

Development of Automated Approaches to High-Throughput Plant Image Analysis

Dissertation am Promotionszentrum IWIT (Ingenieurwissenschaften und Informationstechnologien) der HAW in Sachsen-Anhalt | Hochschule Harz

Narendra Narisetti



Harzer Hochschultexte | Impressum

ISSN (Print): 1619-7232
ISSN (Online): 2627-5708



Zitierbar als

Narisetti, Narendra: Development of Automated Approaches to High-Throughput Plant Image Analysis, Harzer Hochschultexte Nr. 18, Hochschule Harz, Wernigerode, 2025.

Die Harzer Hochschultexte sind lizenziert unter der freien Lizenz Creative Commons Namensnennung 4.0 [CC BY 4.0]. Für weitere Informationen besuchen Sie bitte <https://creativecommons.org/licenses/by/4.0/deed.de>.

Die in diesem Werk enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons-Lizenz, falls sich aus der Legende nichts anderes ergibt. Sofern das Material nicht unter der genannten Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen. Die Wiedergabe von allgemein beschreibenden Bezeichnungen, Marken, Unternehmensnamen etc. in diesem Werk bedeutet nicht, dass diese frei durch jedermann benutzt werden dürfen. Die Berechtigung zur Benutzung unterliegt, auch ohne gesonderten Hinweis hierzu, den Regeln des Markenrechts.

Inhaltlich verantwortlich

Inhaltlich verantwortlich sind die auf dem Cover benannten Autor*innen.

Institution

Die Hochschule Harz ist eine Körperschaft des öffentlichen Rechts.
Sie wird durch den Rektor Prof. Dr. Folker Roland gesetzlich vertreten.

Hochschule Harz
Friedrichstraße 57-59
38855 Wernigerode

E-Mail: info@hs-harz.de
Webseite: www.hs-harz.de
Umsatzsteuer-Identifikationsnummer: DE231052095

Herausgeber / Kontakt

Prorektor für Forschung und Chancengleichheit
Prof. Dr. Frieder Stolzenburg

Tel.: +49 3943 659 333
E-Mail: fstolzenburg@hs-harz.de

Aufsichtsbehörde

Die für die Hochschule Harz zuständige Aufsichtsbehörde ist das Ministerium für Wissenschaft, Energie, Klimaschutz und Umwelt des Landes Sachsen-Anhalt (MWU), Leipziger Straße 58, 39112 Magdeburg.

Haftungsausschluss und Herausgeberhinweise

Die Hochschule Harz ist lediglich für die Veröffentlichung der einzelnen Werke zuständig, sie übernimmt keinerlei Haftung.

Vielmehr gilt Folgendes:

- für den Inhalt der Publikation ist/sind der/die Autor*innen verantwortlich.
- mit der Erfassung in der Schriftenreihe „Harzer Hochschultexte“ verbleiben die Urheberrechte bei dem/der/den Autor*innen
- die Einhaltung von Urheber- und Verwertungsrechten Dritter liegt in der Verantwortung des/der Autor*innen

Vor Veröffentlichung bestätigte/n der/die Autor*innen

- dass mit der Bereitstellung der Publikation und jedes Bestandteils (z.B. Abbildungen) nicht gegen gesetzliche Vorschriften verstoßen wird und Rechte Dritter nicht verletzt werden,
- dass im Falle der Verwendung personenbezogener Daten der Datenschutz (durch Einholen einer Einwilligung des Dritten zur Veröffentlichung und Verbreitung) beachtet wurde,
- dass im Falle einer bereits erfolgten Veröffentlichung (z.B. bei einem Verlag) eine Zweitveröffentlichung dem Verlagsvertrag nicht entgegensteht sowie
- dass die Hochschule Harz von etwaigen Ansprüchen Dritter (z.B. Mitautor*innen, Miturheber*innen, Verlage) freigestellt ist.

DEVELOPMENT OF AUTOMATED APPROACHES TO HIGH-THROUGHPUT PLANT IMAGE ANALYSIS

DISSERTATION

for obtaining the academic degree of

Dr. rer. nat.

accepted at the Doctoral Center for Engineering Sciences and
Information Technologies at the Harz University of Applied
Sciences

Submitted by

M.Sc., Narendra Narisetti

Supervisors:

1. Prof. Dr. Frieder Stolzenburg, Harz University of Applied
Sciences
2. Dr. Evgeny Gladilin, Leibniz Institute of Plant Genetics
and Crop Plant Research

Acceptance of the dissertation was recommended by:

1. Prof. Dr. Korinna Bade, Anhalt University of Applied
Sciences
2. Assoc. Prof. Dr. Oliver Obst, Western Sydney University

The dissertation is submitted on May 22, 2024, and awarding
of the academic degree takes place with passing the oral
defense at Harz University of Applied Sciences on February 14,
2025 with overall assessment **magna cum laude**.

Publication place and year: Wernigerode, March 2025

Acknowledgement

First and foremost, I am incredibly grateful to my supervisors, Prof. Dr. Frieder Stolzenburg and Dr. Evgeny Gladilin. I am thankful for their invaluable advice, continuous support and patience during my PhD study. Their immense knowledge and abundant experience have encouraged me throughout my academic research and daily life. I would also like to thank Dr. Michael Henke and Archana Chelamalashetty for their technical support in my study. I want to thank all the members of the Molecular Genetics Department, IPK Gatersleben. Their kind help and support have made my study and life in Germany wonderful. Finally, I want to thank my parents, in-laws, and my wife for their tremendous understanding and encouragement over the past few years in completing my doctoral dissertation.

Narendra Narisetti

.....
Author's signature

Abstract (Deutsch)

Die bildbasierte Phänotypisierung von ganzen Pflanzen und Pflanzenorganen mit hochauflösenden optischen Sensoren hat sich als bevorzugte Methode zur nicht-destruktiven und hochdurchsatzfähigen Erfassung quantitativer Pflanzenmerkmale etabliert. Aufgrund verschiedener technischer und natürlicher Faktoren unterliegt das optische Erscheinungsbild von Pflanzen- und Hintergrundstrukturen erheblichen Schwankungen. Folglich stellt die automatisierte Analyse großer phänotypischer Daten ein anspruchsvolles Problem dar. Die inhärente Komplexität phänotypischer Bilddaten erfordert höhere kognitive Ansätze für die Musterklassifikation und Segmentierung, wie beispielsweise Convolutional Neural Networks (CNN). Allerdings erfordern die fortschrittlichen CNN-Methoden eine große Menge an repräsentativen Referenzdaten (sog. ground truth), was einen Flaschenhals für ihre direkte Anwendung darstellt. Folglich befasst sich diese Arbeit mit der Untersuchung und Entwicklung halb- und vollautomatischer Bildverarbeitungsansätze zur Segmentierung und Erkennung von komplexen pflanzlichen Strukturen in Anwendung zur Hochdurchsatz-Pflanzenphänotypisierung. Der anwendungsorientierte Aufwand dieser Arbeit liegt insbesondere in der Analyse großer Bilddaten von großen Hochdurchsatz-Bildgebungsplattformen am IPK, einschließlich mehrerer Gewächshausanlagen zur Spross- und Wurzelphänotypisierung. Beginnend mit halbautomatischen Ansätzen wurde eine Genauigkeit von mehr als 80% bei der Wurzel- und Grannenbildsegmentierung erreicht. Aufbauend auf den ground truth Daten aus der semi-automatisierter Bildsegmentierung wurden U-net CNN Modelle entwickelt und trainiert, die bei der vollständigen Automatisierung eine Genauigkeit von mehr als 90% in Wurzel und Sprossegmentierung erreicht haben. In diesen Anwendungen konnte gezeigt werden, dass die vorgeschlagenen U-net-basierten Bildsegmentierungsmodelle verschiedenen modernen CNN-Methoden überlegen sind. Darüber hinaus sind die hier vorgeschlagenen CNN-Modelle weniger komplex und besser in der Lage, unterschiedliche optische Szenen zu segmentieren. Zusammenfassend zeigt diese Arbeit das große Potenzial neuartiger Deep-Learning-Ansätze für die vollautomatische Verarbeitung großer Bilddatenmengen auf, was wesentlich dazu beiträgt, eine Lücke in der Hochdurchsatzpflanzenphänotypisierung und Biodaten-basierenden Pflanzenforschung zu schließen.

Abstract (English)

Image-based phenotyping of whole plants and plant organs with high-resolution optical sensors became a method of choice for assessing quantitative plant traits in a non-destructive, high-throughput manner. Due to several technical and natural factors, optical appearance of plant and background structures undergoes substantial variations. Consequently, automated analysis of large phenotypic data renders a challenging problem. The intrinsic complexity of phenotypic image data requires higher-level cognitive approaches to pattern classification and segmentation, such as convolutional neural networks (CNN). However, the advanced CNN methods require a large amount of representative ground truth data, which poses a bottle-neck for their straightforward application. Consequently, this thesis deals with investigating and developing semi-and fully automated image processing approaches to plant image segmentation and pattern detection in application to high-throughput plant phenotyping. In particular, the application-oriented effort of this work is on the analysis of large image data from major IPK high-throughput imaging platforms, including multiple greenhouse shoot and root phenotyping facilities. Starting with semi-automated approaches, more than 80% accuracy of root and spike image segmentation was achieved. Based on the ground truth data from the semi-automated image segmentation, U-net CNN models were developed and trained, which achieved an accuracy of more than 90% for fully automated root and shoot segmentation. In these applications, the proposed U-net-based image segmentation models were shown to be superior compared to various CNN architectures and state-of-the-art methods. Furthermore, the proposed CNN models are less complex and more capable of segmenting different optical scenes including greenhouse and field-like images. Ultimately, this thesis shows a great potential of novel deep learning approaches to fully automated processing of large image data which essentially helps to bridging a gap in high-throughput plant phenotyping and bio-data driven research.

Appended Publications and Author’s Contribution

Table 1: Overview of the five published articles that are part of this cumulative thesis. The contribution of Narendra Nariseti (NN), the author of this thesis, is stated.

Article	Contribution	Impact factor
Narisetti, N., Henke, M., Seiler, C. et al. Semi-automated Root Image Analysis (saRIA). Sci Rep 9, 19674 (2019). 10.1038/s41598-019-55876-3	NN contributed significantly to the concept, methodology, method implementation, validation and analysis, investigation, data curation, preparation of the manuscript, review and editing, and visualization.	4.38 (according to SCImago Journal Rank, January 2024)
Narisetti, N., Henke, M., Seiler, C. et al. Fully-automated root image analysis (faRIA). Sci Rep 11, 16047 (2021). 10.1038/s41598-021-95480-y	NN contributed significantly to the concept, methodology, method implementation, validation and analysis, investigation, data curation, preparation of the manuscript, review and editing, and visualization.	4.38 (according to SCImago Journal Rank, January 2024)
Narisetti N, Henke M, Neumann K, Stolzenburg F, Altmann T and Gladilin E (2022) Deep Learning Based Greenhouse Image Segmentation and Shoot Phenotyping (DeepShoot). Front. Plant Sci. 13:906410. doi: 10.3389/fpls.2022.906410	NN contributed significantly to the concept, methodology, method implementation, validation and analysis, investigation, data curation, preparation of the manuscript, review and editing, and visualization.	6.6 (according to Clarivate releases the Journal Citation Report (JCR 2023))

Continued on next page...

