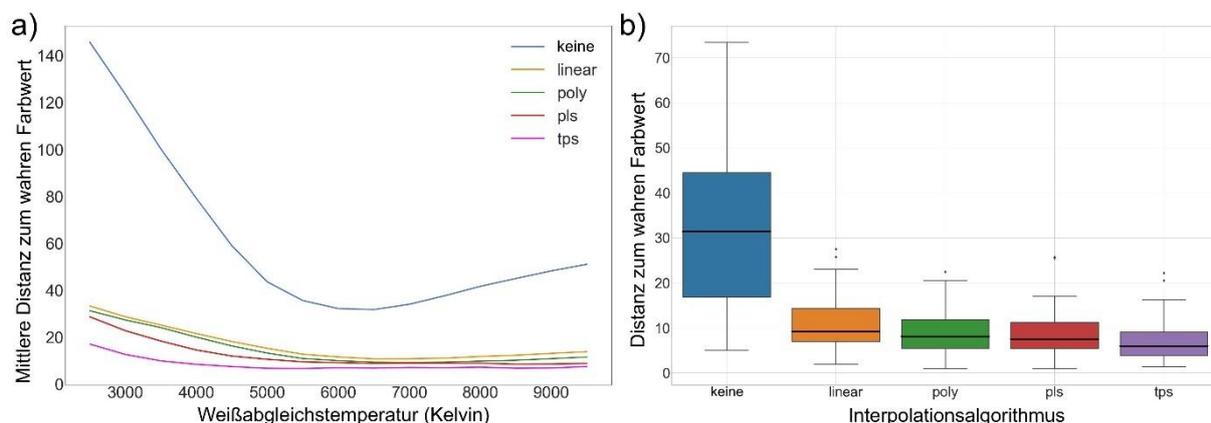


Abbildung 27: Gegenüberstellung der Ergebnisse der Farbkalibrierungsreihe mittels unterschiedlicher Interpolationsverfahren (a) und eine detaillierte Darstellung der Ergebnisse bei einer Weißabgleichstemperatur von 6500°Kelvin

Die Farbdaten der einzelnen Kohlbilder wurden mit der TPS-Interpolation transformiert, welches in der Kalibrierreihe die besten Ergebnisse geliefert hatte. Vergleicht man die Farbanalysen der chromatisch transformierten und nicht-transformierten Bilder, lassen sich nur für die Sorten P2 und R3 statistisch signifikante Unterschiede zwischen den addierten braunen und gelben Pixelmengen mittels einfaktorieller ANOVA-Analyse bei einem Signifikanzlevel von 0,05 feststellen (Abbildung 28). Braune und gelbe Pixelmengen wurden addiert, da diese zusammen eher für ungesundes Blattmaterial stehen und grüne Pixel für gesundes Blattmaterial. In Bezug auf die grünen Pixelmengen wurde nur für die Sorte FL2 ein statistisch signifikanter Unterschied gemessen ($p=0,037$). Für alle anderen Sorten sind die statistischen Unterschiede zwischen den addierten braunen und gelben Pixelmengen größer als für die grünen Pixelmengen. Die größten Unterschiede der addierten braunen und gelben Pixelmengen lassen sich für die Sorte P2 mit 7,3% feststellen. Im Mittel liegen die Abweichungen zwischen den kalibrierten und nicht kalibrierten Bildern für die gelben und braunen Farbanteile bei 2,9% und für die grünen Pixelmengen bei 2,7%. Für die blauen und Cyan Farbanteile lag der mittlere Unterschied zusammen bei 0,2%. Für den Magenta Anteil bei 0,05%.

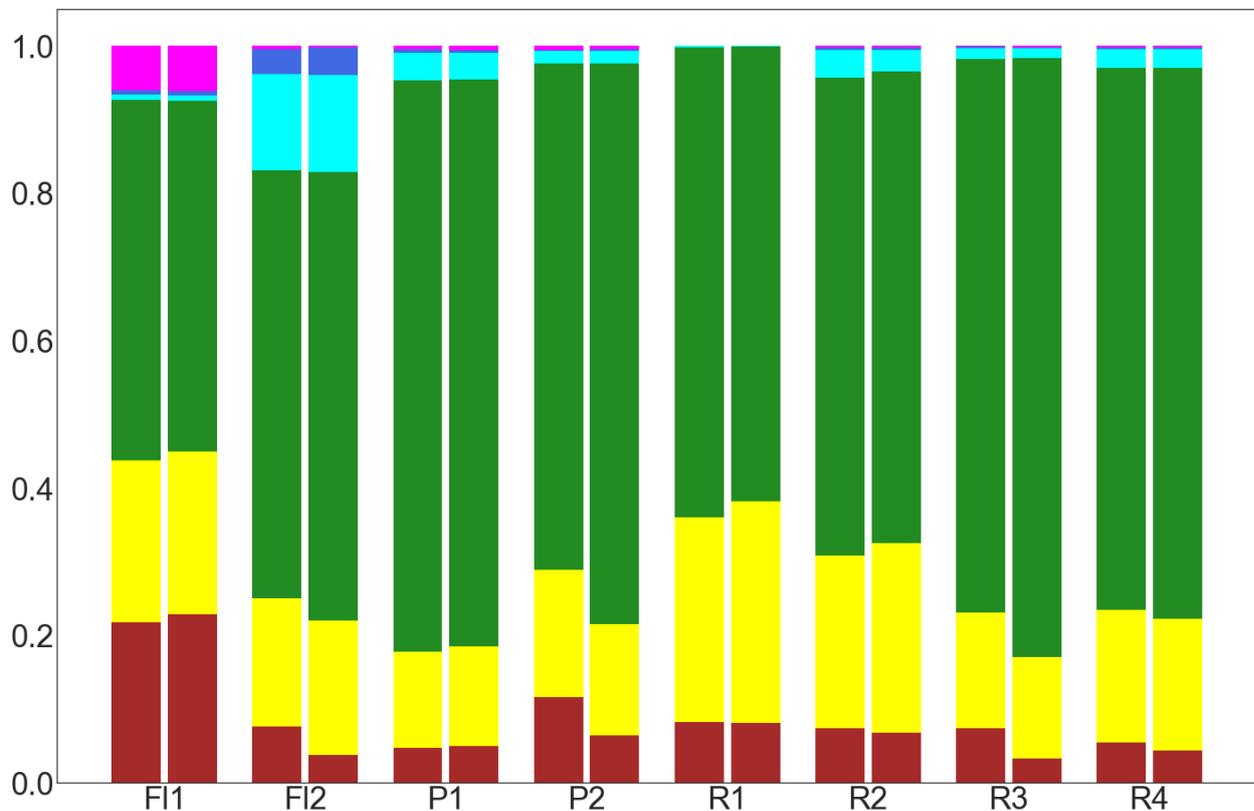


Abbildung 28: Ergebnisse der Farbklassifizierung für die Umblätter der einzelnen Kohlsorten ohne Farbkalibrierung (jeweils linkes Balkendiagramm) und mit einer TPS-Farbkalibrierung (jeweils rechtes Balkendiagramm). Die farbkodierten Abschnitte der Balkendiagramme geben die Anteile der entsprechenden Farbuordnung an ($n=10$ Pflanzen pro Sorte).

3.4.4.3 Diskussion

Die in dieser Studie durchgeführte Farbkalibrierung mittels selbstentwickeltem RHS-Farbpanel führte in Kombination mit allen untersuchten Methoden (linear, polynomisch, PLS, TPS) zu einer Verbesserung der Farbdarstellung (Reduzierung der medianen Entfernung zwischen Bild- und Referenzfarben und deren Streuung). Die besten Ergebnisse lieferte jedoch die lokale TPS-Interpolation, welches stark von der engmaschigen Farbabdeckung des RHS-Farbpanels profitierte. In der Menge an benötigten Referenzfarben liegt jedoch gleichzeitig auch der Nachteil dieses Verfahrens, da die Größe des *RHS-Farbpanels* hinderlich für die

praktische Nutzung im Feld ist. Bei einer Abdeckung des Farbraums mit wenigen Referenzfarben wie bei üblichen ColorCheckern mit 18 Referenzfarben neigt das TPS-Verfahren, welches auf lokaler Interpolation basiert, sehr leicht zum *Overfitting*, sowohl innerhalb des Gamuts, aber insbesondere auch außerhalb des abgedeckten Gamuts der Referenzfarben, wodurch fehlerhafte Farbtransformationen entstehen können, die sich als Farbanomalien im Bild bemerkbar machen. Somit erzielt die TPS-Interpolation in Kombination mit dem *RHS-Farbpanel* in diesem Versuch zwar die präzisesten Ergebnisse, allerdings sind die globalen Verfahren in Kombination mit weniger Referenzfarben für eine Nutzung im Feld gerade dann geeignet, wenn das Gamut des Referenzpanels die erfassten Farben nicht gut abdeckt, was auch durch die Resultate von Sunoj et al. (2018) gestützt wird: Sie nutzten 3 Referenzfarben (Rot, Grün, Blau), um mittels einer globalen Kalibrierungsmethode (lineare Regression) erfolgreiche Farbkalibrierungen unter Feldbedingungen durchzuführen. In Bezug auf die Farbanalysen der Kohlpflanzen in dieser Studie kann man jedoch keine generelle Aussage darüber treffen, ob eine Farbkalibrierung in der züchterischen Praxis sinnvoll ist. So konnten für lediglich 3 der 8 untersuchten Sorten Unterschiede zwischen den chromatisch transformierten und nicht-transformierten Bildern in Bezug auf die Darstellung der Braun-, Gelb- und Grüntöne festgestellt werden. Rottöne waren nicht vorhanden und für blau-, cyan- und magentafarbene Töne konnten keine signifikanten Unterschiede festgestellt werden. Die gemessenen Unterschiede könnten auch aus den Witterungsverhältnissen während der Datenaufnahme resultieren, da die Aufnahmen und Bonituren an einem wechselhaften Tag stattfanden und Teile der Daten unter voller Sonneneinstrahlung und anderer Teile während bewölkten Phasen aufgenommen wurden. So könnten die fluktuierenden Lichtbedingungen einerseits die Qualität der Ergebnisse negativ beeinflusst haben. Weiterhin konnte festgestellt werden, dass jene Sorten, welche die größten Unterschiede aufzeigten, unter voller Sonneneinstrahlung aufgenommen wurden. Für diese Aufnahmen traten innerhalb eines Bildes unterschiedlichste Lichtbedingungen auf, da es schattierte und nicht schattierte Bereiche durch benachbarte Pflanzen oder pflanzeneigene Blätter gab. Schattierungen durch umliegende Objekte wurden jedoch für die Aufnahme des RHS-Panels vermieden, sodass die Kalibrierung für schattierte Pixel in diesen Bildern der Kohlpflanzen wahrscheinlich fehlerbehaftet war. Weiterhin ist die Einteilung von 360 möglichen Farben (gemäß Farbtonkanal des HSV-Farbraums) der Pixel in 7 Farbklassen eine grobe Einteilung, sodass, sofern ein guter Weißabgleich für die Aufnahmen durchgeführt wurde, eine Farbkalibrierung nur einen geringen Einfluss auf das Gesamtergebnis haben kann. Hier sind vor allem Pixel in den Randbereichen der Farbklassen zu nennen, welche, je nach Farbkalibrierung, in die eine oder andere Farbkategorie rutschen können. Diese Bereiche machten in den hier untersuchten Bildern jedoch jeweils nur einen geringen Teil der gesamten Pflanze aus. Auch Divyashree et al. (2015) führten eine Farbklassifizierung an Kohlköpfen ohne Farbkalibrierung durch, um auf verschiedene Krankheiten rückzuschließen, wobei Bilder von den Kohlköpfen (ohne das Umblatt) ebenfalls unter Feldbedingungen aufgenommen wurden. Zu den genauen Aufnahmebedingungen ist jedoch nichts bekannt.

3.5 Automatisierte und nicht-invasive hochdurchsatz-Ertragsabschätzung bei kopfbildendem Blattgemüse

3.5.1 Forschungsstand

Die Ertragsabschätzung bei kopfbildenden Blattgemüsearten ist nicht trivial, da hierfür mehrere Parameter wie das Kopfvolumen und die Kopfdichte ausschlaggebend sind. Das Kopfvolumen seinerseits hängt von weiteren Formfaktoren ab, wie dem Durchmesser, aber auch der Form im vertikalen Schnitt. Naheliegender ist daher die Verwendung von 3D Bildgebungsverfahren, welche in mehreren Studien untersucht wurde (Guo et al., 2023; Kusumam et al., 2017). Alternative Methoden basierend auf multi-/hyperspektraler Information

kommen ebenfalls in Betracht (Ji et al., 2017; Kang et al., 2018). Im Bereich der 2D Bild-Verfahren wurde bereits im Jahr 1996 eine Methode zur Lokalisierung von Weißkohlköpfen (*Brassica oleracea* convar. *capitata* var *alba*) in Bildern entwickelt, die während der Erntezeit aufgenommen wurden. Hierzu wurde eine RGB-Kamera genutzt, welche von oben auf den Kohlkopf gerichtet war. Der Algorithmus beruhte auf der Erkennung von Schatten zwischen dem Kohlkopf und dem Umblatt, wodurch eine Grenzlinie definiert werden konnte, die den Kohlkopf abgrenzt und dessen Größe repräsentiert. 37 von 40 untersuchten Kohlköpfen wurden erfolgreich detektiert und es konnte eine Korrelation mit $R=0,78$ zwischen dem manuell gemessenen und dem errechneten Durchmesser der Kohlköpfe erzielt werden (Hayashi & Sakaue, 1996b). Kurz darauf erweiterten sie ihre Studie auf 130 Kohlköpfen 4 unterschiedlicher Sorten und entwickelten eine neue Methode. Hierzu trafen sie zunächst die Annahme, dass die Kohlköpfe aus der Ansicht von oben kreisrunde Objekte darstellen. Sie nutzten nun eine Diskriminanzanalyse, um das Bild zu segmentieren und kreis-basierte Hough-Transformationen, um den Kohlkopf im Bild zu lokalisieren. Es wurden 61 von 130 Köpfen richtig detektiert und es konnte im Mittel über alle 4 Sorten ein Korrelationskoeffizient von $R=0,75$ zwischen dem manuell gemessenen und dem errechneten Durchmesser erzielt werden. Der Nachteil dieser Methode bestand jedoch in der benötigten Verarbeitungszeit von 35-72 Sekunden pro Bild, was dieses Verfahren unbrauchbar für automatisierte Erntetechniken macht (Hayashi & Sakaue, 1996a). Ein Durchbruch in der Echtzeitdetektion von Kohlköpfen unter Feldbedingungen konnte 1997 durch die Nutzung eines künstlichen Neuronalen Netzwerkwerks realisiert werden. Dieses nutzte die Pixelwerte im HSV-Farbraum, um das Bild zu segmentieren und schließlich die Position und Größe des Kohlkopfes unter Zuhilfenahme zweier Kohlkopfvorlagen zu berechnen. Durch Optimierung des Algorithmus und umfangreiche Parallelisierung auf 8 CPU-Einheiten belief sich die Berechnungszeit auf 8,8 Sekunden für 11 Kohlköpfe, wobei insgesamt 70% der Kohlköpfe erkannt wurden (Murakami et al., 1997). Eine weitere Forschungsgruppe untersuchte nicht direkt den Ertrag, sondern das Wachstum von Weißkohlpflanzen im zeitlichen Verlauf: Hierzu installierten sie eine stationäre RGB-Kamera im Stil einer *PhenoCam* (Milliman, T. et al., 2019), um mittels Schwellenwert Farbsegmentierung im HSV-Farbraum den Grünflächenindex zu extrahieren. Jedoch beobachteten die Autoren in dieser Studie Limitierungen des Verfahrens unter wechselnden Lichtbedingungen im Feld (Low & Tay, 2019).

Des Weiteren wird seit einigen Jahren auch der Einsatz von Drohnen für die Ertragsabschätzung im Feld vermehrt untersucht: Eine Forschungsgruppe stellte einen Zeitreihen-Datensatz zweier Blumenkohlfelder (*Brassica oleracea* var. *botrytis* L.) vor, welche 2 Jahre in Folge mittels georeferenzierter RGB-Drohnenaufnahmen digitalisiert wurden. Der Datensatz dient zur Entwicklung und Evaluierung von maschinellen Lernmethoden zur Klassifikation, Detektion, Segmentierung und ähnlichen Aufgaben. Der Datensatz besteht aus insgesamt 14000 Einzelpflanzen. Von insgesamt 740 Pflanzen gibt es zudem phänotypische Evaluierungen wie das Wachstumsstadium oder die Pflanzen- und Kopfgröße. Anhand des Datensatzes zeigten sie exemplarisch, wie man die Pflanzen mittels eines *Mask R-CNNs* in den Bildern lokalisierten und maskieren kann. Die Resultate verglich man mit manuell annotierten Daten, wobei die Pflanzen mit einem F1 Wert von 0,77 in den Bildern erkannt wurden (Kierdorf et al., 2022). Eine weitere Arbeitsgruppe nutzte Drohnenaufnahmen und neuronale Netzwerke um die Größe von Chinakohlpflanzen im Feld abzuschätzen. Dabei wurden 3 verschiedenen Ansätze miteinander verglichen: Erstens die Detektion der Köpfe mittels 5 kanaliger, multispektraler Bilder, welche die RGB-Bänder, sowie ein Band bei 730-770nm und ein Band bei 780-890nm enthielten. Zweitens die Detektion mit Hilfe der klassischen 3 kanaligen RGB-Bilder. Drittens die Detektion anhand von RGB-Bildern, deren Auflösung mittels Super Resolution Generative Adversarial Networks künstlich erhöht wurde. Zur Detektion und Segmentierung der Pflanzen wurde jeweils ein modifiziertes U-Net CNN genutzt. Dabei zeigte der dritte Ansatz die höchste Genauigkeit mit einer Erkennungsrate der

Pflanzen von 94,43% und einem Determinationskoeffizient von 0,84 für die Breitenmessung, bzw. 0,9 für die Längenmessung der Kohlpflanzen (gemessen an 19 Einzelpflanzen).