Table 1: continued from previous page

Article	Contribution	Impact factor
Narisetti N, Neumann K, Röder MS and Gladilin E (2020) Automated Spike Detection in Diverse European Wheat Plants Using Textural Features and the Frangi Filter in 2D Greenhouse Images. <i>Front. Plant Sci.</i> 11:666. doi: 10.3389/fpls.2020.00666	NN contributed significantly to the concept, methodology, method implementation, validation and analysis, investigation, data curation, preparation of the manuscript, review and editing, and visualization.	6.6 (according to Clarivate releases the Journal Citation Report (JCR 2023))
Narisetti N, Awais M, Khan M, Stolzenburg F, Stein N, Gladilin E. Awn Image Analysis and Phenotyping Using BarbNet. <i>Plant Phenomics</i> 2023;0081. DOI: 10.34133/plantphenomics.0081	NN contributed significantly to the concept, methodology, method implementation, validation and analysis, investigation, data curation, preparation of the manuscript, review and editing, and visualization.	6.96 (according to Clarivate releases the Journal Citation Report (JCR 2023))

List of Contents

Acknowledgement	II
Abstract (Deutsch)	III
Abstract (English)	IV
Appended Publications and Author’s Contribution	V
List of Contents	VII
List of Tables	X
List of Figures	XI
Abbreviations	XIV
Symbols	XVII
1 Introduction	1
1.1 Plant Phenotyping	1
1.2 State-of-the-Art and Research Requirements	3
1.3 Convolutional Neural Network (CNN)	10
1.4 Research Objectives	11
2 Semi-automated Root Image Analysis (saRIA)	12
2.1 Introduction	14
2.2 Methods	15
2.2.1 Image Data	15
2.2.2 Image Analysis	17
2.3 Results	20
2.3.1 Semi-automated Segmentation of Soil-Root Images	20
2.3.2 Evaluation of Phenotypic Traits vs. SmartRoot	23
2.3.3 Visualization of Root Features	24
2.4 Discussion & Conclusion	25
Bibliography	42
3 Fully-automated Root Image Analysis (faRIA)	47
3.1 Introduction	49
3.2 Methods & Computational Setup	52
3.2.1 Deep CNN Model for Root Image Segmentation	52

3.2.2	Performance Metrics	54
3.2.3	Data and Image Annotation	56
3.2.4	Training	57
3.2.5	Prediction	58
3.3	Results	60
3.3.1	Training and Validation of faRIA	60
3.3.2	Evaluation of faRIA vs. SegRoot	61
3.3.3	Segmentation of Further Image Modalities	63
3.3.4	Evaluation of Phenotypic Traits vs. saRIA	63
3.3.5	Graphical User Interface and Runtime	64
3.4	Discussion & Conclusion	66
	Bibliography	79
4	Deep Learning Based Greenhouse Image Segmentation and Shoot Phenotyping (DeepShoot)	86
4.1	Introduction	88
4.2	Materials and Methods	90
4.2.1	Image Data	90
4.2.2	Ground Truth Generation	92
4.2.3	Image Segmentation using CNN	92
4.2.4	Performance Metrics	94
4.2.5	Computational Implementation	95
4.2.6	Method Comparison	98
4.3	Results	99
4.3.1	Training and Validation	99
4.3.2	Evaluation on the Reference Dataset	101
4.3.3	Evaluation of DeepShoot vs. Alternative Solutions	102
4.3.4	DeepShoot GUI Tool	104
4.4	Discussion & Conclusion	106
	Bibliography	113
5	Automated Spike Detection in Diverse European Wheat Plants Using Textural Features and the Frangi Filter in 2D Greenhouse Images	118
5.1	Introduction	120
5.2	Materials and Methods	122
5.2.1	Dataset	122
5.2.2	Methodology	123
5.3	Results and Discussion	131
5.3.1	Spike Emergence	132
5.3.2	Spike Area	135

5.4	Conclusion	139
	Bibliography	147
6	Awn Image Analysis and Phenotyping Using BarbNet	151
6.1	Introduction	153
6.2	Materials and Methods	155
6.2.1	Plant Material	155
6.2.2	Image Data	155
6.2.3	Genotyping	155
6.2.4	Ground Truth Generation	156
6.2.5	Barb Detection using CNN	156
6.2.6	Performance Metrics	159
6.2.7	Computational implementation	159
6.3	Results	161
6.3.1	Training and Validation of BarbNet	161
6.3.2	Evaluation and Comparison of U-net Models	161
6.3.3	Evaluation of Phenotypic Traits	163
6.3.4	Genotypic-Phenotypic Classification	165
6.4	Discussion & Conclusion	167
	Bibliography	176
7	Synthesis of Results	180
7.1	Addressing Research Objectives	180
7.1.1	Objective 1	180
7.1.2	Objective 2	182
7.1.3	Objective 3	184
7.1.4	Objective 4	185
7.1.5	Objective 5	187
7.1.6	Objective 6	188
7.2	Conclusion & Future Work	189
	Bibliography	190
	Curriculum Vitae	197

List of Tables

1	Overview of the five published articles that are part of this cumulative thesis. The contribution of Narendra Nariseti (NN), the author of this thesis, is stated.	V
1	continued from previous page	VI
2.1	Comparison of saRIA image segmentation quality	22
2.2	Feature comparison of saRIA with other software	29
3.1	Convolutional parameters of U-net models	53
3.2	Details of all encoder blocks and corresponding max-pool layer output	55
3.3	Details of all decoder blocks and corresponding transpose convolutional layers	55
3.4	Training parameters of SegRoot, faRIA:1024 and faRIA:256	62
3.5	Training performance of SegRoot, faRIA:1024 and faRIA:256	62
4.1	Convolutional parameters of U-net models	93
4.2	Overview of training data	95
4.3	Overview of growth stages of training data	96
4.4	Training and validation performance of the shoot models	99
4.5	Computational time of shoot segmentation algorithms	104
4.6	Complexity of alternative U-net models	106
5.1	The statistical performance of the neural network in the training stage	129
5.2	Possible structural patterns in 2D images depending on eigen values	130
5.3	Gene classification and comparison of architectural features of 369 plants	136
6.1	Convolutional parameters of the original U-net and BarbNet	158
6.2	Training performance of U-net models	163
6.3	Mean Dice coefficient of U-net models	163
6.4	F1-score of four genotypes with count and area features	165
6.5	F1-score of 4 genotypes with count and length features	166

List of Figures

1.1	Aim of plant phenotyping: Analyze the interaction between genotype and environmental conditions that promote a particular phenotype [1].	2
1.2	Image-based plant phenotyping: (a) Exemplary Arabidopsis shoot system captured in the greenhouse (b) segmented target plant pixels from background pixels. The morphology of the segmented shoot system, called traits, is calculated to study genotype-phenotype characterization.	3
1.3	Segmentation of conventional thresholding techniques on root images from different setups: (a) Barley plant roots in the soil environment (b) Rapeseed roots grown in transparent containers filled with agar gel (c) scanned washed Maize roots on an Epson Expression 10 000 XL scanner. Segmentation of roots from the soil background in the image (d) is very noisy compared to the other setups in (e) and (f).	4
1.4	Histograms of root (blue) and non-root (green) intensity values of different imaging modalities shown in Figure 1.3. The pairwise distance (Pdist) between root and non-root histograms, which serves as a quantitative measure for separability of root and non-root image structures, indicates that soil-root images represent the most challenging modality for image segmentation.	5
1.5	Challenges in root image segmentation: (a) contrast between the root and background pixels varying along the image. The roots at the bottom of the image contain very low contrast compared to the upper part of the image. (b) aged root system architecture contains both primary and lateral roots in the image. However, primary roots are over-exposed compared to the lateral roots, which results in high inhomogeneity between them.	6
1.6	Growth of Arabidopsis plants in top view under different conditions over time: In the early stage (a), the plant has healthier green leaves. Over time, flowers appear in the shoot system (b), and leaves tend to change their colours (c) because of stress conditions like high temperature and drought.	7

1.7	Challenges in segmenting side view shoot images: (a) Due to the reflection of plant structures on the background, the upper region of the Arabidopsis plant exhibits very low contrast compared to the bottom areas (b) also, leaves and tillers exhibit inhomogeneous colours over the plant's ageing.	7
1.8	Visibility of spikes in Wheat plants: segmentation of spikes covered by leaves (c) makes it difficult to see them compared to non-overlapping spikes (b).	8
1.9	Overview of Barley awn barb structures: Due to the location and density of barbs on the awn surface, they exhibit various morphology (c) and contrast (d) compared to the standard barb structures (b).	9
2.1	Examples of root images from different setups	16
2.2	Histograms of root and non-root images	17
2.3	Basic steps of root image processing and analysis in saRIA	18
2.4	Workflow of image processing and analysis in saRIA	20
2.5	Accuracy of root image detection using saRIA vs. manual segmentation	21
2.6	Correlation between root traits calculated using semi-automated saRIA and manual SmartRoot image segmentation	23
2.7	Visualization of root features	24
2.8	Comparison of root image segmentation for young/small vs. large/adult plants	26
2.9	Graphical user interface of saRIA	28
3.1	The proposed U-net architecture for soil-root image segmentation	53
3.2	Exemplary root image from IPK plant phenotyping system	57
3.3	Workflow of the pipeline for image processing and segmentation in faRIA	59
3.4	Training and validation performance of the faRIA	61
3.5	Segmentation performance of CNN models	62
3.6	Segmentation artefacts	63
3.7	Dice coefficient of faRIA:256 over 40 barley soil-root images	64
3.8	Applicability of faRIA on rhizotron soil-root images	65
3.9	Dice coefficient of faRIA:256 over 30 Arabidopsis soil-root images	66
3.10	Applicability of faRIA on UV soil-root images	67
3.11	Correlation between root traits calculated using saRIA and faRIA image segmentation	68
3.12	Graphical user interface of faRIA	69

4.1	Examples of side- and top-view images of arabidopsis	91
4.2	Proposed U-net architecture for shoot image segmentation . . .	93
4.3	Workflow of the pipeline for image processing and segmenta- tion in the DeepShoot tool	97
4.4	Training and validation performance of the shoot model	100
4.5	Segmentation performance of shoot model	101
4.6	Segmentation performance on variational data sets	102
4.7	Evaluation of image segmentation on the references data set .	103
4.8	Examples of segmentation of arabidopsis top-view images from Scharr et al	103
4.9	Performance of neural network and proposed U-net model . . .	104
4.10	Evaluation of neural network segmentation with respect to the proposed U-net	105
4.11	Loss performance of alternative U-net models	106
4.12	Graphical User Interface of the DeepShoot tool	107
5.1	Limitations of wheat spike detection using Qiongyan et al. algorithm	122
5.2	Workflow of the proposed spike detection algorithm	124
5.3	The methodology of the proposed spike detection algorithm with example images	126
5.4	DWT Decomposition	127
5.5	Example of suppression of leaf crossings using the Frangi Filter	131
5.6	The HTP detection using Qiongyan et al. and our proposed method in comparison to the ground truth	133
5.7	Limitations of the proposed method	134
5.8	Comparison of the HTP detection using our and Qiongyan et al method	135
5.9	Spike growth analysis	137
5.10	Detected leaf artifacts	138
5.11	Morphological reconstruction of the spikes	138
6.1	Awn imaging system	156
6.2	Example of barley awn image	157
6.3	Proposed U-net architecture	158
6.4	Training performance of the BarbNet model	161
6.5	Validation performance of BarbNet vs U-net	162
6.6	Correlation between the ground truth and predicted barbs count	164
6.7	Clustering of phenotypic traits	166