3.5.2 Material und Methoden

3.5.2.1 Modellpflanzen

Zur Beantwortung der Forschungsfrage wird der Kopfkohl (*Brassica oleracea convar. capitata* L.) als Modellpflanze genutzt, da diese Unterart verschiedenste Varietäten aufweist, und somit ein breites Spektrum an Phänotypen für die Untersuchung des Forschungsthemas zur Verfügung steht. Weiterhin gibt es auch innerhalb der Varietäten Weißkohl, Rotkohl und Wirsing ein breites Spektrum an Formen und Farben, was eine gute Übertragbarkeit der Ergebnisse auf andere kopfbildende Blattgemüsearten ermöglicht. So können Kopfkohlsorten beispielsweise runde, flache oder spitze Köpfe ausbilden, die Köpfe können unterschiedlich kompakt beziehungsweise dicht sein und der Strunk kann ebenfalls divers ausgeprägt sein. In Bezug auf den Ertrag sind dies wichtige Faktoren, welche durch ein 2D RGB-Verfahren jedoch nicht direkt gemessen werden können und dieses somit vor besondere Herausforderungen stellt. Der Kopfkohl eignet sich somit besonders gut, um die Möglichkeiten aber auch die Limitierungen des 2D Verfahrens für die automatisierte und nicht-invasive Ertragsabschätzung zu untersuchen.

3.5.2.2 Pflanzenanzucht

Die Pflanzenanzucht fand, wie im Abschnitt 3.4.4.1.2 beschrieben, auf den Versuchsflächen des Projektpartners *Rijk Zwaan* in Marne statt. Auch wurde der in diesem Abschnitt beschriebene Datensatz, welche 80 Kohlpflanzen umfasst, zur Beantwortung dieser Forschungsfrage genutzt und wird im Folgenden *Validierungsdatensatz* genannt.

Um das Thema der Automation umfangreich zu untersuchen, wurde ein gesondertes Feldexperiment angelegt. Insgesamt wurden auf einem Areal von ungefähr 500m² 360 Kopfkohlpflanzen angebaut, welche 6 unterschiedlichen Sorten zuzuordnen waren. Unter den 6 Sorten waren 5 Weißkohlsorten und ein Wirsing. Um die Variabilität in der Bodenbeschaffenheit des genutzten Feldes auszugleichen, wurden die einzelnen Sorten in Reihen hintereinander gepflanzt. Ein Versuchsglied bestand aus 3 verschiedenen Kopfkohlsorten mit je 20 Pflanzen pro Sorte, welche in Reihe gepflanzt und von einer Rotkohlreihe umschlossen wurden. Es gab 2 verschiedene Versuchsglieder, welche sich in den Kopfkohlsorten unterschieden, mit jeweils 3 Wiederholungen (Abbildung 29). Vor der Aussaat wurde das Feld 2-mal gedüngt: Am 31. März 2021 mit 500kg/ha Kaliumoxid und am 01.06. mit 600kg/ha Kalkammonsalpeter. Am 2. Juni erfolgte weiterhin eine Pflanzenschutzmaßnahme mit 1,6l/ha *Butisan S* (BASF, Ludwigshafen, Deutschland) gegen auflaufende Unkräuter. Am Aussaattermin am 24. Juni wurden nochmal 300kg/ha Kalkammonsalpeter zugeführt. Die Aussaat erfolgte maschinell in Ost-West Richtung mit einem Reihenabstand von 60cm und einem Pflanzabstand von 48cm. Am 7. Juli erfolgte eine Insektizidbehandlung mit 100 ml/ha *Coragen* (FMC Agricultural Solutions, Philadelphia, USA), 50 ml/ha *Karate* (Syngenta, Basel, Schweiz) und 1 l/ha *Hasten* (Victorian Chemical Company, Victoria, Australien). Eine weitere Gabe von 200kg/ha Kalkammonsalpeter erfolgte am 22. Juli und eine weitere Insektizidbehandlung mit 100 ml/ha *Coragen*, 0,25kg/ha *Pirimor-Granulat* und 150 ml/ha *Karibu* (*Certis Belchim*, Maarssen, Niederlande). Es wurden in der Feldsaison 2021 ungefähr 35000 Bilder über einen Zeitraum von 3 Monaten aufgenommen aus denen schließlich 3700 händisch für die weitere Verarbeitung ausgewählt wurden. Die Aufnahmen fanden jeweils um die Mittagszeit und bei trockenen Witterungsbedingungen an folgenden Terminen statt: 13. Juli, 27. Juli, 03. August, 10. August, 19. August, 24. August, 30. August, 10. September, 14. September, 23. September und am 19. Oktober. Dieser Datensatz wird im folgenden *Entwicklungsdatensatz* genannt und Referenzmessungen wurden für diesen Datensatz nicht erhoben.



Abbildung 29: Feldversuchsaufbau mit 6 unterschiedlichen Sorten, welche zur besseren Abgrenzung von Rotkohlpflanzen umschlossen wurden (eine Farbe pro Sorte).

3.5.2.3 Erhebung der Referenzdaten

Folgende Referenzmessungen wurden nach züchterischer Praxis an den 80 Kohlköpfen des *Validierungsdatensatzes* durchgeführt: Zunächst wurde das Gewicht des geernteten Kopfes mit einer Federwaage gemessen. Anschließend wurde der Kopf von der Spitze bis zum Strunk mittig aufgeschnitten, um eine Abschätzung für die Dichte des Kopfinneren zu machen, welche auf visuellen Eindrücken basierte und in Prozent angegeben wurde. Eine Sorte mit geringer Dichte erhielt beispielsweise einen Wert von 60%, während Sorten mit einer hohen Dichte Werte über 95% erhielten.

3.5.2.4 Erstellung der Bildaufnahmen

Die Erstellung der Bildaufnahmen für den *Validierungsdatensatz* wurde bereits in Abschnitt 3.4.4.1.4 genau beschrieben.

Die Bilder für den *Entwicklungsdatensatz* wurden wie folgt mit einem Field4Cycle aufgenommen (Abbildung 30a): Das Field4Cycle (nachfolgend Fieldbike genannt) stellt ein nicht-motorisiertes Feldfahrzeug dar, an welche verschiedensten Messinstrumente zur Datengenerierung installiert werden können (Muller et al., 2016). Der Name leitet sich von der ursprünglichen Konstruktionsweise ab: 2 Fahrräder mit versteifter Lenkung, welche über einen Stahlrahmen miteinander verbunden sind. Das genutzte Fieldbike stellte eine Weiterentwicklung dieses Prototyps dar (Muller et al., 2016), und zeichnete sich dadurch aus, dass 4 Fahrradreifen über eine Stahlkonstruktion, welche aus 3x3cm Vierkant-Stahlrohren und Flacheisen bestand, verbunden waren. Die Spurbreite des Gefährts betrug 3,50m, sodass insgesamt 5 Reihen Kohlpflanzen gleichzeitig mit dem Gefährt überfahren werden konnten. Dabei dienten die 2 äußeren Reihen lediglich dem Schutz der inneren 3 Reihen vor mechanischer Einwirkung durch die Reifen des Fieldbikes, da von diesen Zeitserienaufnahmen gemacht wurden. Die Höhe, in der drei industriellen *Prosilica GT 2460* (*Allied Vision*, Stadtroda, Deutschland) Kameras an dem Gefährt befestigt wurden, betrug 1,75m. Die Kameras besitzen eine Auflösung von 5,1MP und das Gehäuse ist mit Kühlrippen ausgestattet, welche diesen Kamerateyp besonders geeignet für Aufnahmen im Freien macht. Als Objektiv nutzte man das FL-CC1614-2M F1,4 f16mm 2/3" (*Ricoh*, Tokyo, Japan) mit einer Festbrennweite von 16mm. Der Fokus wurde an dem Objektiv auf die Montagehöhe der Kameras angepasst und die Blendenzahl wurde ebenfalls am Objektiv auf F/8 fest eingestellt, um eine ausreichende Tiefenschärfe zu erhalten, aber auch möglichst viel Licht passieren zu lassen. Die drei Industriekameras wurden parallel über eine selbstentwickelte C++ Applikation mit Benutzeroberfläche angesteuert, welche sich der Vimba C++ API (*Allied Vision*, Stadtroda, Deutschland) bediente. Konzeption, Entwicklung und Tests dieser Steuer-Software erfolgten in Zusammenarbeit mit einem Software-Entwickler. Die Benutzeroberfläche der Ansteuerungssoftware ist in Abbildung 30b dargestellt. Da die Software auch durch Personal des beteiligten Züchtungsunternehmens genutzt werden sollte, welches wenig Erfahrung mit der Nutzung von industriellen Kameras besaß, wurden die Einstellungsmöglichkeiten innerhalb der Benutzeroberfläche auf die wesentlichen Parameter reduziert. So konnte man in der Benutzeroberfläche die verbundenen Kameras auswählen und für diese mit Hilfe einer 18% Graukarte einen Weißabgleich durchführen. Der Weißabgleich wurde am Anfang jeder Inbetriebnahme durchgeführt, zusätzlich auch dann, wenn die Lichtbedingungen sich gravierend änderten (z.B. von bewölkt zu wolkenfrei). Um eine ausreichende Tiefenschärfe und gleichzeitig eine ausreichend kurze Belichtungszeit zu gewährleisten, die trotz der

Kamerabewegung während der Überfahrt Bilder ohne erkennbare Bewegungsunschärfe ermöglicht, wurden die Blendezahl und die Belichtungszeit fest eingestellt. Um dennoch eine gewisse Flexibilität je nach Beleuchtungsbedingung zu erreichen, wurde die Belichtung über den Gain nachjustiert. Die Belichtungszeit wurde auf 2000ms gesetzt, da somit in Verbindung mit der Gain-Einstellung für alle Belichtungsverhältnisse (klarer Himmel vs. Bewölkung) eine gute Ausleuchtung im Bild erreicht werden konnte. Die Framerate, also die Geschwindigkeit, mit der die Kameras Bilder aufnehmen sollten, wurde auf 3 Bilder/Sekunde festgelegt. Da das Gefährt manuell und daher langsam über das Feld geschoben wurde, gab es bei dieser Aufnahmezeit keinen Datenverlust in dem Sinne, dass der Kopf einer jeden Pflanze mindestens einmal relativ mittig im Bild abgelichtet wurde.



Abbildung 30: a) Feldbike während Aufnahmen am Standort Marne. Die drei Industriekameras werden über einen Akku betrieben und über ein Laptop gesteuert; b) Screenshot der Benutzeroberfläche zur parallelen Ansteuerung der Kameras. Die Software komplettierte die Funktionalität des Feldbikes und war auf maximale Benutzerfreundlichkeit ausgelegt, um auch von Nichtfachpersonal genutzt werden zu können. Im oberen Bereich befinden sich die Ansteuerungsfelder der drei Kameras und im unteren Bereich Einstellungsmöglichkeiten wie der Speicherpfad, das Dateiformat sowie den Gain zu Anpassung der Helligkeit im Bild.

3.5.2.5 MS-Methoden zu Ertragsabschätzung Vorhersage des Kohlkopfgewichts

Da das Kohlkopfgewicht durch 2D Bildgebung nicht direkt abgeschätzt werden kann, galt es einen funktionalen Zusammenhang zwischen Gewicht und verschiedenen Formeigenschaften des Kopfes zur Annäherung zu finden. Hierfür wurden drei Parameter identifiziert, welche auch in Hinblick auf eine mögliche Anwendung in der züchterischen Praxis ohne großen Aufwand erfasst werden konnten: Zunächst die projizierte Kopffläche, welche mit einer Kamera aus der Aufsicht auf die Kohlpflanze durch die nachfolgend dargestellten Verfahren abgeschätzt werden kann. Der zweite Parameter ist das vertikale Profil. Es wurde bewusst darauf verzichtet die projizierte Kopffläche unter Annahme einer kugelrunden Kopfform in ein Volumen umzuwandeln, da diese Annahme für flache und spitze Kohlköpfe nicht zutrifft. Das vertikale Profil ergibt sich aus der sortentypischen Kopfform (rund, flach, spitz). Diese fließt nicht direkt in die Gewichtsabschätzung mit ein, jedoch werden die Ergebnisse nach den drei Formen getrennt betrachtet. Zuletzt wurde die Dichte des Kopfes bestimmt, welche innerhalb einer Sorte sehr homogen ist (siehe Tabelle 1), sodass während einer Messreihe pro Sorte nur wenige Kohlköpfe für die Vermessung aufgeschnitten werden müssen. Auf weitere invasive Vermessungen, wie beispielsweise die Bestimmung der Strunklänge, wurde aus Gründen der daraus resultierenden Durchsatzminderung bewusst verzichtet, auch wenn der Strunk als besonders dichte Struktur wahrscheinlich einen Einfluss auf das Gewicht des Kopfes hat.

	R1	R2	R3	R4	FL1	FL2	P1	P2
Mittel	0,85	0,86	0,88	0,62	0,77	0,82	0,78	0,79
StAbw	0,04	0,03	0,02	0,04	0,03	0,04	0,03	0,04

Tabelle 1: Prozentuale mittlere Kohlkopfdichte nach Züchterbewertung. Die mittleren Dichten und Standardabweichungen (StAbw) sind für die 4 runde (R1-R4), zwei flache (FL1, FL2) und 2 spitze (P1, P2) Kohlsorten des Validierungsdatensatz angegeben (n=10/Sorte).

Die Vorhersage des Kohlkopfgewichtes wurde ebenfalls an den 80 Bildern des *Validierungsdatensatzes* getestet. Um herauszufinden, ob die 2-dimensionalen Bilddaten überhaupt ausreichend Informationen für eine solide Ertragsabschätzung bereitstellen, wurden zunächst unter Nutzung von GIMP 2.8, wie in Abschnitt 3.4.4.1.6 beschrieben, binäre Bildmasken von den Kohlköpfen der Pflanzen erstellt (Abbildung 31). Dabei wurde versucht sich dem runden Profil des Kopfes, wie es sich aus einer perfekten Aufsicht des Kopfes ergibt (in einer Flucht mit der vertikalen Achse des Kohlkopfes), anzunähern. Wurden also Teile des Kopfes verdeckt, oder lag ein spitzer Kohl schräg im Bild, wurde das runde Profil des Kopfes während der Annotation geschätzt. Die Fläche der Bildmasken wurde wie in Abschnitt 2.3.6 beschrieben berechnet und mit Hilfe eines Umrechnungsfaktors, welcher sich aus der Größe des *ColorCheckers* im Bild ergab, in das metrische System übertragen. Die projizierte Kopffläche wurde dem gemessenen Gewicht gegenübergestellt. Weiterhin wurde untersucht, ob die Genauigkeit des Verfahrens unter Berücksichtigung der Dichte des Kopfes erhöht werden kann. Dazu wurde die projizierte Blattfläche mit der Dichte des Kopfes multipliziert.



Abbildung 31: Manuell annotierte Masken für den Kohlkopf (rot) sowie für das Umblatt (grün) aus dem Validierungsdatensatz. Die Kopfmasken wurden für die Ertragsanalyse und die Umblattmasken für die Farbanalysen genutzt.

Automatisierung mittels Mask R-CNN

Voraussetzung für die Berechnung der Kopfmasse in einem automatisierten Ansatz mit gesteigertem Durchsatz ist eine zuverlässige automatische Detektion des Umblattes und des Kopfes. Hierfür trainierte und evaluierte man ein geeignetes neuronales Netz, das *Mask R-CNN* aus dem *Detectron2*-Framework. Die genutzte Version bestand aus einem *ResNet-Backbone* mit 101 Schichten, einem *Feature-Pyramid Network* und einer 3x Lernrate. Auch hier bediente man sich wieder der Methode des Transferlernens und initialisierte das Netzwerk mit den *Model Zoo* Gewichten des *Detectron2*-Frameworks⁷.

Training: Für das Training des *Mask R-CNN* wurden insgesamt 100 Bilder von Kohlpflanzen aus dem *Entwicklungsdatensatz* und 50 Bilder von Kohlpflanzen, welche in den Saison 2021 mit einer Handykamera auf den Versuchsflächen des Projektpartners aufgenommen wurden, genutzt. Bei der Auswahl der Bilder wurde darauf geachtet, dass ein breites Spektrum an Phänotypen enthalten war, was sich sowohl auf die Kopfform, die Farbe der Pflanzen als auch auf die Struktur der Blätter bezieht. In den 150 Bildern wurden sowohl die gesamte Pflanze wie auch der Kohlkopf mit Hilfe des *VGG Image Annotator* (Dutta & Zisserman, 2019)(Dutta & Zisserman, 2019)(Dutta & Zisserman, 2019) annotiert. Die Bilder wurden in 125 Trainings- und 25 Testbilder aufgeteilt. Für das Training wurden die Standardeinstellung des *Detectron2*-Frameworks für die Hyperparameter genutzt, außer dass die Batchgröße pro Bild auf 256 und die Lernrate auf 0,00025 angepasst und die Anzahl an Klassen auf 2 eingestellt wurden. Um den Trainingsprozess sowie die Vorhersagen des trainierten Netzwerks zu beschleunigen, wurden die Bilddimensionen auf 500*500 Pixel reduziert, was für eine präzise Kohlkopferkennung eine ausreichende Auflösung darstellt. Um die Vorhersagen des trainierten Netzwerks robuster zumachen, wurden folgende Augmentationen an den Bildern vorgenommen. Zunächst wurden mit einer Wahrscheinlichkeit von 50% von dem bereits reduzierten Bild ein Ausschnitt mit den Dimensionen 300*300 Pixeln generiert, dessen Zentrum im Zentrum des Originalbildes lag und nach dem Zuschnitt wieder auf eine Größe von 500*500 Pixeln hoch skaliert wurde. Diese Augmentation sollte das Netzwerk robuster gegenüber sehr großen Köpfen machen, welche vor allem bei Kohlköpfen des Typs *Flach* auftreten und im Trainingsdatensatz unterrepräsentiert waren. Die Bilder wurden zufällig gespiegelt und um 1 bis 360 Grad gedreht, um alle möglichen Ausrichtungen der Pflanzen zu simulieren. Sättigung, Helligkeit und Kontrast wurden nach dem Zufallsprinzip zwischen 0,75 und 1,25 eingestellt, um variierende Lichtbedingungen im Feld zu simulieren. Eine Evaluierung des Trainingserfolg mit dem Testdatensatz wurde alle 500 Epochen durchgeführt, wobei die Gewichte des Netzwerks gespeichert wurden, wenn die Verlustfunktion für den Testdatensatzes ein neues Minimum erreichte. Das absolute Minimum wurde nach 5000 Epochen erreicht.

Validierung: Zur Validierung der Vorhersagen des trainierten *Mask R-CNNs* wurden die 80 Kohlpflanzen aus dem *Validierungsdatensatz* genutzt, für welche die manuell annotierten Kohlkopfmasken, wie auch das gemessene Kohlkopfgewicht als Referenzmessungen vorlagen. Für die Vorhersagen des Netzwerks wurde eine Konfidenzschwelle von 0,5 verwendet, was der Standardeinstellung entspricht.

3.5.3 Ergebnisse

Vorhersage des Kohlkopfgewichts

Zunächst wurde der Zusammenhang zwischen der projizierten Fläche händisch annotierter Kohlkopfmasken und dem Gewicht des geernteten Kohlkopfes untersucht. Wie in Abbildung 32a dargestellt konnte für den gesamten Datensatz ein Korrelationskoeffizient von 0,89 und ein *MAPE* von 21,9% erzielt werden. Betrachtet man die Ergebnisse typspezifisch (Kopfformen

⁷ Detectron2 – Model Zoo:

https://github.com/facebookresearch/detectron2/tree/main/detectron2/model_zoo (20.02.2024)

Rund, Flach, Spitz), erhält man etwas schwächere Korrelationskoeffizienten zwischen 0,81 und 0,84. Für den Typ *Flach* wurde ein *MAPE* von 10,5%, für den Typ *Spitz* von 15,4% und für den Typ *Rund* von 25,2% erzielt. Durch die Multiplikation der projizierten Kopffläche mit der Dichte des Kopfes konnten die Ergebnisse für Sorten des Typs *Rund* deutlich verbessert werden (Abbildung 32b). Der Korrelationskoeffizient erhöhte sich für diesen Typ auf 0,90 und der *MAPE* sank auf 14,3%, was einer Verbesserung um 10,9% entspricht. Für Kohlköpfe des Typs *Flach* wurden ebenfalls leichte Verbesserungen erzielt (Abbildung 32b). Der Korrelationskoeffizient stieg um 0,04 und der *MAPE* sank um 1%. Für Köpfe des Typs *Spitz* verschlechterte sich das Ergebnis sehr leicht, da der Korrelationskoeffizient um 0,01 sank und der *MAPE* um 0,6% stieg (Abbildung 32b).

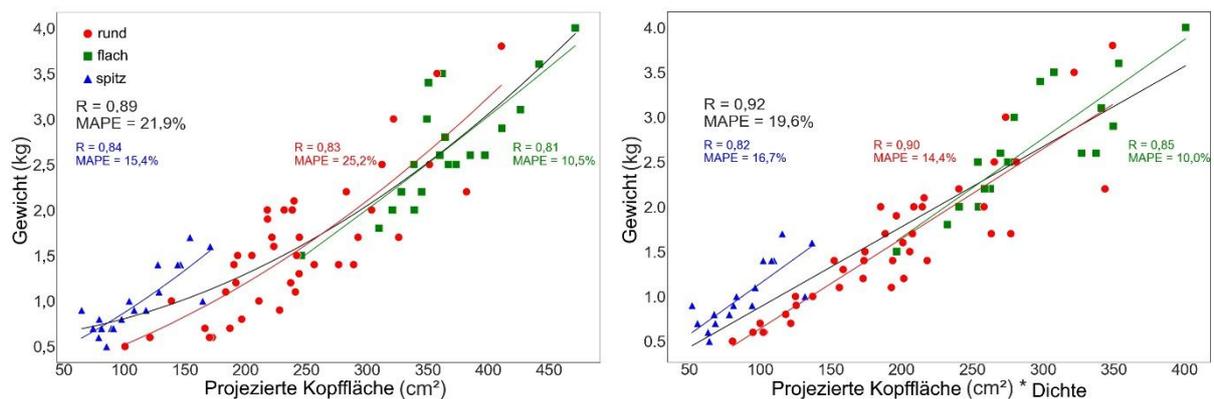


Abbildung 32: Gegenüberstellung des Kohlkopfgewichts zur a) manuell annotierten projizierten Kopffläche und b) manuell annotierten projizierten Kopffläche multipliziert mit der Dichte des Kopfes (nach Züchterbewertung); Die Daten sind nach den einzelnen Kopfformtypen *rund*, *spitz* und *flach* aufgeteilt und es wird jeweils der Korrelationskoeffizient *R* sowie der mittlere absolute Fehler in Prozent (*MAPE*) angegeben

Automatisierung mittels Mask R-CNN

Zur Automatisierung der Instanz-Segmentierung der Kohlköpfe im Bild wurde ein *Mask R-CNN* trainiert, welches für den Testdatensatz während des Trainings einen mAP_{50} von 84,07% erzielte. Insgesamt erkannte das Netzwerk 77 von 80 Kohlköpfen. Bei den 3 nicht-erkannten Köpfen handelte es sich um Fehlerkennungen (falsch-positiv), da das Netzwerk den *ColorChecker* im Bild als Kopf erkannte. Um die Präzision des Netzwerks tiefergehend zu untersuchen, wurden die vorhergesagten Kopfmasken des neuronalen Netzwerks mit den annotierten Masken auf Basis der Maskenflächen verglichen (Abbildung 33). Über den gesamten Datensatz betrachtet konnte zwischen den vorhergesagten und annotierten Maskenflächen ein Korrelationskoeffizient von 0,91 und ein *MAPE* von 22,1% erzielt werden. Betrachtet man die Lage der Datenpunkte zur Identitätslinie, liegen diese für flache Kohlköpfe eher unterhalb dieser. Die Datenpunkte der runden und flachen Kohlköpfe sind gleichmäßiger um die Identitätslinie verteilt, wobei Ausreißer eher darunter liegen. Der orangene Balken zeigt das 15% *MAPE* Intervall, wobei Werte außerhalb dieses Intervalls, in Absprache mit den Zuchtexperten, als Ausreißer angesehen werden können. Es ist jedoch festzustellen, dass anteilmäßig für alle Typen ungefähr gleich viele Ausreißer vorhanden sind. Die extremsten Ausreißer finden sich beim Typ *Rund*. Für die 3 Typen wurden Korrelationskoeffizienten zwischen 0,67 und 0,74 erzielt, wobei der Typ *Spitz* am besten und der Typ *Flach* am schlechtesten abschnitt. Den schlechtesten *MAPE* mit 26,9% erzielten die Köpfe vom Typ *Flach*. Für spitze Köpfe betrug der *MAPE* 21,6% und für flache Köpfe 12,6%.

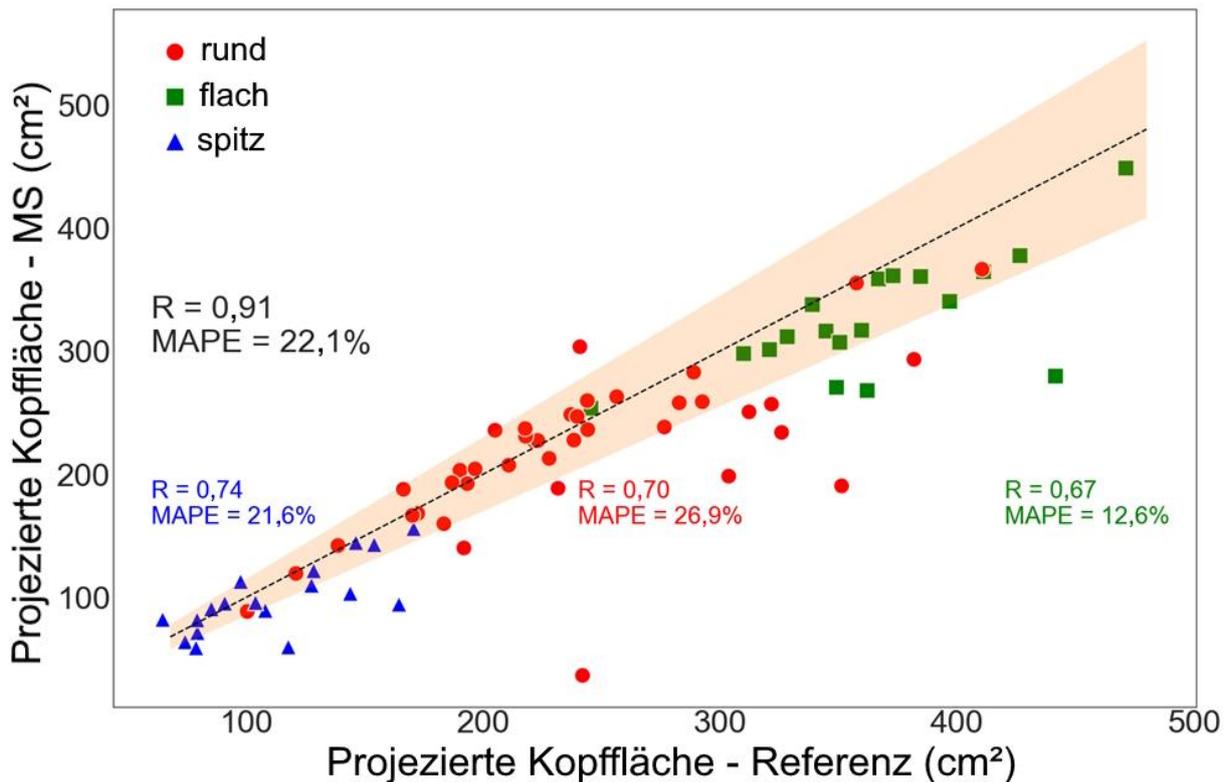


Abbildung 33: Vergleich der projizierten Kopffläche zwischen manuell annotierten und von einem trainierten Mask R-CNN vorhergesagten Kopfmasken; Die Daten sind nach den einzelnen Kopfformtypen rund, spitz und flach aufgeteilt und es wird jeweils der Korrelationskoeffizient R sowie der mittlere Absolute Fehler in Prozent (MAPE) angegeben; Die gestrichelte Linie zeigt die Identitätslinie; Der orangene Balken indiziert den 15% MAPE Bereich, welcher aus Züchtersicht einen akzeptablen Bereich darstellt; n=77

3.5.4 Diskussion

Laut Kenntnisstand des Autors ist dies die erste Studie, welche den direkten Zusammenhang zwischen der projizierten Kohlkopffläche (berechnet in einem 2-dimensionalen RGB-Bild) und dem Gewicht des Kohlkopfes untersucht. Dabei müssen die Ergebnisse in Abhängigkeit zum Kohlkopftyp (*Rund*, *Spitz*, *Flach*) interpretiert werden: So zeigen vor allem Kohlköpfe vom Typ *Spitz* einen anderen Zusammenhang zwischen den untersuchten Parametern im Vergleich zu den Typen *Rund* und *Flach*, wie sich aus der Lage der Regression-Modelle rückschließen lässt (Abbildung 32a). Dies könnte darauf zurückzuführen sein, dass die projizierte Fläche im Vergleich zur Masse klein ist, da Spitzkohl länglich im vertikalen Profil ist. Für Kohlköpfe vom Typ *Rund* wurde weiterhin ein vergleichsweise großer MAPE bestimmt, was 2 Gründe haben kann: Einerseits war es schwierig akkurate Annotationen der Kohlköpfe durchzuführen, da diese oftmals vom Umblatt der Pflanze verdeckt wurden. Zweitens unterschied sich bei den 4 untersuchten runden Sorten eine Sorte (R4) deutlich in ihrer Dichtigkeit von den anderen. Unter Bezugnahme der Dichte konnte der MAPE für runde Köpfe deutlich gesenkt werden. (Abbildung 32b).

Die Detektion der Kohlköpfe mittels eines trainierten *Mask R-CNN* lieferte mit einer Detektionsrate von 96,3% bei 8 untersuchten Sorten sehr gute Ergebnisse zu vergleichbaren Studien (Abbildung 34a). So untersuchten vorausgegangene Studien meist 1-4 unterschiedliche Sorten unter Feldbedingungen und erzielten mit Techniken des maschinellen Sehens Detektionsraten zwischen 46,92% und 94,43% (Hayashi & Sakau, 1996b, 1996a; Murakami et al., 1997; J. Zhang et al., 2022). Für die einzelnen Wuchs-Typen deuten die MAPE-Statistiken zwischen den vorhergesagten und manuell annotierten projizierten Kopfflächen auf Ungenauigkeiten bei den Vorhersagen des neuronalen Netzes. Dazu ist anzumerken, dass der MAPE (genau wie das arithmetische Mittel) empfindlich gegenüber

einzelnen Ausreißern ist und vor allem für die Typen *Spitz* und *Rund* lassen sich Ausreißer feststellen, bei denen die Kopffläche deutlich unterschätzt wurde (Abbildung 34b). Lässt man die Ausreißer außer Acht, kann man weiterhin feststellen, dass die Werte für die runden Köpfe gleichmäßig um die Identitätslinie verteilt sind, die Flächen für die spitzen und flachen Köpfe jedoch tendenziell unterschätzt werden, was vor allem mit einer Unterrepräsentation dieser Typen in den Trainingsdaten zusammenhängen könnte. Weiterhin können Umblätter, welche den Kopf überlappen, oder locker sitzende Deckblätter des Kopfes immer wieder zu Unterschätzungen durch das Mask R-CNN führen: So orientierten sich die manuellen Annotationen des Kopfes im Bild meist an dem dunklen Übergangsbereich zwischen Kopf und Pflanze, welcher sich durch dunkle Pixelwerte und Kanten im Bild definiert. Dieser Bereich wurde bei der manuellen Annotation jedoch dann ignoriert, wenn beispielsweise eine Überlappung des Kopfes mit einem Umblatt zu erkennen war. Zwar wurde dieses Verfahren auch für die Trainingsdaten beachtet, jedoch könnte ein umfangreicherer Trainingsatz die Abschätzung der Kontur eines verdeckten Kohlkopfes weiter verbessern. Für einen Großteil der Pflanzen konnte das Netzwerk jedoch verlässlich Vorhersagen machen, da ein Großteil der Werte im, für den in der züchterischen Praxis akzeptablen 15% MAPE-Intervall liegt und sich auch im Korrelationskoeffizient von 0,91 für alle Messungen widerspiegelt.



Abbildung 34: Vorhersagen des trainierten Mask R-CNNs für den Kohlkopf (weiß) und die gesamte Pflanze (blau); a) zeigt eine starke und b) eine schwache Vorhersage des Netzwerks.