Abbreviations

CNN Convolutional Neural Network

2D Two-Dimensional

GUI Graphical User Interface

IT Information Technology

RSA Root System Architecture

NMR Nuclear Magnetic Resonance

NIR Near-Infrared

μ **CT** Micro-Computed Tomography

3D Three-Dimensional

saRIA Semi-automated Root Image Analysis

LED Light Emitting Diode

UV Ultraviolet

IPK Institute of Plant Genetics and Crop Plant Research

Pdist Pairwise Distance

ROI Region of Interest

CF Conversion Factor

MATLAB Matrix Laboratory

DSC or DC Dice similarity coefficient

TP True Positive

TN True Negative

FP False Positive

FN False Negative

faRIA Fully-automated Root Image Analysis

GPU Graphics Processing Unit

RAM Random-Access Memory

ReLU Rectified Linear Unit

RGB Red, Green, and Blue

HSV Hue, Saturation, and Value

L*a*b CIELAB Colour Space

DeepShoot Deep Learning-based Shoot Image Analysis

VIS Visible Light

API Application Programming Interface

CPU Central Processing Unit

NN Neural Networks

CE Cross Entropy

CVPPP Computer Vision Problems in Plant Phenotyping

HTP Heading Time Point

BBCH Biologische Bundesanstalt, Bundessortenamt und Chemische Industrie

PPD-H1 PHOTOPERIOD-H1

QTL Quantitative Trait Locus

DAS Days After Sowing

DWT Discrete Wavelet Transform

CIS Colour Index-based Segmentation

1D One-Dimensional

PP Plant Pixels

RMSE Root Mean Square Error

GIMP GNU Image Manipulation Program

GWAS Genome-Wide Association Study

SEM Scanning Electron Microscopy

PCR Polymerase Chain Reaction

e!DAL electronic Data Archive Library

PC Personal Computer

Symbols

pH A measure of how acidic/basic water is

w/v Weight / Volume

T or tsh Threshold

A Approximation Coefficients

H Horizontal

V Vertical

D Diagonal

R^2 Coefficient of determination

$h_{11}, h_{12}, h_{21}, h_{22}$ Second-order partial derivatives of Hessian matrix

σ Variable scaling factor

λ_1, λ_2 Eigenvalues

c, β Constants that control the sensitivity of the filter

1. Introduction

1.1 Plant Phenotyping

Crop plants play a central role in ecological systems and represent an essential source of nutrients for 7.5 billion people [2]. However, the increasing world population and climate change pose new challenges for sustainable crop production [3]. To address these challenges, concentrated research efforts towards a deeper understanding of genetic and environmental factors affecting crop yield are required [4, 5]. Crop yield is a critical and multifaceted parameter in agriculture, and it defines the amount of crop grown, such as Wheat and Barley, per unit of area land. The higher the crop yield, productivity and profitability, the higher.

An essential tool for understanding, predicting, and improving crop yield is plant phenotyping. The comprehensive definition of plant phenotyping involves the assessment of plant traits such as growth, development, architecture, physiology, tolerance, resistance, and fundamental measurement of specific quantitative parameters such as leaf area, root length and plant height [6, 5]. The relationship between crop yield and plant phenotyping is intricate, involving the systematic measurement and analysis of various plant traits, encompassing morphological, developmental, and functional properties to enhance crop productivity [7].

Plant phenotyping is considered a basis for any plant breeding selection process, and it requires knowledge integrated from multidisciplinary sciences (covering biological sciences, computer science, mathematics, and engineering). Especially in the case of precision agriculture, plant phenotyping emerged as a powerful tool for examining how phenotypes interact locally and dynamically with the temporally and spatially dynamic environment above (shoot) and below ground (root) as shown in Figure 1.1. Diverse plant structural and functional elements, such as plant biomass [8], root morphology [9, 10], leaf properties [11], and fruit traits [12], are directly quantified in the phenotype. By studying the language of plant traits, researchers

and practitioners can exploit this knowledge to breed resilient crops, enhance agricultural sustainability, and address the challenges of changing climate.

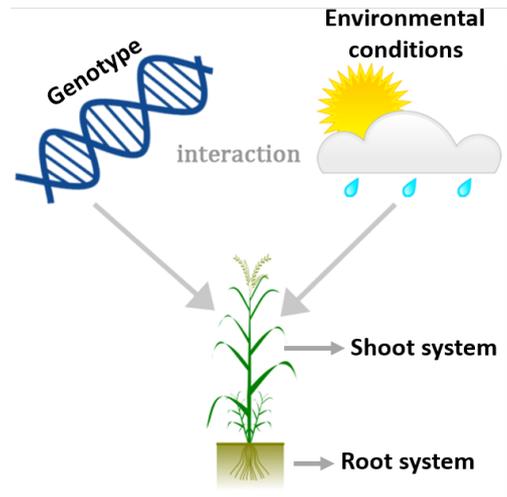


Figure 1.1: Aim of plant phenotyping: Analyze the interaction between genotype and environmental conditions that promote a particular phenotype [1].

The methods for plant phenotyping can be categorized as destructive and non-destructive, depending on whether the plant is harmed or altered during measurement. The destructive phenotyping alters the plant tissue while measuring or collecting plant traits manually, such as height and width, and renders the plant unsuitable for further analysis [13, 14]. These methods are not only destructive but also labour-intensive and time-consuming, increasing the plant population size required in phenotyping studies [15]. On the other hand, non-destructive phenotyping methods enable the repeated measurement of plant traits without causing harm to the plant. Computer-assisted methods for phenotyping, particularly imaging techniques, are becoming more popular [16, 17], enabling the analysis of a large number of plants with less human effort in an automated and high-throughput manner. This technique benefits longitudinal research and high throughput plant phenotyping breeding programs. Further sections will discuss an overview and challenges of various plant image analysis solutions for high-throughput phenotyping.

1.2 State-of-the-Art and Research Requirements

Plant image analysis has emerged as a pivotal tool in various scientific disciplines, revolutionizing how researchers study, monitor, and manage plant life. This multidisciplinary field integrates advanced imaging technologies, image processing algorithms, and data analytics to extract meaningful information from visual data. The importance of plant image analysis is underscored by its applications in precision agriculture [18], genetic research [15], environmental monitoring [19], and automated farming systems [20]. Automated analysis of extensive image data from high-throughput screening studies, such as greenhouse and microscopic measurements, is pivotal for quantitatively evaluating phenotypic plant traits. However, efficient and accurate analysis of extensive image data is a non-trivial task that represents one of the major bottlenecks of quantitative plant biology [21].

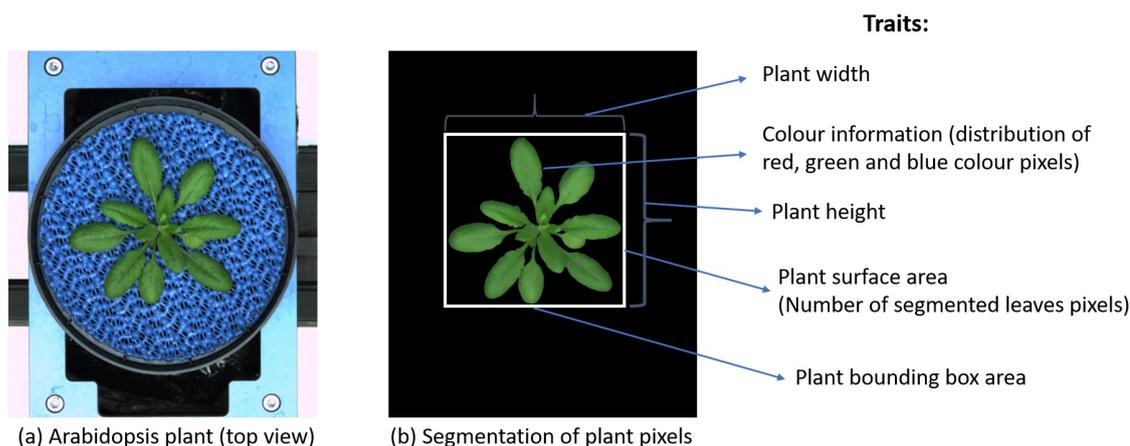


Figure 1.2: Image-based plant phenotyping: (a) Exemplary Arabidopsis shoot system captured in the greenhouse (b) segmented target plant pixels from background pixels. The morphology of the segmented shoot system, called traits, is calculated to study genotype-phenotype characterization.

The primary tasks by analysis of plant image data include pixel-wise identification (segmentation) and/or localization (detection) of whole plants or plant parts/organs, followed by a quantitative description of segmented/detected image regions in terms of morphological, developmental and, functional traits (e.g., diverse shape-, colour-, growth- and health-traits) as shown in Figure 1.2. However, a straightforward segmentation and detection of plant struc-

tures are often hampered by several natural and technical factors, including the complexity of optical scenes (e.g., shadows, reflections), dynamic appearance of growing organisms, presence of statistical and structural noise, differences in the appearance of plant structures in different image modalities. Thereby, it is essential to consider that the reliability of phenotypic traits depends on the accuracy of the first critical step of image processing - the identification of targeted image regions, i.e., pixel-wise image segmentation and object detection.

The thesis focuses on developing image segmentation solutions for major plant organs, such as roots, shoots, and tiny objects in grainy crops (Barley spikes, barbs in Barley awns), for non-destructive plant phenotyping approaches.