4 Fazit und Ausblick

4.1 Präzise Hochdurchsatzanalyse wichtiger Formparameter länglicher Gemüsesorten

Fazit: In dieser Arbeit wurden verschiedene für längliche Gemüsearten typische züchtungsrelevante Formparameter wie die Länge, das Kaliber, die Krümmung und die Eigenschaften der Fruchttenden am Beispiel der Buschbohne untersucht. Mit dem vorgestellten Ansatz, der auf der Analyse von 2D RGB-Bildern mit Methoden des maschinellen Sehens basiert, konnten diese Formparameter mit einer hohen Genauigkeit je nach Parameter quantifiziert oder klassifiziert werden. Die Detektion der Hülsenenden erfolgt mit einer ausreichend hohen Korrektklassifikationsrate, sodass weitere züchtungsrelevante Formparameter im Folgenden genau erfasst werden können. Länge und Kaliber sind als physikalisch korrekte Größen erfassbar. Damit ersetzt das Verfahren die aufwendige Messung mit Lineal und Lochlehre. Die Krümmung wird punktuell und nicht über den Gesamteindruck abgeleitet und ist als kontinuierlicher Wert darstellbar. Gerade hier ist eine deutliche Objektivierung zu erwarten, da als alternatives Verfahren nur die Selbstkalibrierung an wenigen vorgezeichneten Krümmungsklassen in Frage kommt, welche limitiert und fehleranfällig ist.

Mit diesem anwendungsbezogenen Analysesystem stehen nun konkrete Lösungen zur Bewertung züchtungsrelevanter Fruchtqualitätsmerkmale zur Verfügung, welche in verschiedenen Kontexten wie der Gemüsezüchtung, aber auch für diverse Forschungsfragen genutzt werden kann. Die beiden wesentlichen Aspekte, also eine im Vergleich zu gängigen Verfahren höhere Genauigkeit der analysierten Parameter sowie ein vergleichbarer oder auch höherer Durchsatz konnten erreicht bzw. gezeigt werden. Die wesentliche Limitierung besteht in der 2-dimensionalen Repräsentation der Früchte. Dieser Einschränkung kann jedoch durch gute Praxis bei der Positionierung der Früchte oder durch nachfolgende statistische Beurteilungen der Verteilungen innerhalb einer Sorte entgegengewirkt werden. Ein Beispiel für eine Ausreißer-Analyse ist die Identifikation von Hülsen, bei denen die Spitze durch das neuronale Netzwerk nicht richtig detektiert wurde. Da die Länge für die untersuchten Hülsen bei jeder Sorte sehr uniform war, lassen sich Ausreißer leicht durch Toleranzbereiche wie z.B. in Verteilungsstatistiken identifizieren. Bei der Datengenerierung ist immer eine ausreichend große Wiederholungsrate von mindestens 10 Früchten pro Sorte zu empfehlen. Weiterhin wurde in dieser Studie eine große Spanne an unterschiedlichen Phänotypen zur Validierung der vorgestellten Methoden genutzt, womit eine praxisnahe Anwendbarkeit des Systems bewiesen wurde.

Ausblick: Das hier vorgestellte Verfahren zeichnet sich durch eine einfache Bedienung von der Bildgenerierung bis hin zur Analyse aus. So ist die bildgebende Box vollumfänglich kalibriert und benötigt in der Anwendung keine weitere Einstellung durch fachkundiges Personal. Das Verfahren kann somit direkt in die Anwendung, beispielsweise in der Bohnenzucht, gehen und speichert dabei die generierten Bildmasken, Skelette, und Analysedetails probenweise, sodass eine nachträgliche statistische Analyse (z.B. zur Identifikation von Ausreißern) durchgeführt werden kann. Alle Analysen können mit dem vorgestellten System objektiv und wiederholgenau durchgeführt werden. Die ordinalskalierten, subjektiven Bewertungen der Krümmung auf der Grundlage einer Züchterskalen ist durch eine kontinuierliche Skala ersetzt worden, welche prinzipiell eine feinere Bewertung zulässt. Das entwickelte System verbindet die im züchterischen Kontext notwendige Präzision mit einer Erhöhung des Durchsatzes.

Außerdem sollten die entwickelten Methoden problemlos auf längliche Früchte anderer Kulturarten anwendbar sein, welche ebenfalls keine Verzweigungen der Frucht aufweisen. Es

versteht sich von selbst, dass das genutzte *Mask R-CNN* für die Detektion von Blütenansätzen an anderen Kulturarten neu trainiert werden muss, jedoch konnte in dieser Studie gezeigt werden, dass die Methode grundsätzlich sehr gut funktioniert und dass mit Hilfe des Transferlernen-Ansatzes rund 380 Bilder von Einzelfrüchten ausreichen, um den Algorithmus auf die spezifische Kulturart zu trainieren. Zwar ist nicht zu erwarten, dass die hier vorgestellten Methoden für andere längliche Gemüsesorten weniger gute Ergebnisse liefern, nichtsdestotrotz sollten die Methoden bei anderen Kulturarten ebenfalls validiert werden.

Für die dargestellten Ansätze des tiefen Lernens ist zu erwarten, dass das Training mit mehr und diverseren Daten, bei gleichzeitig ausgeglichenen Datensätzen (alle Merkmalsausprägungen kommen gleich oft im Trainingsdatensatz vor) zu besseren Ergebnissen führen kann. Dies gilt beispielsweise für die Detektion der Spitzen und Blütenansätze. So zeigte das *Mask R-CNN* Schwächen bei der Erkennung sehr kleiner Blütenansätze. Dies könnte auf eine unzureichende Generalisierung der Netzwerke durch die Zusammensetzung des Trainingsdatensatzes oder auf eine nicht ausreichende Optimierung der Netzwerkparametrisierungen zurückzuführen sein (dies gilt auch für die Netzwerke, welche in den anderen Bereichen dieser Arbeit genutzt wurden). Das hier entwickelte Konzept zur Erfassung der Bilder mit einer manuell zu bestückenden Aufnahmebox war für diese Arbeit vollkommen ausreichend und könnte im professionellen Einsatz, z.B. in Züchtungsunternehmen, eine deutliche Durchsatzsteigerung ermöglichen. Dennoch könnte man im Sinne einer Vollautomation das System zukünftig als Fließband weiterentwickeln, sodass menschliche Interaktion lediglich für die Bereitstellung der Hülsen in ausreichenden Mengen notwendig wäre. Weitere Anlagenmerkmale wie beispielsweise die Sortenzuordnung mittels QR-Code oder benutzerfreundliche Schnittstellen wären hierfür erforderlich.

4.2 Nicht-invasive Analyse der Samenanlage von Leguminosen

Fazit: Nach Kenntnisstand des Autors ist dies die erste Studie, welche die Evaluierung der Samenanlage von Leguminosen mittels transmittierendem Lichts untersucht. Die hier durchgeführten Versuche zeigen, dass der Ansatz grundsätzlich vielversprechend ist und eine deutliche Verbesserung im Vergleich zur gängigen züchterischen Praxis zur Analyse der Samenanlage darstellen kann, da hierdurch die Samen sichtbar gemacht und das Fehlen dieser erfasst werden kann. Eine präzise Lokalisierung der Samen, Fehlstellen und Schwanenhäule innerhalb der Bohne ist möglich, sofern die Samen eine gewisse Größe erreicht haben und ausreichend Kontrast zwischen Samen und dem Rest der Hülse besteht. Das hier genutzte *Faster R-CNN* zeigte eine gute Performance, wenn diese Bedingungen erfüllt sind. Für eine Sorten mit gelben, lichtdurchlässigen Hülsen werden Fehlstellen, Samen und Schwanenhäule zuverlässig erkannt, sodass das Verfahren in der Praxis genutzt werden kann. Bei grünen Schoten ist dies ebenfalls für hellere Sorten und/oder reiferen Entwicklungsstadien der Fall, jedoch kann das Feintuning der Aufnahmeparameter (Lichtstärke und Belichtungszeit) bei variierenden Helligkeiten und Reifestadien der Hülsen einen zeitlichen Flaschenhals darstellen. Das vorgestellte Verfahren stellt eine realistische Alternative für die nicht-invasive Analyse der Samenanlage von Leguminosen dar und empfiehlt sich für weitere Untersuchungen.

Ausblick: Die Erkennungsraten für weniger transmittierende Sorten und/oder solche mit kleineren Samen könnte durch Verbesserungen im Aufnahmeprozess erhöht werden. Um die entstehenden Datenmengen gering zu halten, wurden die Aufnahmen in dieser Studie mit einem Dynamikumfang von 8-Bit gemacht. Dies ist im züchterischen Kontext von Bedeutung, da viele unterschiedliche Sorten in großem Umfang untersucht werden und die Datenspeicherung viele Ressourcen benötigt. Weiterhin funktionieren viele Verfahren lediglich mit 8-Bit Bildern, sodass Bilder mit größerem Dynamikumfang runtergerechnet werden

müssten, die dann jedoch eine andere Verteilung in den Intensitätswerten hätten. Für die Performance des Durchlichtverfahrens kann eine Erhöhung des Dynamikbereichs nach den Ergebnissen dieser Studie jedoch eine Verbesserung bringen. Ein Verfahren (Kamera + Software-Weiterentwicklung), welches einen höheren Dynamikbereich nutzen kann, ist ein Verbesserungsvorschlag. Weiterhin könnte man mit Belichtungsreihen, also durch die Aufnahme einer Bilderserie mit unterschiedlichen Belichtungseinstellungen, in Kombination mit einer Kontrastanalyse (auf Basis der Intensitätshistogramme) der Hülsen die optimale Einstellung für jede Hülse finden. Hierdurch würden die individuellen und damit sehr zeitaufwendigen Einstellungen der Aufnahmeparameter der Kamera und des Hintergrundlichts entfallen. Weiterhin können die Durchlichtbilder zukünftig für die Analyse weitere Parameter wie die Kornmarkierung der Hülsen oder für Ertragsprognosen genutzt werden.

4.3 Chromatische Analyse von RGB-Bilddaten zur Evaluierung von Sorteneigenschaften sowie Stresssymptomen in der Gemüsezucht

Fazit: Der erste Ansatz, welcher in dieser Arbeit verfolgt wurde, um eine möglichst genaue Farbwiedergabe in Bilder zu erreichen, war die Nutzung einer Fotobox. Durch die kontrollierbaren Bedingungen konnte die Hardware präzise aufeinander eingestellt werden, um eine genau Farbwiedergabe zu gewährleisten. Es wurde eine neue Methode entwickelt, mit der nicht nur zwischen grünen und gelben Bohnenhülsen durch Schwellenwertanalyse des H-Kanals (HSV-Farbraum) unterschieden werden kann, sondern auch seltene zweifarbige Hülsen mittels RF-Klassifizierung identifiziert werden können, welche in Zuchtprogrammen unerwünscht sind. Obwohl für die zweifarbige Klasse nur begrenzte Daten zur Verfügung standen (da diese selten ist), konnten wir mit Hilfe des trainierten RF-Modells $83\pm 7\%$ der Hülsen richtig klassifizieren. Das Verfahren kann zukünftig der Züchtung helfen, jene gelben Bohnensorten zu identifizieren, welche zur Bildung zweifarbiger Hülsen neigen.

Der Anwendungsfall *Alternaria* infizierter Möhrenblätter wurde genutzt, um die technischen Limitierungen eines kontrollierbaren RGB-Verfahrens zu untersuchen. Diese Beurteilung muss allerdings auch unter Berücksichtigung der Limitierungen erfolgen, welche die züchterische, d.h. visuelle Bewertung des Infektionsstadiums beeinflussen und welche einen erheblichen Einfluss auf die Vorhersagekraft der trainierten Modelle haben. So liegt den visuellen Bewertungen des Zuchtextperten eine nicht-lineare Skala zu Grunde, in die verschiedene Kriterien einfließen, deren Gewichtung ebenfalls nicht bekannt ist. Die Kriterien konnten vom Züchter zwar benannt, aber nicht genauer definiert werden. Zunächst muss man sich fragen, ob die Auflösung des Bewertungssystems mit den Klassen 1 bis 9 nicht zu fein graduiert ist und dementsprechend die Bewertungsmöglichkeiten des Zuchtextperten nicht überfordert. In anderen Worten: Ist der Experte überhaupt in der Lage zwei in der Züchterskala benachbarte Klassen in jedem Bereich (eindeutig) zu unterscheiden. Diese Punkt muss zumindest berücksichtigt werden, wenn man die Streuung der MS-Analysen innerhalb der Klassen 2-6 beurteilt. Weiterhin stellte die Klasse 9 einen Spezialfall dar, da diese keinen Wertebereich abdeckte, sondern das untere Ende der Klasse 8 darstellt. Diese Limitierungen in den Referenzbewertungen erschwerten schließlich auch die Entwicklung und Bewertung der Methoden des maschinellen Sehens. So können fehlerhafte Referenzwerte, welche aus der Subjektivität des Bewertungssystems resultieren, dafür sorgen, dass die Modelle des maschinellen und tiefen Lernens nicht angemessen generalisieren können. Weiterhin können einfache MS-Methoden, wie der Segmentierungsansatz, die vielschichtigen Einflussfaktoren der Züchterbewertungen nicht abbilden. Eine weitere Limitierung liegt in dem genutzten 2D Bildgebungsverfahren, da Blattverdeckungen jene Blattsegmente verdecken können, welche für die Bewertung des Züchters unter Umständen ausschlaggebend sind. Es konnten jedoch auch Vorteile für die Nutzung des RGB-Verfahrens gezeigt werden, die auch im Ausblick noch weiter erläutert werden. So ließ sich der Trend der Züchterwerte mit allen getesteten Verfahren

gut abbilden (hohe Korrelationen), und es konnte mit dem System ein hoher Durchsatz bei geringen Systemkosten und einer einfachen Handhabung gewährleistet werden. In Bezug auf die Bildanalyseverfahren erzielte das RF-Verfahren die besten Ergebnisse und war trotzdem mit vergleichsweise geringem Ressourcenaufwand zu etablieren, was die Flexibilität und Anwendungsvielfalt dieses Verfahrens unterstreicht. Im Gegensatz dazu benötigte das Verfahren des tiefen Lernens einen sehr hohen Ressourcenaufwand (Material, Rechenleistung, Energie), wurde jedoch durch die geringe Auflösung der Eingangsbilder limitiert.

Der zweite Ansatz, welcher in dieser Arbeit verfolgt wurde, um eine möglichst genaue Farbwiedergabe in Bildern zu erreichen, war die nachträgliche Farbkalibrierung der Bilder. Dieser Ansatz ist vor allem für Feldversuche interessant, da dort nicht ohne Weiteres kontrollierbare Bedingungen geschaffen werden können. Für eine genauere Einschätzung der Farbeigenschaften von Gemüse wurde in dieser Arbeit eine validierte Methode zur Farbkalibrierung erstellt, mit welcher sehr präzise Resultate unter Feldbedingungen erzielt werden können. Auch wenn die Unterschiede zwischen Farb-transformierten und nicht-transformierten Realbildern im Kohl unter den gegebenen Bedingungen gering waren, besitzt das Verfahren ein großes Potenzial die Analyse der Farbanteile zu verbessern, gerade dann, wenn die Lichtbedingungen hinsichtlich der chromatischen Zusammensetzung stärker fluktuieren. Allerdings können für grobe Farbanalysen, wie es Farbklassifizierungen in die 6 Grundfarben sind, bereits gute Aufnahmebedingungen in Kombination mit angemessenen Kameraeinstellungen ausreichend sein. Dies inkludiert beispielsweise nicht bei direkter Sonneneinstrahlung zu fotografieren, um Schattierungen zu vermeiden, sowie einen guten Weißabgleich und eine angemessene Belichtung im Bild zu schaffen. Dabei ist zu beachten, dass Bilder, welche unter denselben Aufnahmebedingungen aufgenommen wurden, grundsätzlich erstmal den gleichen Fehler beinhalten und dadurch vergleichbar sind. Sollte eine Farbkalibrierung, beispielsweise für Zeitreihenanalysen nötig sein, empfehlen sich die hier vorgestellten Kalibrierungsmethoden, gerade in Bezug auf die zeitliche Veränderung der Lichtfarbe. Während die globalen Methoden weniger Referenzfarben benötigen und dadurch einfacher zu etablieren sind, ist die lokale TPS-Interpolation in Kombination mit ausreichend Referenzfarben in der Lage äußerst genaue Farbtransformationen zu erzielen. Auf Grund der Größe des RHS-Farbpanels ist eine automatisierte Feldanwendungen schwerer zu etablieren, könnte aber beispielsweise durch Nutzung einer Zweitkamera realisiert werden. Diese kann parallel zu den Aufnahmen der Kohlpflanzen Bilder des RHS-Farbpanels erzeugen, sodass zu jedem Aufnahmezeitpunkt eine korrekte Farbtransformation möglich wird.

Ausblick: Wie aus dieser Studie hervor geht, ist es für Evaluierungen von Sorteneigenschaften und Stresssymptomen an Pflanzen, welche auf der Farbe beruhen, wichtig, zuverlässige Referenzdaten zu erheben (dies gilt allgemein für visuell erhobene Referenzdaten). Um zukünftig qualitativ hochwertigere Referenzdaten zu erhalten besteht die Möglichkeit, ein angepasstes Protokoll für die Erhebung visueller Referenzdaten zu entwickeln. Da die Züchterbewertungen ein nicht klar definierbares Protokoll verfolgen, könnte man die Bewertungen direkt am Pflanzgut durch Bewertungen an den aufgenommenen Bildern ersetzen. So wären zunächst für beide Verfahren die gleichen Voraussetzungen gegeben. Weiterhin sollten die visuellen Bewertungen an den Bildern randomisiert und zeitlich und personell mehrfach unabhängig voneinander erfolgen, um Varianzen in den Züchterbewertungen quantifizierbar zu machen. Eine weitere Lösung könnte sein, dass man die Staffelung der ordinalen Skalen in sehr schwere, schwere, moderate und leichte Effekte zusammenfasst, wie es auch in anderen Studien gemacht wurde (Prabhakar et al., 2020; G. Wang et al., 2017; Zeng et al., 2020).

Die Leistung des neuronalen Klassifikationsnetzwerk für die Bestimmung des Infektionsstadiums von Möhrenblättern kann zukünftig verbessert werden, indem man die

Bilder in kleinere Segmente aufteilt, einzeln klassifiziert und daraus eine Gesamtbewertung für ein Blatt in voller Auflösung generiert, wobei dieser Prozess jedoch ein Vielfaches an Aufwand hinsichtlich der Erstellung der Referenzdaten für die Bildsegmente bedeutet, sowie deutlich mehr Rechenaufwand benötigt. Eine zusätzliche Verbesserung verspricht weiterhin die Augmentierung der Trainingsdaten, z.B. mittels *Deep Convolutional Generative Adversarial Network* (GAN) (Zeng et al., 2020). Der deutlich geringere Ressourcenaufwand und flexible Anwendbarkeit empfiehlt in jedem Fall zukünftig auch eine Weiterverfolgung des RF-Ansatzes für Klassifizierungsaufgaben von Objekten in Bildern anhand bestimmter Farbmerkmale. Aber auch bezüglich des Versuchsaufbaus könnte eine alternative Herangehensweise interessant sein, z.B. für die Bestimmung der Infektionsstadien von Möhrenblättern mit *A. dauci* direkt im Anbau. Im Feld werden die Pflanzen sortenweise in einzelnen Versuchsgliedern, meistens bestehend aus 5 Anbaureihen je 3 Meter Länge angebaut, sodass auch eine Evaluierung des Infektionsstadiums der einzelnen Versuchsglieder denkbar ist. In diesem Szenario, indem viele Pflanzen gleichzeitig in die Bewertung einfließen, könnten die infizierte Blattfläche oder der Blattflächenindex geeignete Kenngrößen zur Bestimmung darstellen. Hierfür könnte man ebenfalls ein Fieldbike-System nutzen, an dem man RGB-Kameras montieren kann.

Im Bezug auf die Farbkalibrierung von Bildern muss man immer Aufwand und Nutzen gegeneinander abwägen und eine für die Aufgabenstellung angepasste Lösung wählen. So kann die vorgestellte TPS-Methode in Kombination mit dem RHS-Farbpanel zukünftig für Referenzierungs- und Kalibrierungszwecke genutzt werden. Sollen beispielsweise zeitreihen Farbanalysen unter Feldbedingungen erfolgen, kann ein globaler Kalibrierungsansatz die richtige Wahl sein.

4.4 Automatisierte und nicht-invasive Hochdurchsatz-Ertragsabschätzung bei kopfbildendem Blattgemüse

Fazit: Für die Ertragsabschätzung in Kopfkohl mittels 2D RGB-Verfahren sind verschiedene Verfahren, wie die Detektion der Köpfe und die Abschätzung der Biomasse aus mehreren Formparametern entwickelt worden, welche Potential besitzen die traditionellen Methoden zu unterstützen bzw. zu ersetzen. Die Detektion der Köpfe und anschließende Berechnung der projizierten Kopffläche stellt ein wichtiges Maß dar, reicht jedoch nicht in allen Fällen aus, um eine genaue Abschätzung der Kopfmasse zu gewährleisten, da sich Sorten durch unterschiedliche Dichten auszeichnen und diese je nach Entwicklungsstadium des Kopfes noch variieren können. Daher konnte unter Zuhilfenahme einer Dichteabschätzung die Genauigkeit der Ertragsabschätzungen teilweise deutlich verbessert werden. Dies empfiehlt sich vor allem auch für die Auswertung von Zeitreihenaufnahmen. Um die Limitierungen des 2D RGB-Verfahrens aber noch besser zu kompensieren, müssen weitere Informationen in ein Vorhersagemodell einfließen, worauf im Ausblick näher eingegangen wird. Die Automatisierung der Kopfsegmentierung mittels *Mask R-CNN* stellt einen vielversprechenden Ansatz dar. Hier lagen 75,4% der Vorhersagen in einem Toleranzbereich (<15% MAPE), der z.B. aus Züchtersicht als praxistauglich bewertet wurde. Die Ergebnisse bei flachen Köpfen (systematische Unterschätzung der projizierten Kopffläche) deuten jedoch darauf hin, dass individuelle Modelle zumindest für verschiedene Kopftypen sinnvoll sein können.

Ausblick: Im Folgenden sollen verschiedene Ideen zur Weiterentwicklung aus den Resultaten dieser Arbeit abgeleitet werden, durch die eine Ertragsabschätzung bei kopfbildendem Blattgemüse im größeren Stil automatisiert und im betrieblichen Kontext durchführbar sein könnte: Diese setzen bereits bei dem Aufnahmeverfahren an. Hier könnte zukünftig die Bilderfassung durch die Nutzung einer nach unten offenen Fotobox erfolgen, welche an einen Traktor gehangen wird. Die Box besitzt dabei ähnliche Dimensionen wie das hier genutzte Fieldbike, sodass 3 Kameras für einen hohen Durchsatz parallel geführt werden. Dabei sollte

die Box knapp über dem Pflanzbestand geführt werden, sodass die Aufnahmebereiche stets, und unabhängig der Witterungsbedingungen, schattiert sind. Weiterhin entfällt, nach einer Kalibrierung des Systems, eine nachträgliche Farbtransformation der Bilder, da, durch in der Decke der Box montierte Halogen-Beleuchtung, sehr einheitliche Aufnahmebedingungen herrschen. Außerdem ermöglicht eine solche Beleuchtung einen umfangreichen Einsatz von beispielsweise Multispektralkameras, welche durch das enge Lichtspektrum von LED-Dioden limitiert würden. Um die entstehenden Datenmengen so gering wie möglich zu halten, lösen die Kameras nur dann aus, wenn sich eine Pflanze in einem bestimmten Ausschnitt im Zentrum des Bildes befindet, wobei man für die Erkennung in Echtzeit (während der Überfahrt) ein Faster R-CNN oder ein YOLO-Netzwerk nutzen kann. Die Bilder werden auf einem Bordcomputer des Traktors zwischengespeichert und bei Bedarf in eine Cloud hochgeladen und automatisiert verarbeitet und analysiert, um den Züchtern die Daten schnellstmöglich verfügbar zu machen. In den frühen Wachstumsstadien, wenn noch keine Köpfe entwickelt sind, kann das Pflanzenwachstum über die Berechnung der projizierten Pflanzenfläche (z.B. durch Nutzung des *Excess Green* Vegetationsindex in Kombination mit klassischen Segmentierungsmethoden) erfasst werden (Kim et al., 2018; Low & Tay, 2019). Sobald die Köpfe als solche erkennbar sind, kann deren projizierte Fläche mit Hilfe eines *Mask R-CNN* vorhergesagt werden. Zur Verbesserung dieser Netzwerkvorhersagen kann man entweder manuell mehr Trainingsdaten erzeugen, oder eine GAN-basierte Augmentierung der Trainingsdaten in Betracht ziehen (Zeng et al., 2020). Die projizierten Kopfflächen fließen dann mit weiteren Parametern in ein neuronales Vorhersagemodell (ANN), um den Ertrag jeder einzelnen Pflanze präzise abzuschätzen. Dabei ist denkbar, die Ertragsmodellierung durch Nutzung nicht-invasiver Daten zu optimieren. So könnte man anstatt der Dichte des Kopfes, welche an jedem Aufnahmetag an einigen wenigen Köpfen manuell gemessen werden muss, beispielsweise Sortencharakteristika wie Kopfform (rund, flach, spitz), oder die Art der Pflanze (z.B. Weißkohl, Wirsing, Chinakohl) einfließen lassen. Auch Daten wie die Wachstumsgradtage oder die aufsummierte Niederschlagsmenge oder Sonnenstunden können einfließen (Ji et al., 2017). Weitere mögliche Parameter für die Modellierung ist beispielsweise die Pflanzenhöhe, welche durch 3d Modelle (z.B: aus *Structure-from-Motion* Analysen) ermittelt werden kann (Kim et al., 2018).

5 Résumé

In dieser Arbeit wurden in vier verschiedenen Themengebieten methodische Ansätze zur Charakterisierung unterschiedlicher Merkmale des Pflanzensprosses anhand von drei Kulturarten entwickelt. Alle Merkmale, die entweder physikalisch messbar sind oder verschiedenen Form- und Farbkategorien zugeordnet werden können, sind für die Gemüsezüchtung von großer Relevanz und konnten nicht-invasiv mittels intelligenter, bildgebender 2D RGB-Verfahren erfasst, quantifiziert oder kategorisiert werden. Es konnte gezeigt werden, dass die Kombination aus Verfahren der klassischen Bildverarbeitung (z.B. Pixelwertanalysen), des klassischen maschinellen Lernens (z.B. Random Forest Modelle) und des tiefen Lernens (neuronale Netzwerke) ein Katalysator für die günstige, objektive, durchsatzgesteigerte und dennoch präzise Evaluierung wichtiger Züchtungsmerkmale sein kann. So konnte die Genauigkeit und Praxisauglichkeit der Methoden für die meisten Anwendungsfälle trotz der Limitierung auf RGB und 2 Dimensionen gezeigt werden. Auf der anderen Seite zeigt diese Arbeit auf, dass man sich den technischen Limitierungen eines 2D RGB-Verfahrens bewusst sein muss. Am Beispiel der Quantifizierung des Befallsgrads und der deutlichen Korrelation der Messergebnisse mit den Einschätzungen des Zuchtexperten wird deutlich, dass es prinzipiell möglich ist mit der RGB Methode diese Art von Fragestellungen anzugehen, für deren Beantwortung üblicherweise multispektrale Methoden verwendet werden. Es ist aber auch klar, dass die Anwendbarkeit auf (für den Menschen) sichtbare Symptome beschränkt bleibt und dass man gerade im Bereich der Früherkennung von Stresssymptomen in vielen Fällen eher auf die multispektralen Verfahren zurückgreifen muss (Stoll et al., 2008; D. Wang et al., 2019). Weiterhin muss beachtet werden, dass man bei einem 2D Verfahren nur eine Perspektive nutzen kann. Hier kann entweder die Position und Orientierung des Untersuchungsobjekts optimiert werden, wobei es auch sinnvoll sein kann die Größe des möglichen Fehlers vorher zu bestimmen, sollte nur eine Aufnahme aus Gründen des Durchsatzes erwünscht sein. Oder man muss mehrere Aufnahmen machen, um aus der Vielzahl an gewonnenen Daten bestimmte Merkmale unter Zuhilfenahme statistischer Kenngrößen zu quantifizieren.

Auch wenn es keine Patentlösung für die Evaluierung von züchtungsrelevanten Pflanzenmerkmalen mittels intelligenter, bildgebender 2D Verfahren gibt und auf Grund der vielseitigen Fragestellungen immer wieder individuell angepasste Lösungsansätze gefunden werden müssen, so lassen sich aus dieser Arbeit dennoch grundsätzliche Regeln ableiten, die bei der Erarbeitung neuer Lösungsansätze helfen können: Zunächst muss man die Problemstellung präzise formulieren. Dies beinhaltet das Pflanzenmerkmal und das Maß, welches für die Charakterisierung verwendet werden kann. Daraus muss abgeleitet werden, ob ein auf 2D RGB-Sensorik basierendes Verfahren mit seinen Limitierungen ein geeigneter Lösungsansatz sein kann. Stellt sich eine Herangehensweise mittels 2D RGB-Verfahren als sinnvoll heraus, gilt es vor allem zu klären, welche Nebenbedingungen wichtig sind. Dies kann beispielsweise eine Durchsatzsteigerung oder eine Objektivierung des bisherigen Verfahrens sein und damit verbunden eine Erhöhung der Wiederholgenauigkeit oder eine Steigerung der Genauigkeit. Auf Grundlage dessen kann man entscheiden, auf welchen Abstraktionslevel (Pflanzenorgan, Gesamtpflanze, Nahaufnahmen, Fernaufnahmen) man mit dem Verfahren ansetzt und ob man nicht-invasiv (z.B. Pflanzen im Feldbestand), teil-invasiv (z.B. Ernten von Früchten) oder invasiv (destruktiv, um z.B. in das Innere einer Frucht zu schauen) vorgeht. Die Ergebnisse dieser Arbeit legen weiterhin nahe, dass Referenzmessungen zur Validierung neuer entwickelter Verfahren nach Möglichkeit nicht auf erfahrungsbasierten Schätzwerten auf einer ordinalen Skala beruhen sollten, auch wenn dies oft der Realität in der züchterischen Praxis entspricht. In solchen Fällen sollten die Referenzwerte personell und zeitlich unabhängig voneinander gemacht werden, oder man etabliert aufwendigere Messverfahren. So könnte die Farbe beispielsweise mit Hilfe eines Multispektralsensors gemessen werden. Bei der Aufnahme der Bilder sollte zunächst eine räumliche Kalibration erfolgen, wenn durch

das genutzte Objektiv nennenswerte Verzerrungen entstehen. Der spezifische Anwendungsfall gibt vor, ob und welche farbliche Kalibration notwendig ist. Verfahren für eine angemessene farbliche Kalibration wurden ebenfalls in dieser Arbeit vorgestellt.

Teilt man das Spektrum der in dieser Arbeit verwendeten Methoden auf, so werden vermutlich zunächst die neuronalen Netze genannt. Dennoch empfiehlt es sich an die Problemstellung zunächst mit klassischen Bildverarbeitungsverfahren heranzugehen, da diese schnell und einfach zu implementieren sind und einen geringen Ressourcenverbrauch haben. So benötigen diese beispielsweise keine Trainingsdaten. Beachtet man diese Herangehensweise, dann kann ein bildgebendes RGB-Verfahren eine kostengünstige und praxisnahe Lösung für objektive Evaluierungen von Gemüseeigenschaften darstellen, einen Transfer von invasiven hin zu nicht-invasiven Verfahren ermöglichen, gleichzeitig aber auch eine Durchsatzsteigerung und eine Steigerung der Genauigkeit bewirken.

6 Danksagung

Zunächst möchte ich der AiF danken, welche dieses Projekt durch Bereitstellung der finanziellen Mittel erst möglich gemacht hat. Des Weiteren der *GFPi* sowie den beteiligten Züchtungsunternehmen *Van Waveren Saaten GmbH*, *Rijk Zwaan Zaadteelt en Zaadhandel B.V.*, sowie *Nunhems B.V.*, deren Zusammenarbeit maßgeblich zum Erfolg dieser Arbeit beigetragen haben. Weiterhin möchte ich meinen Kollegen des Forschungszentrum Jülich und insbesondere Mark und Laura danken, welche mir während der vergangenen Jahre immer mit ihrem fachlichen Rat zur Seite standen. Zuletzt gilt mein Dank meinen Freunden, meiner Familie, sowie Verena, die stets durch sämtliche Höhen und Tiefen mit mir geht und meine größte Stütze ist.

7 Abkürzungsverzeichnis

AiF Arbeitsgemeinschaft industrieller Forschungsvereinigungen „Otto von Guericke“

ANOVA Analysis of Variance (Varianzanalyse)

APS-C Advanced Photo System type-C

C Celsius

CIE Commission Internationale de l'Eclairage

CM Centimeter

CMOS Complementary metal–oxide–semiconductor

CNN Convolutional Neural Network (Neuronales Faltungsnetzwerk)

CoV Variationskoeffizient

FN Falsch Negativ

FP Falsch Positiv

GAN Generative Adversarial Network

GFPI Gemeinschaft zur Förderung von Pflanzeninnovation e.V.