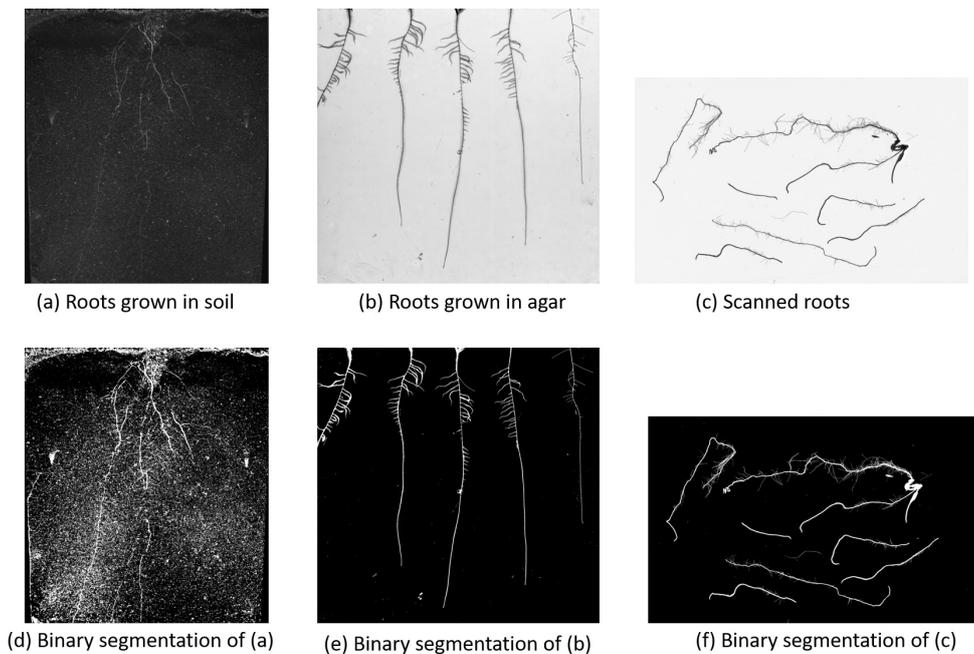


Figure 1.3: Segmentation of conventional thresholding techniques on root images from different setups: (a) Barley plant roots in the soil environment (b) Rapeseed roots grown in transparent containers filled with agar gel (c) scanned washed Maize roots on an Epson Expression 10 000 XL scanner. Segmentation of roots from the soil background in the image (d) is very noisy compared to the other setups in (e) and (f).

To date, most of the root image analysis approaches WinRHIZO [22],

EZ-RHIZO [23], SmartRoot[24] and, IJ_Rhizo [25] have focused on segmenting roots from well-contrasting background media such as transparent liquids (agar) and manually scanned roots as shown in Figure 1.3b and 1.3c respectively. However, optically contrasting substrates do not represent a natural root environment. In recent imaging systems [26, 27], roots are grown and imaged in their natural opaque soil environment (soil-root image in Figure 1.3a). However, conventional thresholding techniques [28] are insufficient to accurately differentiate between root and non-root pixels of these soil-root images compared to the high contrast background mediums as shown in Figure 1.3 d, e, f. Moreover, the distribution of root and non-root pixels in these three imaging modalities was analysed using pairwise distances ($Pdist = pdist2(A, B)$), where A and B are root and non-root pixel vectors, respectively. The histogram study in Figure 1.4 briefs that the soil-root image exhibits a substantial overlap between grayscale values of root and non-root pixels compared to other non-soil imaging modalities. Therefore, it is evident that both root and non-root pixels in the soil-root images exhibit similar or even the same intensity. This results in significantly higher background noise pixels in the segmented soil-root image (Figure 1.3b).

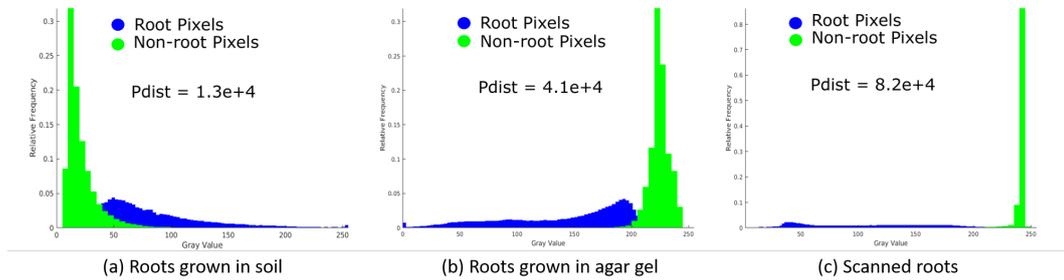


Figure 1.4: Histograms of root (blue) and non-root (green) intensity values of different imaging modalities shown in Figure 1.3. The pairwise distance (Pdist) between root and non-root histograms, which serves as a quantitative measure for separability of root and non-root image structures, indicates that soil-root images represent the most challenging modality for image segmentation.

Furthermore, the intensity among the root structures varies rapidly across the soil-root image of different plant types and time points of root growth. Figure 1.5a shows the sparse root system architecture (RSA) in the soil environment where contrast among the root and background soil pixels is reduced from top to bottom of the image because of varying soil distributions. Furthermore, the aged plants tend to have high inhomogeneity between the

primary (or thin) and lateral (or fine) roots, as shown in Figure 1.5b. These addressed limitations make the segmentation of roots under soil background a challenging task.

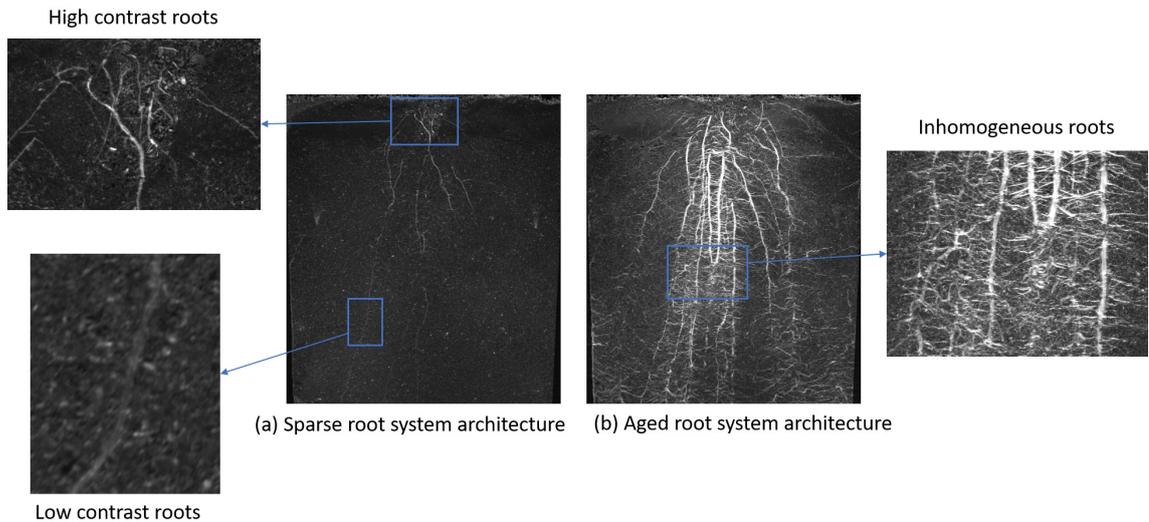


Figure 1.5: Challenges in root image segmentation: (a) contrast between the root and background pixels varying along the image. The roots at the bottom of the image contain very low contrast compared to the upper part of the image. (b) aged root system architecture contains both primary and lateral roots in the image. However, primary roots are over-exposed compared to the lateral roots, which results in high inhomogeneity between them.

Along with the root system, the shoot system also undergoes significant changes in appearance during plant development. This study focuses on segmenting plant shoot images from Arabidopsis, Barley, and Maize, in which the optical appearance of shoot systems dynamically varies a lot in real time. Figure 1.6 displays the growth of Arabidopsis plants under different conditions over time. Initially, the shoots are light green in colour (Figure 1.6a), but as they mature and the leaves grow larger, they become darker and may have white-coloured flowers (1.6b). Eventually, due to stress and/or ageing, the colours of the shoots change completely to brown, yellow, and red, as shown in Figure 1.6c. Likewise, the side view shoot images of Arabidopsis and Barley in Figure 1.7 exhibit varying contrast shoot structures due to the reflection of plant structures on the background walls of greenhouse photo chambers and inhomogeneous colours of leaves and tilters over the plant ageing. Because of this complexity, state-of-the-art solutions based on reference

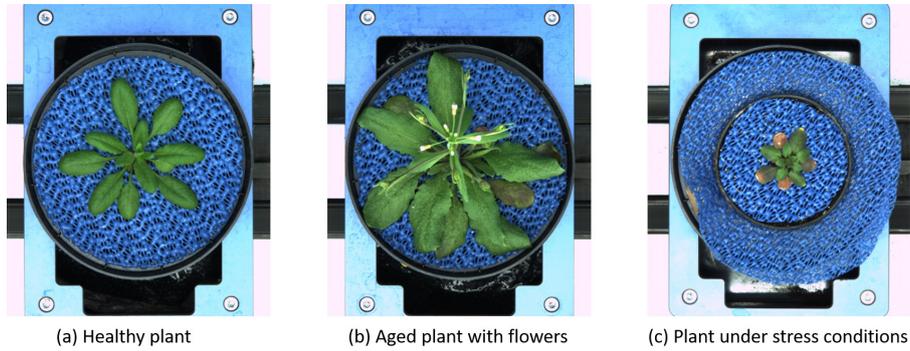


Figure 1.6: Growth of Arabidopsis plants in top view under different conditions over time: In the early stage (a), the plant has healthier green leaves. Over time, flowers appear in the shoot system (b), and leaves tend to change their colours (c) because of stress conditions like high temperature and drought.

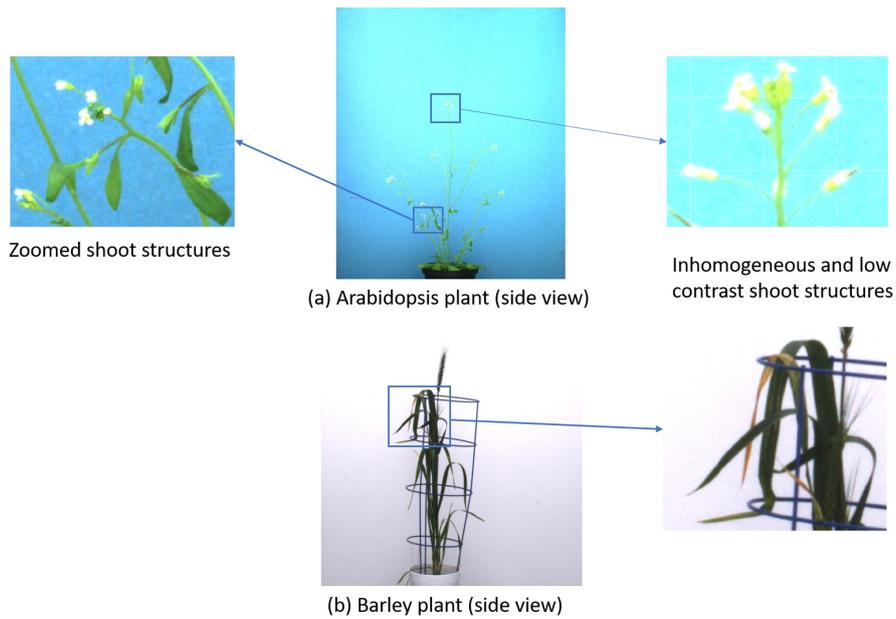


Figure 1.7: Challenges in segmenting side view shoot images: (a) Due to the reflection of plant structures on the background, the upper region of the Arabidopsis plant exhibits very low contrast compared to the bottom areas (b) also, leaves and tillers exhibit inhomogeneous colours over the plant's ageing.

frames [29, 30] and pixel-based supervised learning methods [31, 32] are gen-

erally not sufficiently robust for unsupervised analysis of sizeable variable image data. Segmenting plant root and shoot systems poses a challenging task. Therefore, more advanced methods with higher abstraction capabilities are required to cope with the tasks of noisy image segmentation.