GIMP GNU Image Manipulation Program (<https://www.gimp.org/>)

K Kelvin

MAE Mean Absolute Error (mittlerer absoluter Fehler)

MAPE Mean Absolute Percentage Error (mittlerer absoluter Fehler in Prozent)

mAP₅₀ Mean Average Precision at 50 Percent Overlap

MLR Multiple Lineare Regression

MM Millimeter

MP Megapixel

MS Maschinelles Sehen

NDVI Normalized Difference Vegetation Index

NM Nanometer

QR-Code Quick Response-Code

PLS(R) Partial Least Square (Regression) (Regression der partiellen kleinsten Quadrate)

ReLU Rectified Linear Unit (gleichgerichtete lineare Einheit)

RF Random Forest

RHS Royal Horticultural Society

RN Richtig Negativ

RP Richtig Positiv

SGD Stochastic Gradient Descent (stochastischer Gradientenabstieg)

TPS Thin Plate Spline

8 Literaturverzeichnis

- Abed, S. H., Al-Waisy, A. S., Mohammed, H. J., & Al-Fahdawi, S. (2021). A modern deep learning framework in robot vision for automated bean leaves diseases detection. *International Journal of Intelligent Robotics and Applications*, 5(2), 235–251. <https://doi.org/10.1007/s41315-021-00174-3>
- Ali, J., Khan, R., Ahmad, N., & Maqsood, I. (2012). Random forests and decision trees. *IJCSI/ International Journal of Computer Science Issues*, 9(5), 272–278.
- Arakeri, M. P., & Lakshmana. (2016). Computer Vision Based Fruit Grading System for Quality Evaluation of Tomato in Agriculture industry. *Procedia Computer Science*, 79, 426–433. <https://doi.org/10.1016/j.procs.2016.03.055>
- Araújo, S. A. De, Pessota, J. H., & Kim, H. Y. (2015). Beans quality inspection using correlation-based granulometry. *Engineering Applications of Artificial Intelligence*, 40(4), 84–94. <https://doi.org/10.1016/j.engappai.2015.01.004>
- Armi, L., & Fekri Ershad, S. (2019). Texture image analysis and texture classification methods-A Review. *International Online Journal of Image Processing and Pattern Recognition*, 2(1), 1–29. <https://doi.org/https://doi.org/10.48550/arXiv.1904.06554>
- Baek, J., Lee, E., Kim, N., Kim, S. L., Choi, I., Ji, H., Chung, Y. S., Choi, M. S., Moon, J. K., & Kim, K. H. (2020). High throughput phenotyping for various traits on soybean seeds using image analysis. *Sensors (Switzerland)*, 20(1), 1–9. <https://doi.org/10.3390/s20010248>
- Becker, H. (2019). *Pflanzenzüchtung* (3. Auflage). utb GmbH. <https://doi.org/10.36198/9783838549507>
- Blasco, J., Aleixos, N., & Moltó, E. (2003). Machine vision system for automatic quality grading of fruit. *Biosystems Engineering*, 85(4), 415–423. [https://doi.org/10.1016/S1537-5110\(03\)00088-6](https://doi.org/10.1016/S1537-5110(03)00088-6)
- Bock, C. H., Chiang, K. S., & Del Ponte, E. M. (2022). Plant disease severity estimated visually: a century of research, best practices, and opportunities for improving methods and practices to maximize accuracy. *Tropical Plant Pathology*, 47(1), 25–42. <https://doi.org/10.1007/s40858-021-00439-z>
- Bodner, G., Alsalem, M., Nakhforoosh, A., Arnold, T., & Leitner, D. (2017). RGB and spectral root imaging for plant phenotyping and physiological research: Experimental setup and imaging protocols. *Journal of Visualized Experiments*, 2017(126). <https://doi.org/10.3791/56251>
- Boor, C. de. (1993). B(asic)-Spline Basics. In *Fundamental Developments of Computer-Aided Geometric Modeling* (1st ed.). Les Piegl (Ed.). <https://ftp.cs.wisc.edu/Approx/bsplbasic.pdf>
- Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*. <https://www.drdoobs.com/open-source/the-opencv-library/184404319>
- Brown, C. E. (1998). Coefficient of Variation. In *Applied Multivariate Statistics in Geohydrology and Related Sciences* (pp. 155–157). Springer. https://doi.org/10.1007/978-3-642-80328-4_13
- Chaerle, L., Hagenbeek, D., De Bruyne, E., Valcke, R., & Van Der Straeten, D. (2004). Thermal and chlorophyll-fluorescence imaging distinguish plant-pathogen interactions at

- an early stage. *Plant and Cell Physiology*, 45(7), 887–896.
<https://doi.org/10.1093/pcp/pch097>
- Cheshkova, A. F. (2022). A review of hyperspectral image analysis techniques for plant disease detection and identification. *Вавиловский Журнал Генетики и Селекции*.
<https://doi.org/10.18699/VJGB2225>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates. <https://doi.org/https://doi.org/10.4324/9780203771587>
- Da Rugna, J., Colantoni, P., & Boukala, N. (2003). Hybrid color spaces applied to image database. *Proceedings of SPIE - The International Society for Optical Engineering*, 5304, 254–263. <https://doi.org/10.1117/12.526934>
- Dastres, R., & Soori, M. (2021). Artificial Neural Network Systems. *International Journal of Imaging and Robotics (IJIR)*, 21(2), 13–25. www.ceserp.com/cp-jour
- del Val, L., & de Agreda, A. P. (1993). Principal components analysis. *Atencion Primaria / Sociedad Española de Medicina de Familia y Comunitaria*, 12(6), 333–338.
<https://doi.org/10.5455/ijlr.20170415115235>
- Deng, J., Dong, W., Socher, R., Li, L.-J., Kai Li, & Li Fei-Fei. (2010). ImageNet: A large-scale hierarchical image database. *IEEE Conference on Computer Vision and Pattern Recognition, May 2014*, 248–255. <https://doi.org/10.1109/cvpr.2009.5206848>
- Ding, S., Li, H., Su, C., Yu, J., & Jin, F. (2013). Evolutionary artificial neural networks: A review. *Artificial Intelligence Review*, 39(3), 251–260. <https://doi.org/10.1007/s10462-011-9270-6>
- Divyashree, N., Ramsha, S., Sahana, S., Praveen, K., Suvarna, K., & Shamarao, J. (2015). Detection and Quantification of Disease in Cabbage using Clustering and RGB Colour Features. *International Journal of Emerging Technology in Computer Science & Electronics*, 14(2).
- Dorić, D., Nikolić-Dorić, E., Jevremović, V., & Mališić, J. (2009). On measuring skewness and kurtosis. *Quality and Quantity*, 43(3), 481–493. <https://doi.org/10.1007/s11135-007-9128-9>
- Dutta, A., & Zisserman, A. (2019). The VIA Annotation Software for Images, Audio and Video. *Proceedings of the 27th ACM International Conference on Multimedia*, 2276–2279. <http://www.robots.ox.ac.uk/>
- Elfatimi, E., Eryigit, R., & Elfatimi, L. (2022). Beans Leaf Diseases Classification Using MobileNet Models. *IEEE Access*, 10, 9471–9482.
<https://doi.org/10.1109/ACCESS.2022.3142817>
- Evgeniou, T., & Pontil, M. (2000). Statistical Learning Theory: A Primer. *International Journal of Computer Vision*, 38, 9–13. <https://doi.org/https://doi.org/10.1023/A:1008110632619>
- Fagan, B., Cowan, C. W., & Watson, P. J. (1994). The Origins of Agriculture: An International Perspective. *Journal of Field Archaeology*, 21(3). <https://doi.org/10.2307/530342>
- Fang, Z., Liu, Y., Lou, P., & Liu, G. (2004). Current trends in cabbage breeding. *Journal of New Seeds*, 6(2–3), 75–107. https://doi.org/10.1300/J153v06n02_05
- Fu, L., Feng, Y., Majeed, Y., Zhang, X., Zhang, J., Karkee, M., & Zhang, Q. (2018). Kiwifruit detection in field images using Faster R-CNN with ZFNet. *IFAC-PapersOnLine*, 51(17), 45–50. <https://doi.org/10.1016/j.ifacol.2018.08.059>

- Fu, X., Li, A., Meng, Z., Yin, X., Zhang, C., Zhang, W., & Qi, L. (2022). A Dynamic Detection Method for Phenotyping Pods in a Soybean Population Based on an Improved YOLO-v5 Network. *Agronomy*, *12*(12). <https://doi.org/10.3390/agronomy12123209>
- Gehan, M. A., Fahlgren, N., Abbasi, A., Berry, J. C., Callen, S. T., Chavez, L., Doust, A. N., Feldman, M. J., Gilbert, K. B., Hodge, J. G., Hoyer, J. S., Lin, A., Liu, S., Lizárraga, C., Lorence, A., Miller, M., Platon, E., Tessman, M., & Sax, T. (2017). PlantCV v2: Image analysis software for high-throughput plant phenotyping. *PeerJ*, *2017*(12), 1–23. <https://doi.org/10.7717/peerj.4088>
- Gignoux, C. R., Henn, B. M., & Mountain, J. L. (2011). Rapid, global demographic expansions after the origins of agriculture. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(15). <https://doi.org/10.1073/pnas.0914274108>
- Gonzales, R. C., & Woods, R. E. (2018). *Digital Image Processing* (4th ed.). Pearson PLC.
- Guo, R., Xie, J., Zhu, J., Cheng, R., Zhang, Y., Zhang, X., Gong, X., Zhang, R., Wang, H., & Meng, F. (2023). Improved 3D point cloud segmentation for accurate phenotypic analysis of cabbage plants using deep learning and clustering algorithms. *Computers and Electronics in Agriculture*, *211*, 108014. <https://doi.org/10.1016/j.compag.2023.108014>
- Han, S., & Cointault, F. (2013). Early detection of disease on leaves by image processing. *EFITA-WCCA-CIGR Conference: Sustainable Agriculture through ICT Innovations*. <http://proceedings.cigr.org/uploads/2013/0366.pdf>
- Haque, S., Lobaton, E., Nelson, N., & Yencho, C. (2021). Computer vision approach to characterize size and shape phenotypes of horticultural crops using high-throughput imagery. *Computer and Electronics in Agriculture*, *182*. <https://doi.org/https://doi.org/10.1016/j.compag.2021.106011>
- Harakannanavar, S. S., Rudagi, J. M., Puranikmath, V. I., Siddiqua, A., & Pramodhini, R. (2022). Plant leaf disease detection using computer vision and machine learning algorithms. *Global Transitions Proceedings*, *3*(1), 305–310. <https://doi.org/10.1016/j.gltip.2022.03.016>
- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., ... Oliphant, T. E. (2020). Array programming with NumPy. In *Nature* (Vol. 585, Issue 7825, pp. 357–362). Nature Research. <https://doi.org/10.1038/s41586-020-2649-2>
- Hassan, E., El-Rashidy, N., & M. Talaa, fatma. (2022). Review: Mask R-CNN Models. *Nile Journal of Communication and Computer Science*, *3*(1), 17–27. <https://doi.org/10.21608/njccs.2022.280047>
- Hastings, G. D., & Rubin, A. (2012). Colour spaces - a review of historic and modern colour models*. *African Vision and Eye Health*, *71*(3), 133–143. <https://doi.org/10.4102/aveh.v71i3.76>
- Hayashi, S., & Sakaue, O. (1996a). Cabbage head size estimation by image processing using circle recognition. *Journal of Shita*, *8*(4), 271–279.
- Hayashi, S., & Sakaue, O. (1996b). Detection of cabbage head by image processing. *Journal of Shita*, *8*(1), 20–27.

- He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). *Mask R-CNN*. Facebook Research. <https://doi.org/https://doi.org/10.48550/arXiv.1703.06870>
- He, L., Ren, X., Gao, Q., Zhao, X., Yao, B., & Chao, Y. (2017). The connected-component labeling problem: A review of state-of-the-art algorithms. *Pattern Recognition*, *70*, 25–43. <https://doi.org/10.1016/j.patcog.2017.04.018>
- Herath. (2019). Consumer Behavior and Attitudes in Purchasing Vegetables. *Agricultural Research & Technology: Open Access Journal*, *20*(2), 90–95. <https://doi.org/10.19080/ARTOAJ.2019.20.556123>
- Hidayatullah, P., & Zuhdi, M. (2015). Color-texture based object tracking using HSV color space and local binary pattern. *International Journal on Electrical Engineering and Informatics*, *7*(2), 161–174. <https://doi.org/10.15676/ijeei.2015.7.2.1>
- Huang, H., Tang, Y., Tan, Z., Zhuang, J., Hou, C., Chen, W., & Ren, J. (2022). Object-Based Attention Mechanism for Color Calibration of UAV Remote Sensing Images in Precision Agriculture. *IEEE Transactions on Geoscience and Remote Sensing*, *60*, 1–13. <https://doi.org/10.1109/TGRS.2022.3224580>
- Huang, H., Yang, A., Tang, Y., Zhuang, J., Hou, C., Tan, Z., Dananjayan, S., He, Y., Guo, Q., & Luo, S. (2021). Deep color calibration for UAV imagery in crop monitoring using semantic style transfer with local to global attention. *International Journal of Applied Earth Observation and Geoinformation*, *104*, 102590. <https://doi.org/10.1016/j.jag.2021.102590>
- Huang, Z., & Leng, J. (2010). Analysis of Hu's moment invariants on image scaling and rotation. *ICCET 2010 - 2010 International Conference on Computer Engineering and Technology, Proceedings*, *7*(October). <https://doi.org/10.1109/ICCET.2010.5485542>
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science and Engineering*, *9*(3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>
- Ilakiya, T., Parameswari, E., Davamani, V., Swetha, D., & Prakash, E. (2020). High-throughput crop phenotyping in vegetable crops. *The Pharma Innovation*, *9*(8), 184–191. <https://doi.org/10.22271/tpi.2020.v9.i8c.5035>
- Jakusch, P., Kocsis, T., Székely, I. K., & Hatvani, I. G. (2018). The application of magnetic resonance imaging (MRI) to the examination of plant tissues and water barriers. *Acta Biologica Hungarica*, *69*(4), 423–436. <https://doi.org/10.1556/018.69.2018.4.5>
- Ji, R., Ju, M., Wang, Y., Hu, C., Zhang, H., & Shi, W. (2017). In-season yield prediction of cabbage with a hand-held active canopy sensor. *Sensors (Switzerland)*, *17*(10), 1–14. <https://doi.org/10.3390/s17102287>
- Jin, C. qian, & Cai, Z. yu. (2020). A circular arc approximation algorithm for cucumber classification with image analysis. *Postharvest Biology and Technology*, *165*(October 2019). <https://doi.org/10.1016/j.postharvbio.2020.111184>
- Jollet, D., Junker-Frohn, L. V., Steier, A., Meyer-Lüpken, T., & Müller-Linow, M. (2023). A new computer vision workflow to assess yield quality traits in bush bean (*Phaseolus vulgaris* L.). *Smart Agricultural Technology*, *5*(100306). <https://doi.org/10.1016/j.atech.2023.100306>
- Jollet, D., Rascher, U., & Müller-Linow, M. (2021). Assessing yield quality parameters in bush bean via RGB imagery. *Acta Horticulturae*, *1327*(1327_56), 421–428. <https://doi.org/10.17660/ActaHortic.2021.1327.56>

- Jusoh, M. (2017). *The validation of QTL for pod quality and domestication traits through association mapping and genomics in snap beans* [University of Wisconsin-Madison]. <https://api.semanticscholar.org/CorpusID:90462794>
- K. Sing, M., & Chetia, S. (2017). Detection and classification of plant leaf diseases in image processing using MATLAB. *International Journal of Life Sciences Research*, 5(4), 2347–2812.
- Kang, Y.-S., Ryu, C. S., Kim, S. H., Jun, S. R., Jang, S. H., Park, J. W., Sarka, T. K., & Song, H. young. (2018). Yield Prediction of Chinese Cabbage (Brassicaceae) Using Broadband Multispectral Imagery Mounted Unmanned Aerial System in the Air and Narrowband Hyperspectral Imagery on the Ground. *Journal of Biosystems Engineering*, 43(2), 138–147.
- Katiyar, S. K., & Arun, P. V. (2014). A review over the applicability of image entropy in analyses of remote sensing datasets. *Photogrammetric Engineering and Remote Sensing*, 79(2b), 21–36. <http://arxiv.org/abs/1405.6133>
- Keller, W., & Borkowski, A. (2019). Thin plate spline interpolation. *Journal of Geodesy*, 93(9), 1251–1269. <https://doi.org/10.1007/s00190-019-01240-2>
- Khadabadi, G. C., Kumar, A., & Rajpurohit, V. S. (2016). Identification and classification of diseases in carrot vegetable using Discrete Wavelet Transform. *2015 International Conference on Emerging Research in Electronics, Computer Science and Technology, ICERECT 2015*, 59–64. <https://doi.org/10.1109/ERECT.2015.7498988>
- Kierdorf, J., Junker-Frohn, L. V., Delaney, M., Olave, M. D., Burkart, A., Jaenicke, H., Muller, O., Rascher, U., & Roscher, R. (2022). GrowliFlower: An image time-series dataset for GROWth analysis of cauLIFLOWER. *Journal of Field Robotics*, October. <https://doi.org/10.1002/rob.22122>
- Kim, D.-W. W., Yun, H. S., Jeong, S.-J. J., Kwon, Y.-S. S., Kim, S.-G. G., Lee, W. S., & Kim, H.-J. J. (2018). Modeling and testing of growth status for Chinese cabbage and white radish with UAV-based RGB imagery. *Remote Sensing*, 10(4), 1–25. <https://doi.org/10.3390/rs10040563>
- King, R., & Vining, K. (2019). *Breeding Snap Beans for Organic Agriculture: Genomic Shifts under Different Agricultural Management Systems Resistance and resistance breeding View project Small Fruits Genetics View project* [Oregon State University]. <https://doi.org/10.13140/RG.2.2.12594.94408>
- Kılıç, K., Boyacı, İ. H., Köksel, H., & Küsmenoğlu, İ. (2007). A classification system for beans using computer vision system and artificial neural networks. *Journal of Food Engineering*, 78(3), 897–904. <https://doi.org/10.1016/j.jfoodeng.2005.11.030>
- Kullarni, V. Y., & Sinha, P. K. (2013). Random Forest Classifier: A Survey and Future Research Directions. *International Journal of Advanced Computing*, 36(1), 1144–1156. <https://api.semanticscholar.org/CorpusID:18001819>
- Kusumam, K., Krajník, T., Pearson, S., Duckett, T., & Cielniak, G. (2017). 3D-vision based detection, localization, and sizing of broccoli heads in the field. *Journal of Field Robotics*, 34(8), 1505–1518. <https://doi.org/10.1002/rob.21726>
- Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>