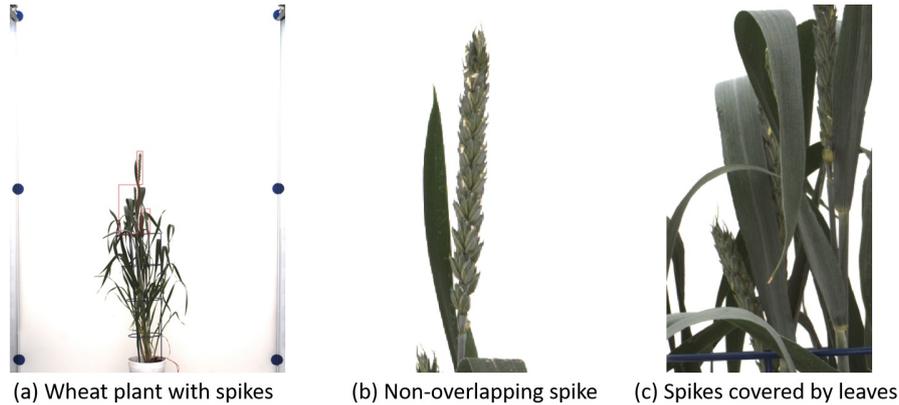


Figure 1.8: Visibility of spikes in Wheat plants: segmentation of spikes covered by leaves (c) makes it difficult to see them compared to non-overlapping spikes (b).

Phenotyping plant organs in grain crops is crucial for improving crop quality and ensuring food security for the growing population [33]. In addition to studying plant roots and shoot systems, researchers need to study spikes in Wheat and Barley plants from 2D images, as shown in Figure 1.8a. Till date, many state-of-the-art solutions are based on single spikes and require cutting off spikes from plants to classify different Wheat varieties using morphological image processing algorithms, Hu moments, and neural networks [34, 35, 36]. For the first time, Qiongyan et al. [37] proposed a novel approach for detecting (young) spikes in digital images of Wheat plants based on Law’s textural (energy) features and a neural network. However, the process of phenotyping spikes is challenging due to the variational spike texture and visibility of spikes in the plants. Leaves partially or entirely cover spikes; some leaves cross over spikes, making it difficult to segment them from the plants. One can see this example in Figure 1.8c, where the leaves obscure the spikes, compared to Figure 1.8b, where the spike is fully visible.

Another novel application in the phenotyping plant organs is the detection and phenotyping of barbs in microscopic Barley awn images to study awn roughness, as illustrated in Figure 1.9a. Awns, which are bristle-like extensions of the glumes or husks of various grass species, including significant

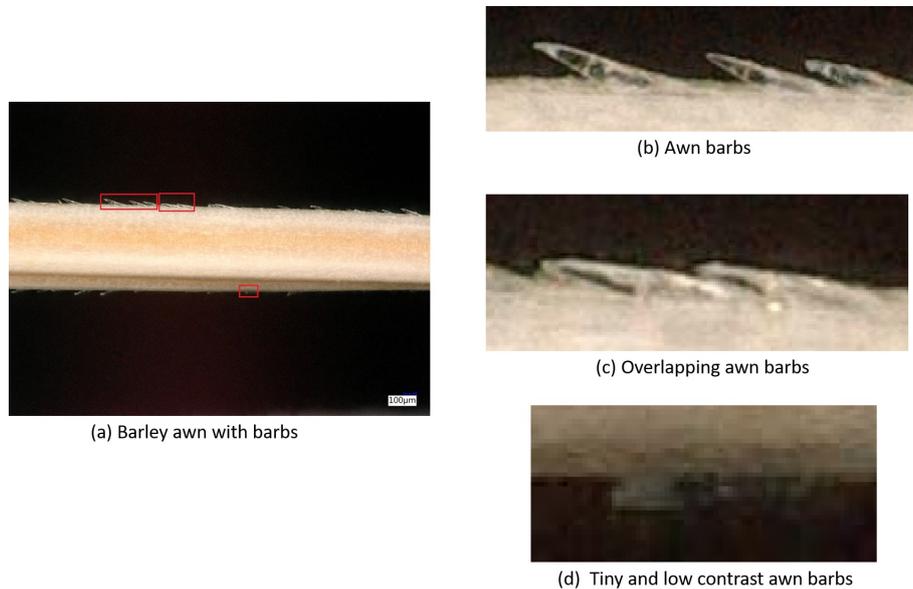


Figure 1.9: Overview of Barley awn barb structures: Due to the location and density of barbs on the awn surface, they exhibit various morphology (c) and contrast (d) compared to the standard barb structures (b).

crop plants like Wheat and Barley, are the focus of this research. The presence of barbs gives awns a rough texture that can be harmful during manual harvest or lowers the quality of Barley as a feed crop [38]. Awn barbs generally are upward-oriented single-cell structures, see Figure 1.9b. However, accurate segmentation of barbs is a critical and challenging task because of its limited spatial resolution, low contrast, and considerable variability, as shown in Figure 1.9c and d. To our knowledge, no appropriate state of the art image processing solutions are available for accurate barb segmentation and phenotyping in awn images.

Although image modalities and applications in plant phenotyping may differ substantially, common challenges are observed across a broad spectrum of root-and-shoot image analysis. They include, for example, (i) low contrast between fore- and background image regions, (ii) dynamic appearance of plant structures in the course of plant development (over time), and, as a result, absence of trivial invariant features that could be universally used for segmentation, detection and classification of plant from non-plant structures. Therefore, more advanced computer vision methods with a higher level of abstraction capabilities such as deep learning, are required to overcome these challenges.

1.3 Convolutional Neural Network (CNN)

Technologies related to computer vision are growing in popularity. They are utilized in a wide range of systems and applications, including data analysis systems in medicine, face, fingerprint, and iris recognition for personal identification tasks, security video surveillance systems that use license plate recognition to identify cars, and many more [39, 40, 41]. The development of large image bases and increased processing power have made it feasible to train deep neural networks to handle computer vision tasks like recognition and classification. According to research (2020) [42], the deep learning industry's global market value will increase to \$60.5 billion by 2025 from \$12.3 billion in 2020, representing a 37.5% growth rate. Among industrial sectors, agriculture is one of the slowest to adopt deep learning but has a high potential for its use to combat global food insecurity.

Convolutional Neural Networks (CNNs) are a deep learning model designed for processing structured grid data, such as images and videos [43, 41, 39]. They are neural networks with convolution layers that have been shown to outperform traditional methods associated with higher-level cognitive abilities in many computer vision applications [44, 45]. Typically, convolutional neural networks include a fully connected layer and a sub-sampling layer, also known as a pooling layer. Convolutional neural networks are employed in various tasks such as semantic segmentation, object detection, pattern recognition, and image classification [43, 46]. Several "levels" of neurons, referred to as feature maps or channels, make up convolution layers and sub-samplings in convolutional neural networks. A tiny area of the preceding layer, known as the receptive field, is connected to every neuron in this layer. A feature map in the context of an image is a two-dimensional neural array or just a matrix. Simultaneously, despite the practical application of such networks, the issue of selecting the best architecture and configuring the neural network's parameters continues to be open. In this view, the work aims to develop the most efficient convolutional neural networks and practical recommendations on choosing the best architecture and setting the neural network's parameters according to the problem's specifics.

In recent years, an emerging trend in quantitative plant sciences has been using CNNs for automated image analysis and plant phenotyping such as the detection of Wheat roots grown in germination paper [47], segmentation of roots from the soil in X-ray tomography [48] and segmentation of spikes in Wheat plants [49]. However, because of the wide variation in optical plant appearance, variations in experimental setups [50, 48], and lack of

labelled ground truth data [51, 52], dependable software tools appropriate for a specific plant species are rarely available. As a result, very few software tools are currently available for high-throughput plant image analysis and phenotyping.

1.4 Research Objectives

The thesis aims to develop efficient and automatic plant image segmentation algorithms and phenotyping tools for a broad spectrum of applications including root-and-shoot image analysis, spike detection in Wheat plants and barbs in microscopic images of Barley awn cultivars. In brief, this thesis focuses on the development of image processing and deep learning solutions for automatic high-throughput plant phenotyping. The objectives of the thesis are:

- Development of a semi-automated root image segmentation algorithm and phenotyping tool using traditional image processing algorithms and evaluating the tool’s performance with state-of-the-art tools. This tool can generate ground truth data for subsequent training of deep learning models. (👉 Chapter 2)
- Research on CNN models for the segmentation of roots from soil background and compare the performance of the proposed model with the state-of-the-art CNN models for root image analysis. (👉 Chapter 3)
- Investigation of shallow learning-based neural networks and different deep learning models for the segmentation of different plant shoot images (Arabidopsis, Barley and Maize) at different developmental stages (e.g., juvenile, adult) in different views (e.g., top and side views) and evaluate their performance with the proposed CNN model. (👉 Chapter 4)
- Enhancement of wheat spike detection algorithm in 2D images using shallow learning-based neural networks and studying the robustness of the model on European cultivars. (👉 Chapter 5)
- Developing a CNN model for the segmentation of barbs in microscopic images of Barley awn cultivars and studying the robustness of the proposed model by genotype-phenotypic characterization and classification of awns. (👉 Chapter 6)
- Developing a GUI-based front end for efficiently handling the above algorithmic frameworks is also suitable for IT-unskilled users.

2. Semi-automated Root Image Analysis (saRIA)

Paper 1: Narisetti, N., Henke, M., Seiler, C. et al. Semi-automated Root Image Analysis (saRIA). *Sci Rep* 9, 19674 (2019). <https://doi.org/10.1038/s41598-019-55876-3>

Impact Factor: 4.38 (according to SCImago Journal Rank, January 2024)

Abstract

Quantitative characterization of root system architecture and its development is important for the assessment of a complete plant phenotype. To enable high-throughput phenotyping of plant roots efficient solutions for automated image analysis are required. Since plants naturally grow in an opaque soil environment, automated analysis of optically heterogeneous and noisy soil-root images represents a challenging task. Here, we present a user-friendly GUI-based tool for semi-automated analysis of soil-root images which allows to perform an efficient image segmentation using a combination of adaptive thresholding and morphological filtering and to derive various quantitative descriptors of the root system architecture including total length, local width, projection area, volume, spatial distribution and orientation. The results of our semi-automated root image segmentation are in good conformity with the reference ground-truth data (mean dice coefficient = 0.82) compared to IJ_Rhizo and GiAroots. Root biomass values calculated with our tool within a few seconds show a high correlation (Pearson coefficient = 0.8) with the results obtained using conventional, pure manual segmentation approaches. Equipped with a number of adjustable parameters and optional correction tools our software is capable of significantly accelerating quantitative analysis and phenotyping of soil-, agar- and washed root images.

2.1 Introduction

Plant roots are key drivers of plant development and growth. They absorb the water and inorganic nutrients from the soil [1, 2, 3] and provide anchoring of the plant body [4, 5]. Root system architecture (RSA), the spatial configuration of a root system [1] is known to be an important phenotypic feature closely related to crop yield variability upon changes in environmental conditions [6, 7]. In general, the RSA and its response to the environment are known to be dependent on multiple factors including the plant species, the plant genotype, composition of the soil, availability of nutrients and the environmental conditions [8]. The emerging discipline of plant phenomics aims to extract the plant anatomical and physiological properties to study the plant performance under given conditions [1]. In the case of roots, the relevant traits include descriptors of global and local root morphology (like total length, are, volume, and diameter, or lateral branching, the direction of a tangent, etc.) [9, 10, 11, 12]. Monitoring of these traits enables conclusions about the ability of plants response to variable environmental factors such as drought, cold, starvation, etc., [13].

In recent years, a number of approaches to root imaging and image analysis were suggested [14]. However, most of these works rely on the measurement of washed roots or roots grown in artificial, optically transparent media such as liquids or gels [15, 16] that allow a straightforward image analysis. Further non-destructive methods including X-ray computed tomography [17, 18, 19, 20], nuclear magnetic resonance (NMR) microscopy [21], magnetic resonance imaging [22, 23] and laser scanning [24] provide unique insights into 3D organization of living root architecture, however, their throughput capabilities are presently rather limited. To enable 2D imaging of roots in a soil-like environment, near-infrared (NIR) imaging of roots growing along surfaces of transparent pots or minirhizotron was designed and tested [25, 26]. Special long pass filters were used to block root exposure to visible light and the images were taken by NIR sensitive camera with suitable illumination. The system allows a non-invasive acquisition of root images in darkness [26].

To analyze a large number of root images in an automated high-throughput manner, a number of software tools are available. Most of these tools were, however, designed to extract RSA traits from specific imaging systems, e.g., images from minirhizotron [25] and images of roots grown in agar [27]. In addition, some general tools are available for in-depth analysis of monocot root systems regardless of root structure. These tools depends on significant user input for processing even though they can be used in a batch mode

[28, 29, 30, 31]. Moreover, some tools can be used as a plugin for general image processing platforms like ImageJ to perform specific tasks for the manual segmentation of roots in the image [32].

The majority of software for root image analysis is rather tailored to artificial setups such as transparent growing media that cannot be applied to the analysis of heterogeneous and noisy soil-root images. With exception of software for analysis of X-ray micro-computed tomography (μ CT) images [33] and Root1 [34], which still requires extensive human-computer interaction and suitable for X-ray tomography 3D images. In recently published works [35, 36, 37], novel machine and deep learning approaches to automated segmentation of soil-grown root image were presented. However, the presented approaches rely on colour information and require substantial amount of ground truth training data as well as substantial computational resources.

In this work, we present a GUI-based handy tool for semi-automated root image analysis (saRIA) which enables rapid segmentation of diverse 2D root images including potting soil and artificial media setups in a high-through manner. Based on a combination of adaptive image enhancement, adjustable thresholding and filtering as well as optional manual correction, saRIA represents a broadly applicable tool for quantitative analysis of diverse root image modalities as well as generation of quality ground truth reference images for training of advanced machine learning/deep learning algorithms.

The paper is structured as follows: Materials and Methods section describes the methodological framework of saRIA including data preparation, segmentation algorithm, and root trait computation. Results section shows the segmentation capabilities of saRIA with respect to freely available tools and presents the results of roots traits derivation from manually and saRIA segmented root images. In Discussion, we summarize the results of an evaluation study using the saRIA root image segmentation and give an outlook of possible future improvements.

2.2 Methods

2.2.1 Image Data

Three different modalities for imaging of root system architecture were analyzed in this study including

- **Soil-root Image:** This type of digital image is taken by a monochrome

camera (UI-5490SE-M-GL, IDS) with LED illumination (UV, 380nm) in a custom-made imaging box similar to our previously published setup [26]. In brief, plants are grown in transparent pots [77x77x97mm (WxLxH)] filled with potting substrate (Potgrond P, Klasmann). An example of soil-root images acquired with this system is shown in Figure 2.1a. Depending on the developmental stage, plant health, environmental factors (e.g., temperature, humidity), these images may, in general, exhibit diverse artifacts including low contrast between the root architecture and heterogeneous soil, inhomogeneous scene illumination (i.e. vertical intensity gradient), water condensation at the pot walls, see Figure 2.1b. Identification of relevant root architecture in such structurally and statistically noisy images represents a challenging task.

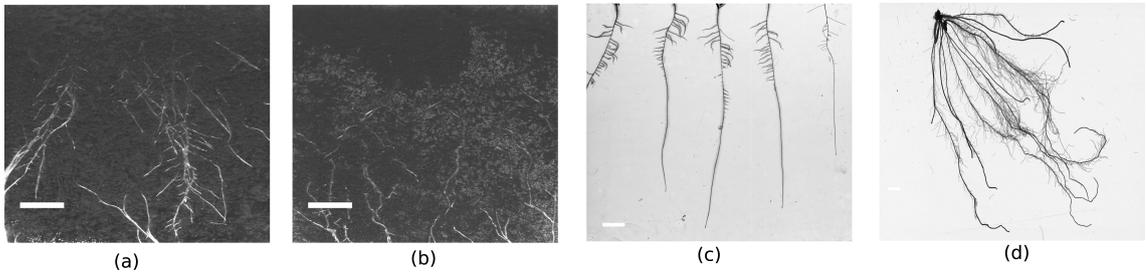


Figure 2.1: Examples of root images from different setups: (a) Arabidopsis plant roots 28 days after sowing, (b) Arabidopsis plant roots with condensation noise 28 days after sowing, (c) roots grown in agar, (d) washed roots. The white colour bar on each image represents the scale of 1cm in length.

- **Agar-root Image:** During this experiment, the plants were grown on 1/2 MS, 1.5% (w/v) agar medium (pH 5.6 without sugar) in Petri dishes for 5 days. The images were captured by scanning the dishes in grayscale at 300 dots per inch resolution using an Epson Expression 10 000 XL scanner (Seiko Epson) [26]. An exemplary image of Rapeseed roots is shown in Figure 2.1c. This image has a clear contrast between roots and homogeneous background. Nevertheless, the background pixels have some morphological artifacts which will be discussed in the next subsection.
- **Washed-roots Image:** Maize plants were grown in transparent pots filled with mixture of substrate (self-made compost, IPK) and sand (1:1) for 3 weeks. This digital image is obtained by scanning the washed

maize roots on an Epson Expression 10 000 XL scanner (Seiko Epson) as shown in Figure 2.1d. Compared to the above two types of root images, it is less noisy and the contrast between roots and background is significantly higher.

Consequently, the classification between root and non-root pixels in these three image modalities is unequally difficult. Figure 2.2 shows histograms and pairwise distances ($Pdist = pdist2(A, B)$) between root and non-root grayscale distributions corresponding to soil-root, agar and washed roots images in Figure 2.1, respectively. As expected, soil-root images exhibit a strong overlap between grayscale values of root and non-root pixels and the distance between them is the lowest among these three imaging modalities. Therefore, here we focused on soil-root image analysis only. Application of our approach to higher contrast image modalities (i.e. agar, washed roots images) is, however, trivial, and only requires inverting the grayscale values.

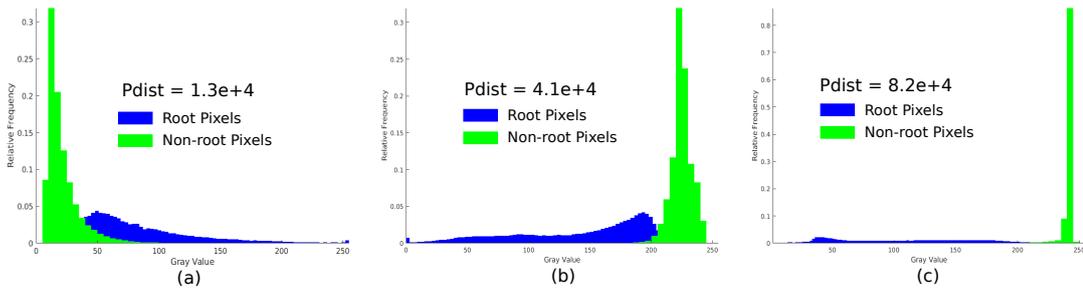


Figure 2.2: Histograms of root (blue) and non-root (green) intensity values of different imaging modalities shown in Figure 2.1: (a) soil-root, (b) agar grown roots and (c) washed roots images. The pairwise distance ($Pdist$) between root and non-root histograms, which serves as a quantitative measure for separability of root and non-root image structures, indicates that soil-root images represent the most challenging modality for image segmentation.

2.2.2 Image Analysis

Image analysis algorithms, as well as the graphic user interface (GUI) were implemented under the MATLAB 2018b environment [38]. The major goal of image analysis consists of segmentation of root architecture and calculation of phenotypic features of root architecture and image intensity (i.e. colour). In case of colour images, the input image is converted to a grayscale image using *rgb2gray* Matlab routine. In general, the pipeline of image analysis includes the following steps:

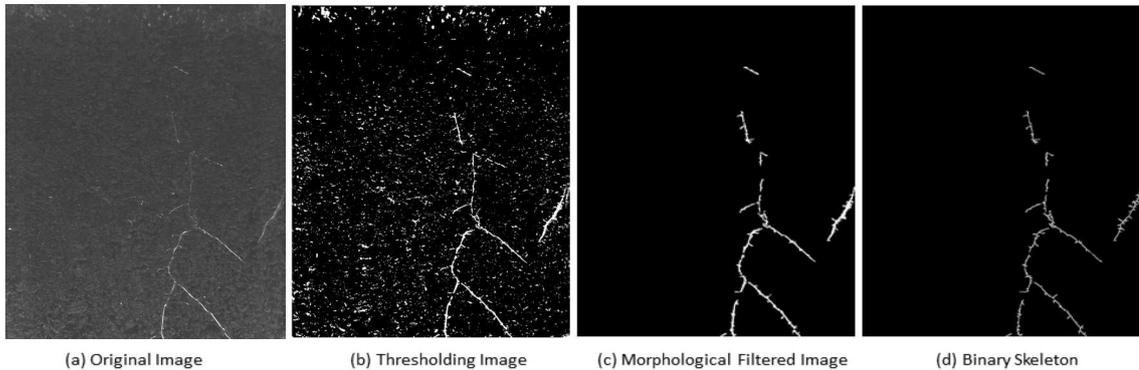


Figure 2.3: Basic steps of root image processing and analysis in saRIA: (a) Arabidopsis soil-root image, (b) adaptive thresholding, (c) morphological filtering, (d) root skeleton.