- Lee, D.-J., Archibald, J. K., & Xiong, G. (2011). Rapid Color Grading for Fruit Quality Evaluation Using Direct Color Mapping. *IEEE Transactions on Automation Science and Engineering*, 8(2), 292–302. <https://doi.org/10.1109/TASE.2010.2087325>
- Li, X., Lai, T., Wang, S., Chen, Q., Yang, C., & Chen, R. (2019). Weighted feature pyramid networks for object detection. *Proceedings - 2019 IEEE Intl Conf on Parallel and Distributed Processing with Applications, Big Data and Cloud Computing, Sustainable Computing and Communications, Social Computing and Networking, ISPA/BDCLOUD/SustainCom/SocialCom 2019*, 1500–1504. <https://doi.org/10.1109/ISPA-BDCLOUD-SustainCom-SocialCom48970.2019.00217>
- Li, Z., Li, Y., Yang, Y., Guo, R., Yang, J., Yue, J., & Wang, Y. (2021). A high-precision detection method of hydroponic lettuce seedlings status based on improved Faster RCNN. *Computers and Electronics in Agriculture*, 182, 106054. <https://doi.org/10.1016/j.compag.2021.106054>
- Lindner, U., & Billmann, B. (2006). Planung, Anlage und Auswertung von Versuchen im ökologischen Gemüsebau. Handbuch für die Versuchsanstellung. In *Das Taspo Magazin*. <http://orgprints.org/id/eprint/9863>
- Liu, C., Zhang, X., Nguyen, T. T., Liu, J., Wu, T., Lee, E., & Tu, X. M. (2022). Partial least squares regression and principal component analysis: Similarity and differences between two popular variable reduction approaches. *General Psychiatry*, 35(1), 1–5. <https://doi.org/10.1136/gpsych-2021-100662>
- Lopez-Marcano, S., L. Jinks, E., Buelow, C. A., Brown, C. J., Wang, D., Kusy, B., M. Ditria, E., & Connolly, R. M. (2021). Automatic detection of fish and tracking of movement for ecology. *Ecology and Evolution*, 11(12), 8254–8263. <https://doi.org/10.1002/ece3.7656>
- Low, J. J., & Tay, L. C. (2019). Cabbage growth monitoring using time series images. *IEEE 9th International Conference on Systems Engineering and Technology*, 6–10. <https://doi.org/10.1109/ICSEngT.2019.8906367>
- Lu, W., Du, R., Niu, P., Xing, G., Luo, H., Deng, Y., & Shu, L. (2022). Soybean Yield Preharvest Prediction Based on Bean Pods and Leaves Image Recognition Using Deep Learning Neural Network Combined With GRNN. *Frontiers in Plant Science*, 12. <https://doi.org/10.3389/fpls.2021.791256>
- Luxburg, U. von, & Schölkopf, B. (2011). Statistical Learning Theory: Models, Concepts, and Results. *Handbook of the History of Logic*, 10, 651–706. <https://doi.org/10.1016/B978-0-444-52936-7.50016-1>
- Maalouf, M. (2011). Logistic regression in data analysis: An overview. *International Journal of Data Analysis Techniques and Strategies*, 3(3), 281–299. <https://doi.org/10.1504/IJDATS.2011.041335>
- Madani, R., Ayremlou, A., Amini, A., & Marvasti, F. (2010). Optimized Spline Interpolation. *IEEE Transactions on Signal Processing*. <https://doi.org/https://doi.org/10.48550/arXiv.1105.0011>
- Manninen, H., Paakki, M., Hopia, A., & Franzén, R. (2015). Measuring the green color of vegetables from digital images using image analysis. *Lwt*, 63(2), 1184–1190. <https://doi.org/10.1016/j.lwt.2015.04.005>
- Maschal, R. A., Young, S. S., Reynolds, J., Krapels, K., Fanning, J., & Corbin, T. (2010). *Review of Bayer Pattern Color Filter Array (CFA) Demosaicing with New Quality Assessment Algorithms*. <https://api.semanticscholar.org/CorpusID:120233629>

- McKinney, W. (2010). Data Structures for Statistical Computing in Python. *Proceedings of the 9th Python in Science Conference, 1(Scipy)*, 56–61. <https://doi.org/10.25080/majora-92bf1922-00a>
- Meier, U. (2018). *Entwicklungsstadien mono-und dikotyler Pflanzen BBCH Monographie*. Julius Kühn-Institut. <https://doi.org/10.5073/20180906-075119>
- Menesatti, P., Angelini, C., Pallottino, F., Antonucci, F., Aguzzi, J., & Costa, C. (2012). RGB color calibration for quantitative image analysis: The “3D Thin-Plate Spline” warping approach. *Sensors (Switzerland)*, *12*(6), 7063–7079. <https://doi.org/10.3390/s120607063>
- Methun, N. R., Yasmin, R., Begum, N., Rajbongshi, A., & Islam, M. E. (2021). Carrot Disease Recognition using Deep Learning Approach for Sustainable Agriculture. *International Journal of Advanced Computer Science and Applications*, *12*(9), 732–741. <https://doi.org/10.14569/IJACSA.2021.0120981>
- Mikołajczyk, A., & Grochowski, M. (2018). Data augmentation for improving deep learning in image classification problem. *2018 International Interdisciplinary PhD Workshop, IIPhDW 2018, August 2019*, 117–122. <https://doi.org/10.1109/IIPHDW.2018.8388338>
- Miller-Butler, M. A., Smith, B. J., Curry, K. J., & Blythe, E. K. (2019). Evaluation of Detached Strawberry Leaves for Anthracnose Disease Severity Using Image Analysis and Visual Ratings. *HortScience*, *54*(12), 2111–2117. <https://doi.org/10.21273/HORTSCI14321-19>
- Milliman, T. B., Seyednasrollah, A. M., Young, K., Hufkens, M. A., Friedl, S., Froking, A. D., Richardson, M., Abraha, D. W., Allen, M., Apple, M. A., Arain, J., Baker, J. M., Baker, D., Baldocchi, C. J., Bernacchi, J., Bhattacharjee, P., Blanken, D. D., Bosch, R., Boughton, E. H., & Zona, D. (2019). *PhenoCam Dataset v2.0: Digital Camera Imagery from the PhenoCam Network, 2000-2018*.
- Ming-Kuei Hu. (1962). Visual pattern recognition by moment invariants. *IEEE Transactions on Information Theory*, *8*(2), 179–187. <https://doi.org/10.1109/TIT.1962.1057692>
- Muller, O., Mateo, M. P. C., Albrecht, H., Keller, B., Francisco, P., Schickling, A., Müller-Linow, M., Pieruschka, R., Schurr, U., & Rascher, U. (2016). Field phenotyping infrastructure in a future world - Quantifying information on plant structure and function for precision agriculture and climate change. *Proceedings of the 13th International Conference on Precision Agriculture*.
- Murakami, N., Otsuka, K., Inoue, K., & Sugimoto, M. (1997). Image processing algorithm for developing selective harvester of cabbage. *農業機械学会誌*, *59*(2), 65–72.
- Mutka, A. M., & Bart, R. S. (2015). Image-based phenotyping of plant disease symptoms. *Frontiers in Plant Science*, *5*(1), 1–8. <https://doi.org/10.3389/fpls.2014.00734>
- Oerke, E. C., Herzog, K., & Toepfer, R. (2016). Hyperspectral phenotyping of the reaction of grapevine genotypes to *Plasmopara viticola*. *Journal of Experimental Botany*, *67*(18), 5529–5543. <https://doi.org/10.1093/jxb/erw318>
- O’Shea, K., & Nash, R. (2015). *An Introduction to Convolutional Neural Networks*. December. <http://arxiv.org/abs/1511.08458>
- Pagés, G., Deborde, C., Lemaire-Chamley, M., Moing, A., & Bonny, J. M. (2021). Correction to: MRSI vs CEST MRI to understand tomato metabolism in ripening fruit: is there a better contrast? *Analytical and Bioanalytical Chemistry*, *413*(6), 1777. <https://doi.org/10.1007/s00216-021-03180-3>

- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., ... Chintala, S. (2019). PyTorch: An imperative style, high-performance deep learning library. *Advances in Neural Information Processing Systems*, 32. <https://doi.org/https://doi.org/10.48550/arXiv.1912.01703>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Müller, A., Nothman, J., Louppe, G., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2012). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830. <http://scikit-learn.sourceforge.net>.
- Pirouz, D. M. (2012). An Overview of Partial Least Squares. *SSRN Electronic Journal, March*. <https://doi.org/10.2139/ssrn.1631359>
- Polder, G., Blokker, G., & Van Der Heijden, G. W. A. M. (2012). An ImageJ plugin for plant variety testing. *Proceedings of the ImageJ User and Developer Conference*, 168–173.
- Prabhakar, M., Purushothaman, R., & Awasthi, D. P. (2020). Deep learning based assessment of disease severity for early blight in tomato crop. *Multimedia Tools and Applications*, 79(39–40), 28773–28784. <https://doi.org/10.1007/s11042-020-09461-w>
- Rahman, A. (2014). A Review on Semiconductors Including Applications and Temperature Effects in Semiconductors. *American Scientific Research Journal for Engineering, Technology, and Sciences*, 7(1), 50–70. <http://asrjetsjournal.org/>
- Rahman, S. M. E., Mele, M. A., Lee, Y. T., & Islam, M. Z. (2021). Consumer preference, quality, and safety of organic and conventional fresh fruits, vegetables, and cereals. *Foods*, 10(1), 1–17. <https://doi.org/10.3390/foods10010105>
- Ray, S. D., Natasha, M. K. T. K., Hakim, M. A., & Nur, F. (2022). Carrot Cure: A CNN based Application to Detect Carrot Disease. *2022 6th International Conference on Trends in Electronics and Informatics, ICOEI 2022 - Proceedings*, 1043–1049. <https://doi.org/10.1109/ICOEI53556.2022.9776947>
- Raza, S.-A., Prince, G., Clarkson, J. P., & Rajpoot, N. M. (2015). Automatic Detection of Diseased Tomato Plants Using Thermal and Stereo Visible Light Images. *PLOS ONE*, 10(4), e0123262. <https://doi.org/10.1371/journal.pone.0123262>
- Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), 1137–1149. <https://doi.org/10.1109/TPAMI.2016.2577031>
- Riemann, B. (2013). *Bernard Riemann's gesammelte mathematische Werke und wissenschaftlicher Nachlass* (R. Dedekind & H. M. Weber, Eds.). Cambridge University Press. <https://doi.org/10.1017/CBO9781139568050>
- Riera, L. G., Carroll, M. E., Zhang, Z., Shook, J. M., Ghosal, S., Gao, T., Singh, A., Bhattacharya, S., Ganapathysubramanian, B., Singh, A. K., & Sarkar, S. (2020). Deep multiview image fusion for Soybean yield estimation in breeding applications. *Plant Phenomics*, 2021. <https://doi.org/10.34133/2021/9846470>
- Ronneberger, O., Fischer, P., & Brox, T. (2015, May). U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention*. <http://arxiv.org/abs/1505.04597>

- Roy, P., Dutta, S., Dey, N., Dey, G., Chakraborty, S., & Ray, R. (2014). Adaptive thresholding: A comparative study. *2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT)*, 1182–1186. <https://doi.org/10.1109/ICCICCT.2014.6993140>
- Sarstedt, M., & Mooi, E. (2014). Regression Analysis. In *Nurse Researcher* (Vol. 4, Issue 1, pp. 193–233). https://doi.org/10.1007/978-3-642-53965-7_7
- Sathya, R., & Abraham, A. (2013). Comparison of Supervised and Unsupervised Learning Algorithms for Pattern Classification. *International Journal of Advanced Research in Artificial Intelligence*, 2(2), 34–38. <https://doi.org/10.14569/ijarai.2013.020206>
- Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, 6(1). <https://doi.org/10.1186/s40537-019-0197-0>
- Simko, I., & Hayes, R. J. (2018). Accuracy, reliability, and timing of visual evaluations of decay in fresh-cut lettuce. *PLOS ONE*, 13(4). <https://doi.org/10.1371/journal.pone.0194635>
- Sliney, D. H. (2016). What is light? The visible spectrum and beyond. *Eye*, 30(2), 222–229. <https://doi.org/10.1038/eye.2015.252>
- Smith, A. R. (1978). Color gamut transform pairs. *Proceedings of the 5th Annual Conference on Computer Graphics and Interactive Techniques - SIGGRAPH '78*, 12–19. <https://doi.org/10.1145/800248.807361>
- Soleimanipour, A., & Chegini, G. R. (2019). Three-dimensional reconstruction of cucumbers using a 2D computer vision system. *Journal of Food Measurement and Characterization*, 13(1), 571–578. <https://doi.org/10.1007/s11694-018-9970-6>
- Srisha, R., & Khan, A. (2013). Morphological Operations for Image Processing : Understanding and its Applications. *NCVSComs-13, December*, 17–19.
- Stoll, M., Schultz, H. R., Baecker, G., & Berkelmann-Loehnertz, B. (2008). Early pathogen detection under different water status and the assessment of spray application in vineyards through the use of thermal imagery. *Precision Agriculture*, 9(6), 407–417. <https://doi.org/10.1007/s11119-008-9084-y>
- Sunoj, S., Igathinathane, C., Saliendra, N., Hendrickson, J., & Archer, D. (2018). Color calibration of digital images for agriculture and other applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 146(September), 221–234. <https://doi.org/10.1016/j.isprsjprs.2018.09.015>
- Suzuki, S., & Abe, K. (1985). Topological structural analysis of digitized binary images by border following. *Computer Vision, Graphics and Image Processing*, 30(1), 32–46. [https://doi.org/10.1016/0734-189X\(85\)90016-7](https://doi.org/10.1016/0734-189X(85)90016-7)
- Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. *36th International Conference on Machine Learning, ICML 2019, 2019-June*, 10691–10700.
- Tarika, B., Kaur, M., & Singh, I. (2014). Review of Histogram Equalization Techniques. *International Journal Of Engineering And Computer Science*, 3(11), 9205–9210.
- Tort, Ö. Ö., Vayvay, Ö., & Çobanoğlu, E. (2022). A Systematic Review of Sustainable Fresh Fruit and Vegetable Supply Chains. *Sustainability (Switzerland)*, 14(3), 1–38. <https://doi.org/10.3390/su14031573>