- **Image I/O:** Most standard image formats (such as *.jpg, *.png, *.bmp, *.tif) can be imported for further processing and analysis. Stepwise single image, as well as automated processing of large image datasets is implemented.
- **Image preprocessing:** Depending on the imaging modality (e.g., soil-root or agar-root images) and presence of noisy or structural artifacts, preprocessing steps may include cropping of the region of interest (ROI), inverting of image intensity, despeckling and smoothing. In case of agar and washed-root images, inversion of image intensity has to be performed prior to image analysis. Otherwise, the procedure of agar and washed-root image analysis is similar to soil-root images.
- **Adaptive image thresholding:** Preprocessed images are segmented into a foreground (roots) and background using adaptive thresholding based on Gaussian weighted mean as suggested by [39]. This technique tolerates global inhomogeneity of image intensity such as vertical image gradient in our soil-root images. An example of an adaptive thresholding of an Arabidopsis soil-root image (Figure 2.3a) is shown in Figure 2.3b.
- **Morphological filtering:** To remove white noise and small non-root blob-like structures (such as sand, gravel or water condensation arti-

fact in Figure 2.1b) morphological filtering is applied. Thereby, roots are considered to be elongated line-/curve-like structures that differ from this kind of non-root blobs with respect to their area, length, and shape (i.e. eccentricity). In the case the root region represents a single connected structure, filtering can be performed merely by applying intensity and area thresholds. If roots are represented by disconnected structures, differentiation of root from non-root structures is performed using additional shape descriptors such as length and eccentricity, i.e. a descriptor of the object's Eigenellipse elongation, which is zero for an absolutely round and 1 for a line-object. By appropriate setting of thresholds for these three parameters, non-root blobby structures are removed. Figure 2.3c shows an example of a morphological filtering of a preliminary segmented root image.

- **Skeletonization:** Root skeleton is calculated on the basis of the segmented and filtered image. In addition to the filtering steps described above, additional thinning or eroding of the binary image is applied to suppress high-frequency noise. The exemplary image for the skeleton extraction is shown in Figure 2.3d.
- **Root feature calculation:** The distance transform of the cleaned binary image is calculated for assessment of the local root width (or diameter) measured in pixels of the root skeleton. Further root features include root length, root angles with respect to a vertical axis, branching and end points of the roots skeleton, the intensity of root pixels and their standard statistical descriptors (i.e. mean, stdev values). The complete list of a total of 44 root traits can be found in Supplementary Information (Table S1). Note that all traits are extracted using pixel-wise calculation irrespective of number of root systems in the image. In addition, the RSA traits can also be written out in mm. For this purpose, the user have to set the pixel-to-mm conversion factor in the saRIA GUI. Here, we derived the pixel-to-mm factor by measuring a reference line (white colour bar) in the images as shown in Figure 1. The pixel-to-mm conversion factor (CF) is then defined as follows:

$$\mathbf{CF} = \frac{\text{length of the reference line in mm}}{\text{length of the reference line in pixels}} \quad (2.1)$$

The workflow of image analysis is also shown in Figure 2.4

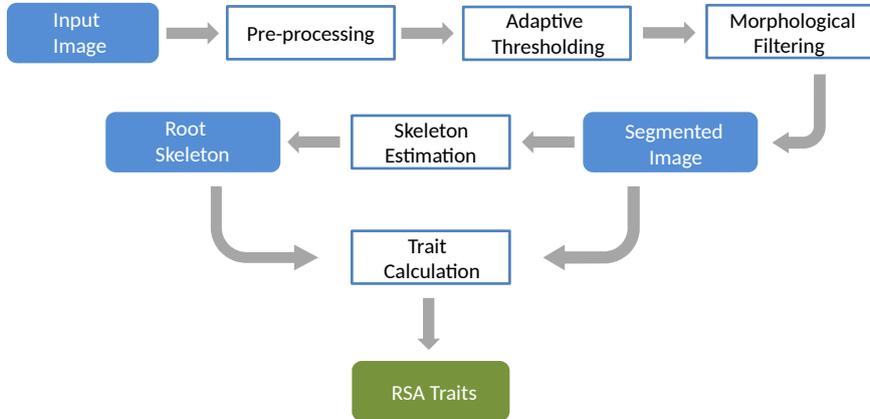


Figure 2.4: Workflow of image processing and analysis in saRIA. Colour filled icons indicate the data modalities, framed rectangles describe image processing steps.

- **Evaluation:** To examine the performance of image segmentation, a standard statistical metric, the Dice similarity coefficient (DSC) [40], is used. The DSC evaluates the spatial overlap between two binary images and its value ranges between 0 (no overlap) to 1 (perfect overlap). The DSC is defined as follows:

$$DSC = \frac{2 * TP}{2 * TP + FP + FN}, \quad (2.2)$$

Where TP, FP, and FN are true positive, false positive and false negative pixels, respectively.

2.3 Results

2.3.1 Semi-automated Segmentation of Soil-Root Images

To evaluate the performance of our algorithms, segmentation of 100 *Arabidopsis* soil-root images was performed automatically and compared with the results of fully manual segmentation. Thereby, manual segmentation was also carried out with saRIA by applying a low intensity threshold for

selection of all high- as well as low-intensity roots, and subsequently followed by manual removal of all remaining artifacts including solitary objects as well as noise regions attached to the root (which cannot be otherwise identified and quantified as a separate object) using the ‘clearInside’ saRIA tool. This step was done by two biologists (co-authors of this manuscript) with an expertise in RSA. In contrast, for semi-automated image segmentation user merely has to adjust the four basic algorithmic parameters (controlled by the four GUI sliders) according to the subjectively best result of visual inspection of a few test images. Once the best combination of algorithmic parameters is defined, segmentation of all remaining images can be performed in a fully unsupervised manner.

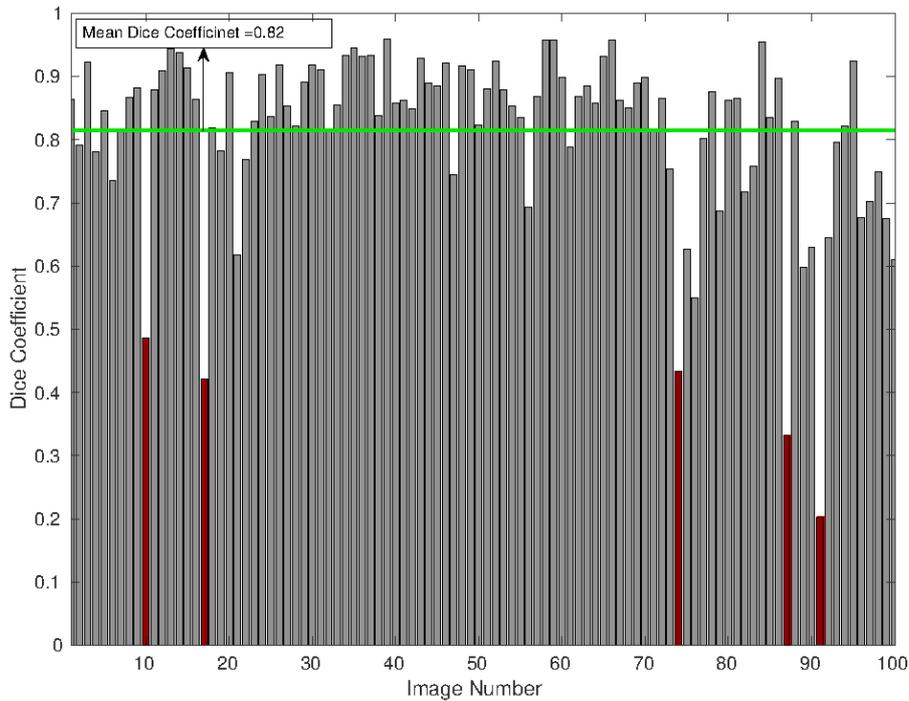


Figure 2.5: The accuracy of root image detection using saRIA vs. manual segmentation (ground-truth data) in terms of the dice similarity coefficient (DSC). The green line points to the mean DSC value over 100 tested soil-root images. The red colour bars indicate a few cases of poor saRIA performance with a low DSC value that corresponds to roots in the early stage of plant development.

Figure 2.5 shows the performance of image segmentation compared to manual segmentation of 100 soil-plant images for the intensity threshold (T) 0.12, minimum area 450, minimum length 46 and minimum eccentricity 0.49. This figure shows that approximately 90% images have DSC value greater than 0.7 and the mean DSC value is 0.82.

To validate the robustness of the saRIA image segmentation of the above 100 images, we have also compared with two other freely available tools called IJ_Rhizo [41] and GiA roots [42]. Table 2.1 shows the mean DSC value for the subjectively best possible configuration of IJ_Rhizo, GiA roots with three different thresholding methods and saRIA. The table briefs that saRIA significantly outperformed with the combination of Gaussian adaptive thresholding and all morphological parameters (area, length, circularity) compared to the IJ_Rhizo and GiA roots. The brief discussion on parameter configuration of IJ_Rhizo and GiA roots can be found at [41] and [42] respectively.

Parameter	IJ_Rhizo	GiAroots- Global threshold	GiAroots- Adaptive thresh- old	GiAroots- Double adaptive threshold	saRIA
Mean DSC	0.43	0.50	0.54	0.69	0.82
Threshold	T1=50, T2=255	T=50	Mean shift=- 2.0096	Bound drop value=5	T=0.12
Minimum area	5 mm	450	450	450	450
Minimum length	x	x	x	x	46
Circularity	0.7 (0-line, 1-circle)	x	x	x	0.49 (0-circle, 1-line)

Table 2.1: Comparison of saRIA image segmentation quality (i.e. mean DSC) and parameters with IJ_Rhizo and GiA roots.

2.3.2 Evaluation of Phenotypic Traits vs. SmartRoot

Here, the results of phenotypic root characterization obtained with saRIA are evaluated in comparison with SmartRoot [32]. The SmartRoot is the most widely used for the traits quantification of disconnected RSA and each part of the root is traced manually by placing multiple landmarks finally interconnected to the root skeleton. However, the SmartRoot doesn't deliver the single segmented binary (reference) image for the comparison with saRIA. Therefore, the root traits derived from such manually segmented images can be seen as a reference ground-truth data. To quantify the (dis)similarity between saRIA and SmartRoot results, the correlation coefficient of determination R^2 and significance level p-value are used. They represent the percent of the saRIA calculated traits that is closest to the ground-truth data and model validation respectively. Figure 2.6 shows the correlation between the SmartRoot (x-axis) and saRIA (y-axis) outputs for three traits where each point denotes one particular image out of 126 Arabidopsis root images acquired with our in-house soil-root imaging system. Note that the images used for traits evaluation are different from above segmentation evaluation data. The three traits used for evaluation are the total root length, total root surface area, and total root volume. As one can see, for all three traits correlation between saRIA semi-automatically and SmartRoot manually segmented images exhibit R^2 values higher than 0.84, 0.86, 0.77 and p-values $7.7e-53$, $5.25e-55$, $3.35e-42$ respectively.

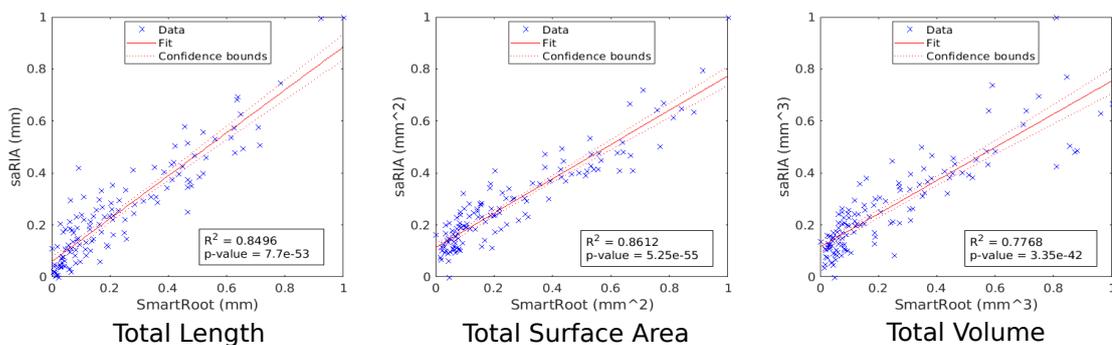


Figure 2.6: Correlation between root traits calculated using semi-automated saRIA (y-axis) and manual SmartRoot (x-axis) image segmentation. Each point represents a trait value estimated from one of 126 soil-root images. The red colour solid line and dotted lines represent a fitted curve and 95% confidence bounds, respectively. The R^2 value indicates good conformity between saRIA and SmartRoot results of image segmentation and trait calculation.