- Tripathi, M. K., & Maktedar, D. D. (2020). A role of computer vision in fruits and vegetables among various horticulture products of agriculture fields: A survey. *Information Processing in Agriculture*, 7(2), 183–203. <https://doi.org/10.1016/j.inpa.2019.07.003>
- Uyanık, G. K., & Güler, N. (2013). A Study on Multiple Linear Regression Analysis. *Procedia - Social and Behavioral Sciences*, 106, 234–240. <https://doi.org/10.1016/j.sbspro.2013.12.027>
- van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., & Yu, T. (2014). scikit-image: image processing in Python. *PeerJ*, 2(e453). <https://doi.org/10.7717/peerj.453>
- Van Eck, J. W., Van Der Heijden, G. W. A. M., & Polder, G. (1998). Accurate measurement of size and shape of cucumber fruits with image analysis. *Journal of Agricultural and Engineering Research*, 70(4), 335–343. <https://doi.org/10.1006/jaer.1998.0285>
- Velimirovic, A., Jovovic, Z., & Pržulj, N. (2021). From neolithic to late modern period: Brief history of wheat. *Genetika*, 53(1). <https://doi.org/10.2298/GENSR2101407V>
- Vinoth Kumar, K., & Jayasankar, T. (2019). An Identification of Crop Disease Using Image Segmentation. *International Journal of Pharmaceutical Sciences and Research*, 10(3), 1054. [https://doi.org/10.13040/IJPSR.0975-8232.10\(3\).1054-64](https://doi.org/10.13040/IJPSR.0975-8232.10(3).1054-64)
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., ... Vázquez-Baeza, Y. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3), 261–272. <https://doi.org/10.1038/s41592-019-0686-2>
- Visa, S., Ramsay, B., Ralescu, A., & van der Knaap, E. (2011). Confusion Matrix-based Feature Selection. *Proceedings of the 22nd Midwest Artificial Intelligence and Cognitive Science Conference*.
- Wang, D., Vinson, R., Holmes, M., Seibel, G., Bechar, A., Nof, S., & Tao, Y. (2019). Early Detection of Tomato Spotted Wilt Virus by Hyperspectral Imaging and Outlier Removal Auxiliary Classifier Generative Adversarial Nets (OR-AC-GAN). *Scientific Reports*, 9(1), 4377. <https://doi.org/10.1038/s41598-019-40066-y>
- Wang, G., Sun, Y., & Wang, J. (2017). Automatic Image-Based Plant Disease Severity Estimation Using Deep Learning. *Computational Intelligence and Neuroscience*, 2017. <https://doi.org/10.1155/2017/2917536>
- Wang, Q., & Qi, F. (2019). Tomato Diseases Recognition Based on Faster RCNN. *2019 10th International Conference on Information Technology in Medicine and Education (ITME)*, 772–776. <https://doi.org/10.1109/ITME.2019.00176>
- Wang, S. X., Hu, R. F., Gao, K., Wali, F., Zan, G. Bin, Wang, D. J., Pan, Z. Y., & Wei, S. Q. (2017). Non-destructive study of fruits using grating-based X-ray imaging. *Nuclear Science and Techniques*, 28(2), 1–4. <https://doi.org/10.1007/s41365-016-0169-4>
- Wiley, V., & Lucas, T. (2018). Computer Vision and Image Processing: A Paper Review. *International Journal of Artificial Intelligence Research*, 2(1), 22. <https://doi.org/10.29099/ijair.v2i1.42>
- Wilhelm, J., Wojciechowski, T., Postma, J. A., Jollet, D., Heinz, K., Böckem, V., & Müller-Linow, M. (2022). Assessing the Storage Root Development of Cassava with a New Analysis Tool. *Plant Phenomics*, 2022. <https://doi.org/10.34133/2022/9767820>

- Wu, Y., Kirillov, A., Massa, F., Lo, W.-Y., & Girshick, R. (2019). *Detectron2* (0.2.1). <https://github.com/facebookresearch/detectron2>
- Yang, C. J., Samayoa, L. F., Bradbury, P. J., Olukolu, B. A., Xue, W., York, A. M., Tuholski, M. R., Wang, W., Daskalska, L. L., Neumeyer, M. A., de Jesus Sanchez-Gonzalez, J., Romay, M. C., Glaubitz, J. C., Sun, Q., Buckler, E. S., Holland, J. B., & Doebley, J. F. (2019). The genetic architecture of teosinte catalyzed and constrained maize domestication. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(12). <https://doi.org/10.1073/pnas.1820997116>
- Zeder, M. A. (2011). The origins of agriculture in the Near East. *Current Anthropology*, *52*(SUPPL. 4). <https://doi.org/10.1086/659307>
- Zeng, Q., Ma, X., Cheng, B., Zhou, E., & Pang, W. (2020). GANS-based data augmentation for citrus disease severity detection using deep learning. *IEEE Access*, *8*, 172882–172891. <https://doi.org/10.1109/ACCESS.2020.3025196>
- Zhang, J., Wang, X., Liu, J., Zhang, D., Lu, Y., Zhou, Y., Sun, L., Hou, S., Fan, X., Shen, S., & Zhao, J. (2022). Multispectral Drone Imagery and SRGAN for Rapid Phenotypic Mapping of Individual Chinese Cabbage Plants. *Plant Phenomics*, *2022*. <https://doi.org/10.34133/plantphenomics.0007>
- Zhang, T. Y., & Suen, C. Y. (1984). A fast parallel algorithm for thinning digital patterns. *Communications of the ACM*, *27*(3), 236–239. <https://doi.org/10.1145/357994.358023>
- Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(11), 1330–1334. <https://doi.org/10.1109/34.888718>
- Zheng, B., Sun, G., Meng, Z., & Nan, R. (2022). Vegetable Size Measurement Based on Stereo Camera and Keypoints Detection. *Sensors*, *22*(1617). <https://doi.org/10.3390/s22041617>
- Zhu, H., Chu, B., Zhang, C., Liu, F., Jiang, L., & He, Y. (2017). Hyperspectral Imaging for Presymptomatic Detection of Tobacco Disease with Successive Projections Algorithm and Machine-learning Classifiers. *Scientific Reports*, *7*(4125). <https://doi.org/10.1038/s41598-017-04501-2>
- Zhu, H., Deng, L., Wang, D., Gao, J., Ni, J., & Han, Z. (2019). Identifying carrot appearance quality by transfer learning. *Journal of Food Process Engineering*, *42*(6), 1–10. <https://doi.org/10.1111/jfpe.13187>

9 Abbildungsverzeichnis

Abbildung 1: Phänotypischen Unterschiede der Pflanzen (A), sowie der Fruchtkörper (B) zwischen Teosinte und Mais. Quelle: (Yang et al., 2019).....	8
Abbildung 2: Das elektromagnetische Spektrum. Quelle: https://upload.wikimedia.org/wikipedia/commons/6/62/Electromagnetic_spectrum_-de_c.svg (30.12.2023).....	14
Abbildung 3: Skalierte 3-Dimensionale Repräsentationen des a) RGB, b) HSV und c) CIE-LAB Farbraums.	16
Abbildung 4: Beispiele für klassische Bildtransformationen. a) Originales Graubild; b) Skalierung; c) Zoom; d) Rotation; e) Spiegelung; f) Translation; g) Histogramm Entzerrung.	18
Abbildung 5: a) Schema zur Funktionsweise einer 3x3 Faltungsmatrix. Die Faltungsmatrix wird mit dem Zentrum über das Initialpixel (Rot) gelegt und jeweils die entsprechenden Elemente der Bildmatrix (grün) und der Faltungsmatrix miteinander multipliziert und anschließend aufsummiert, woraus sich der neue Wert für das Initialpixel ergibt. Abbildung inspiriert aus dem GIMP Benutzerhandbuch (https://www.gimp-handbuch.de/). b) Beispiele für 3x3 Faltungsmatrizen für 3 einfache Bildfilter	18
Abbildung 6: a) Binärbild eines Apfels; b) Kontur eines Apfels; c) Skelett eines Apfels	21
Abbildung 7: Struktur eines vollständig vernetzten künstlichen neuronalen Netzwerks (Dastres & Soori, 2021).....	26
Abbildung 8: Struktur eines neuronalen Faltungsnetzwerks zur Bildklassifikation (O'Shea & Nash, 2015).....	27
Abbildung 9: a) Struktureller Aufbau eines U-Net CNNs zur Segmentierung von Bildern; b) Struktureller Aufbau eines R-CNNs. Dieses besteht aus mehreren miteinander verbundenen CNNs zur Objekterkennung und Klassifizierung.	28
Abbildung 10: Referenzskala zur Bewertung der Krümmung (Lindner & Billmann, 2006). Der Blütenansatz der Hülsen liegt jeweils am oberen Bildrand.....	34
Abbildung 11: Diffus beleuchtete Aufnahmebox mit einer im LED-Panel verbauten Industriekamera, welche über einen externen Rechner gesteuert wird; Bei geschlossener Tür können bis zu 30 Bohnenhülsen in einer Aufnahme erfasst werden.	36
Abbildung 12: Vorverarbeitung der Bohnenbilder, welche in einer für diesen Zweck entwickelten Fotobox aufgenommen werden. Zunächst werden die Bilder gerade ausgerichtet und zugeschnitten und, falls vorhanden, der ColorChecker aus dem Bild entfernt und die Helligkeitsverteilung mittels einer Gaussfunktion optimiert.	37
Abbildung 13: Generierung der Zentrallinie aus der a) Bildmaske einer Hülse; b) Skelett nach Zhang & Suen (1984); c) Zentrallinie nach iterativer Entfernung von Ästen; d) Zentrallinie nach Subtraktion der Spitze und des Blütenansatzes.	40
Abbildung 14: Gewichtungsfunktion zur Berechnung der Krümmung für eine hypothetische Bohnenhülse mit einer Länge der Zentrallinie von 600 Pixeln. Sie dient dazu die Computerberechnungen der Krümmung an die Züchterbewertung anzugleichen, welche Krümmungen an den Enden der Hülse weniger stark bewerten	41
Abbildung 15: Wahrheitsmatrizen zur a) Erkennung der Blütenansätze und b) Erkennung der Spitzen durch ein trainiertes Mask R-CNN; c) Beispiele der Detektion bzw. Segmentierung durch das Mask R-CNNs für I-III: Richtig-Positive Blütenansatzerkennung (rote Maske), IV-VI: Richtig-Negative Blütenansatzerkennung, VII-IX: Falsch-Negative Blütenansatzerkennung, X-XI: Falsch-Positive Blütenansatzerkennung, XII-XIV: Richtig-	

Positive Spitzenerkennung (cyane Maske), XV-XVII: Richtig-Negative Spitzenerkennung, XVIII-XX: Falsch-Negative Spitzenerkennung, XXI-XXIII: Falsch-Positive Spitzenerkennung.	42
Abbildung 16: Vergleich zwischen den Referenzmessungen und den Messungen der MS-Methoden für a) die Hülsenlänge, b) das Kaliber und c) die Hülsenkrümmung (n=399). Der mittlere absolute Fehler in Prozent (MAPE) sowie der Korrelationskoeffizient (R) sind für die drei Merkmale angegeben. Ausreißer (graue Punkte) zeigen für a) Falsch-Negative Spitzenerkennungen des Mask R-CNN (n=17) und für b) Hülsen der Bohnentypen Oval und Flach, welche auf Grund der Krümmung ihre flache Seite zur Kamera zeigen (n=26). Ausreißer wurden von der Berechnung der Regression ausgeschlossen. Sterne in c) zeigen Box-Plots mit geringer Stückzahl (n<5).....	45
Abbildung 17: Transmissionsaufnahme von gekben Bohnenhülsen unter Nutzung einer LED-Flächenleuchte	49
Abbildung 18: Evaluierung der Analyse der Samenanlage mittels Faster R-CNN. Die Ergebnisse sind für die gelbe Bohnensorte (a, b, n=60) und die grüne Bohnensorte (c, d, n=57) aufgeteilt, wobei a) und c) die Ergebnisse für die Erkennung der Samen und Fehlstellen und b) und d) die Ergebnisse für die Erkennung von Schwanenhälsen zeigen. ...	51
Abbildung 19: Faster R-CNN-Vorhersagen für Samen (blaue Rechtecke), Fehlstellen (grüne Rechtecke) und Schwanenhälse (rote Rechtecke). Die Werte in den Kästchen geben die Konfidenzwerte an. a) und b) zeigen gelbe Bohnenhülsen, c) eine dunklere grüne BohnenHülse mit größeren Samen und d) eine hellere grüne BohnenHülse mit kleineren Samen.....	51
Abbildung 20: a) Referenzskala zur visuellen Bewertung der Helligkeit grüner Bohnenhülsen; b) Referenzskala zur Beurteilung der Farbe von Bohnenhülsen.....	56
Abbildung 21: Zusammenhang zwischen den Züchterbewertungen und den Messungen der MS-Methode für die Helligkeit von grünen Bohnenhülsen. Die orangen Sterne deuten ovale oder flache Hülsen an, welche mit der Schmalen Seite zur Kamera ausgerichtet waren, während die visuellen Bewertungen an einer der breiten Seite durchgeführt wurden	57
Abbildung 22: Infektionslevel einzelner Möhrenblätter mit <i>Alternaria dauci</i> nach Züchterskala. Diese reicht von 1, abgestorbene Blätter, bis 9, vollkommen gesunde Blätter ohne jegliche Läsionen. Diese Skala beruht auf Erfahrungswerten sehr erfahrener Züchter.	59
Abbildung 23: Aufnahmebox für den Teilbereich Möhre. Das LED-Panel der Aufnahmebox wurde mit einem Camping Akku betrieben, wobei dessen Laufzeit mit einem Solarmodul verlängert werden kann. Die Aufnahmen wurden mit einer Digitalkamera gemacht, welche durch ein Loch im LED-Panel auf die Grundfläche der Aufnahmebox gerichtet war.....	60
Abbildung 24: Vergleich der visuellen Bewertung des Infektionsstadiums einzelner Möhrenblätter mit der Vorhersage durch eine MS-Segmentierungsmethode im HSV-Farbraum (a), einer Random Forest Klassifikation (b) und der Vorhersage eines EfficientNet V2 (c). Angegeben sind jeweils der mittlere absolute Fehler (MAE) für die einzelnen Infektionsstadien, sowie für die gesamten Daten, der Korrelationskoeffizient R und die Korrekturklassifikationsrate.....	64
Abbildung 25: Aufnahme einer Kohlpflanze (a) zusammen mit der Aufnahme eines selbstentwickelten RHS-Farbpanels (b), um nachträgliche eine präzise Farbkalibrierung der Aufnahmen durchführen zu können.....	68
Abbildung 26: Eigenentwickeltes RHS-Farbpanel mit 144 Referenzfarben des Colour Charts Edition V der Royal Horticulture Society. Die wahren Farbwerte der einzelnen Farbfelder sind über einen Link aufrufbar, welche in dem großen QR-Code gespeichert ist....	70

Gegenüberstellung der Ergebnisse der Farbkalibrierungsreihe mittels unterschiedlicher Interpolationsverfahren (a) und eine detaillierte Darstellung der Ergebnisse bei einer Weißabgleichstemperatur von 6500°Kelvin	71
Abbildung 28: Ergebnisse der Farbklassifizierung für die Umblätter der einzelnen Kohlsorten ohne Farbkalibrierung (jeweils linkes Balkendiagramm) und mit einer TPS-Farbkalibrierung (jeweils rechtes Balkendiagramm). Die farb-kodierten Abschnitte der Balkendiagramme geben die Anteile der entsprechenden Farbzuordnung an (n=10 Pflanzen pro Sorte).	72
Abbildung 29: Feldversuchsaufbau mit 6 unterschiedlichen Sorten, welche zur besseren Abgrenzung von Rotkohlpflanzen umschlossen wurden (eine Farbe pro Sorte).	76
Abbildung 30: a) Fieldbike während Aufnahmen am Standort Marne. Die drei Industriekameras werden über einen Akku betrieben und über ein Laptop gesteuert; b) Screenshot der Benutzeroberfläche zur parallelen Ansteuerung der Kameras. Die Software komplettierte die Funktionalität des Fieldbikes und war auf maximale Benutzerfreundlichkeit ausgelegt, um auch von Nichtfachpersonal genutzt werden zu können. Im oberen Bereich befinden sich die Ansteuerungsfelder der drei Kameras und im unteren Bereich Einstellungsmöglichkeiten wie der Speicherpfad, das Dateiformat sowie den Gain zu Anpassung der Helligkeit im Bild.	77
Abbildung 31: Manuell annotierte Masken für den Kohlkopf (rot) sowie für das Umblatt (grün) aus dem Validierungsdatensatz. Die Kopfmasken wurden für die Ertragsanalyse und die Umblattmasken für die Farbanalysen genutzt.....	78
Abbildung 32: Gegenüberstellung des Kohlkopfgewichts zur a) manuell annotierten projizierten Kopffläche und b) manuell annotierten projizierten Kopffläche multipliziert mit der Dichte des Kopfes (nach Züchterbewertung); Die Daten sind nach den einzelnen Kopfformtypen rund, spitz und flach aufgeteilt und es wird jeweils der Korrelationskoeffizient R sowie der mittlere absolute Fehler in Prozent (MAPE) angegeben.....	80
Abbildung 33: Vergleich der projizierten Kopffläche zwischen manuell annotierten und von einem trainierten Mask R-CNN vorhergesagten Kopfmasken; Die Daten sind nach den einzelnen Kopfformtypen rund, spitz und flach aufgeteilt und es wird jeweils der Korrelationskoeffizient R sowie der mittlere Absolute Fehler in Prozent (MAPE) angegeben; Die gestrichelte Linie zeigt die Identitätslinie; Der orangene Balken indiziert den 15% MAPE Bereich, welcher aus Züchtersicht einen akzeptablen Bereich darstellt; n=77	81
Abbildung 34: Vorhersagen des trainierten Mask R-CNNs für den Kohlkopf (weiß) und die gesamte Pflanze (blau); a) zeigt eine starke und b) eine schwache Vorhersage des Netzwerks.	82

10 Tabellenverzeichnis

Tabelle 1: Prozentuale mittlere Kohlkopfdichte nach Züchterbewertung. Die mittleren Dichten und Standardabweichungen (StAbw) sind für die 4 runde (R1-R4), zwei flache (FL1, FL2) und 2 spitze (P1, P2) Kohlsorten des Validierungsdatensatz angegeben (n=10/Sorte).....78