2.3.3 Visualization of Root Features

In addition to numerical outputs, saRIA software generates root features (e.g., distance maps, skeletons, width distributions, etc) for visualization purposes. Figure 2.7 shows an example of images of root width and orientation. The root width is calculated as the Euclidean distance between the root skeleton and root boundary pixels. Figure 2.7b depicts the Euclidean distance map of the root object Figure 2.7a where high-intensity central pixels represent the root width. The corresponding width colour map is shown in Figure 2.7c. The width feature is useful to calculate the root volume and surface area for biomass estimation.

Figure 2.7d displays the absolute orientation of each root skeleton pixel with respect to the horizontal axis in black-yellow colour-map representation. Here, a validated linear regression model was used to calculate the slope of a pixel in the skeleton image. The local slope in i -th pixel is obtained by fitting a tangent line to the fraction of root skeleton framed by a 15x15 pixel mask around the i -th pixel. An exemplary figure for the local linear fit can be found in Supplementary Information (Figure S1). The validated regression model means that only pixels satisfying the regression model with a high confidence level (i.e. p -value < 0.05 and $R^2 > 0.5$) were accepted. Pixels with a low confidence of the linear fit model such as root branches with non-linear distribution of skeleton pixels were excluded by the calculation of global statistics of root orientation.

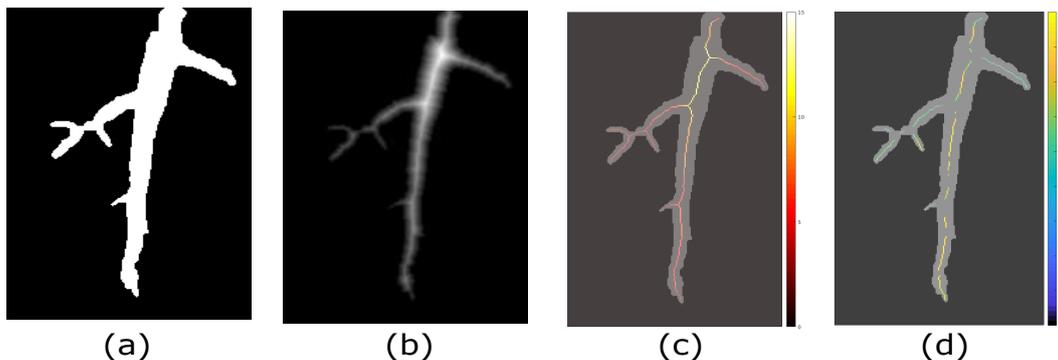


Figure 2.7: Visualization of root features: (a) binary root image, (b) corresponding Euclidean distance map, (c) overlay of the binary image with the root skeleton coloured according to the local root width (gray-scale map), (d) overlay of the binary image with the root skeleton coloured according to the local root tangent (black-yellow colour map).

2.4 Discussion & Conclusion

The objectives of our GUI-based saRIA tool are to automatize the time-consuming manual segmentation of structurally complex and noisy root images and to enable calculation of RSA traits from different root imaging modalities including soil, agar and washed roots images. Using this approach, root architecture can be rapidly segmented and quantified by adjusting a small set of algorithmic parameters. Segmentation with saRIA is particularly efficient when background structures differ from roots in geometrical parameters (such as shape and size) and grayscale intensity. Artifacts resembling optical root appearance, e.g., scratches on the pot surface or high-intensity areas resulting from water condensation, are, in contrast, difficult to eliminate in a fully automated manner. Such artifacts can, however, be removed using manual segmentation tool also available with saRIA.

The accuracy of trait estimation in saRIA depends on the quality of semi-automated image segmentation. Our solution for analyzing a large number of images is to define the best possible set of algorithmic parameters for a subset of representative root images and then to apply this configuration to all remaining images in a fully automated manner. Here, 15% of input data with different scenarios (low, medium and high dense root images) are considered for the best possible configuration settings.

The quality of image segmentation from Table 2.1 explains that the global thresholding methods in IJ_Rhizo (bi-level threshold) and GiAroots (single-level threshold) under-performed than adaptive thresholding methods. Since the global thresholding methods contain one or two threshold values for a complete image that preserve the high-intensity noisy objects and removes the low-intensity roots in the soil-root image. However, the GiAroots also implemented based on adaptive thresholding but it lacks the Gaussian smoothing filter in the preprocessing step and morphological constraints (i.e. length and circularity) on binary root objects, represented as x in Table 2.1. Therefore, the combination of adaptive thresholding and morphological filtering promising more accurate segmentation in saRIA for soil-root images.

From the summary of our automated image analysis in Figure 2.5, it is evident that accuracy of root image segmentation is tendentially higher ($DSC > 0.9$) for images with large root architecture. Figure 2.8 shows example images of small and large root architecture that exhibit low and high DSC of automated segmentation vs. ground-truth, respectively. Because the large root architecture requires a low value for threshold and high value

for morphological parameters compared to the small root architecture image where most of the disconnected root components are small in morphology. It results in the removal of small sized roots and keeping the (both disconnected root and heterogeneous soil) structures which are big in morphology. This observation confirms that the relative error in segmentation of small root architecture (see red colour bars in Figure 2.5) from background pixels is higher for automated segmentation in comparison to the large roots. However, these artifacts can be overcome by setting low values for morphological parameters in the segmentation configuration.

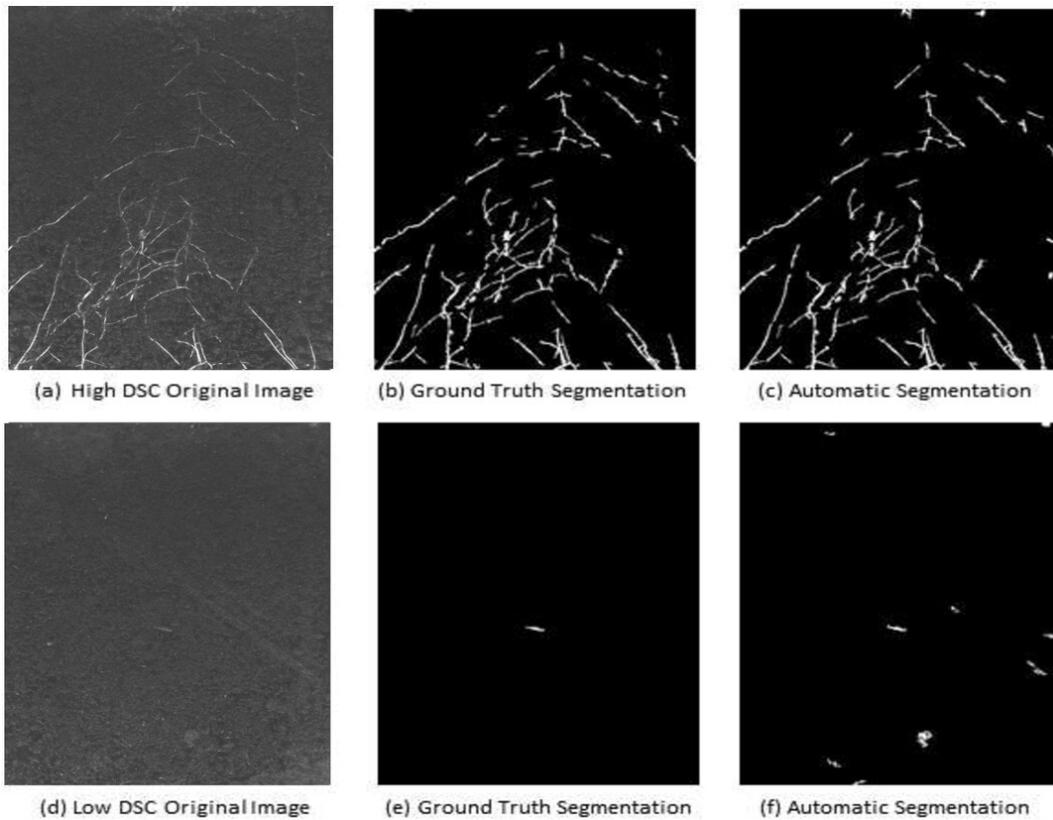


Figure 2.8: Comparison of root image segmentation for young/small vs. large/adult plants. Top row shows (a) original, (b) ground-truth and high-DSC saRIA-segmented images of a large root, bottom row (d-f) shows an example of a relatively small root at the early stage of Arabidopsis plant development with a low DSC corresponding to the red bar number 91 in Figure 2.5.

The quantitative comparison of saRIA is limited to SmartRoot because other software solutions for root image analysis are either tailored to non-interrupted representation of root architecture like RootNav [43] or restricted to high contrast imaging modalities, e.g., agar grown or washed roots, like GiARoots [42] that is no longer under development and closed source and not promising more accurate segmentation as shown in Table 2.1.

As mentioned earlier, saRIA is capable of calculating 44 number of root traits. In brief, they are categorized into 11 features named area (number of root pixels), number of disconnected root objects, total length, surface area, volume, number of branching and ending points, statistical distribution (mean, median, standard deviation, skewness, kurtosis and percentile) of root geometry in horizontal and vertical direction, width and orientation. Among those, three important features for root biomass calculation are presented for saRIA traits evaluation.

The results of our evaluation tests in Figure 2.6 show that root traits obtained using saRIA are highly correlating ($R^2 > 0.8$) and significant (p -value < 0.05) with manual segmentation in SmartRoot software. Consequently, one can perform comparatively high-quality root phenotyping with saRIA 20 or more times faster than with manual annotation of root structures pixel by pixel.

The difference in trait estimates between saRIA and SmartRoot might result from the image segmentation parameters and root thickness. First, the morphological parameters remove tiny roots which have small area and length. Second, the local thickness of roots in saRIA is defined as the average diameter automatically segmented roots which includes high-intensity structures originating from tiny root hairs that can be avoided in manual segmentation. This may lead to differences in average root width, length and volume assessed with saRIA vs. SmartRoot. However, the ability to interactively adjust parameters is available in the saRIA to improve the trait extraction and even to produce a set of alternative image segmentation in an automated manner, i.e. automated root tracing and trait extraction for selected configuration.

Here, we present a user-friendly GUI-based software solution for high-throughput analysis of root images of different image modalities, including challenging soil-root images. Figure 2.9 shows the GUI of saRIA software which is freely available as a precompiled executable program from <https://ag-ba.ipk-gatersleben.de/saria.html>. Further examples of agar and

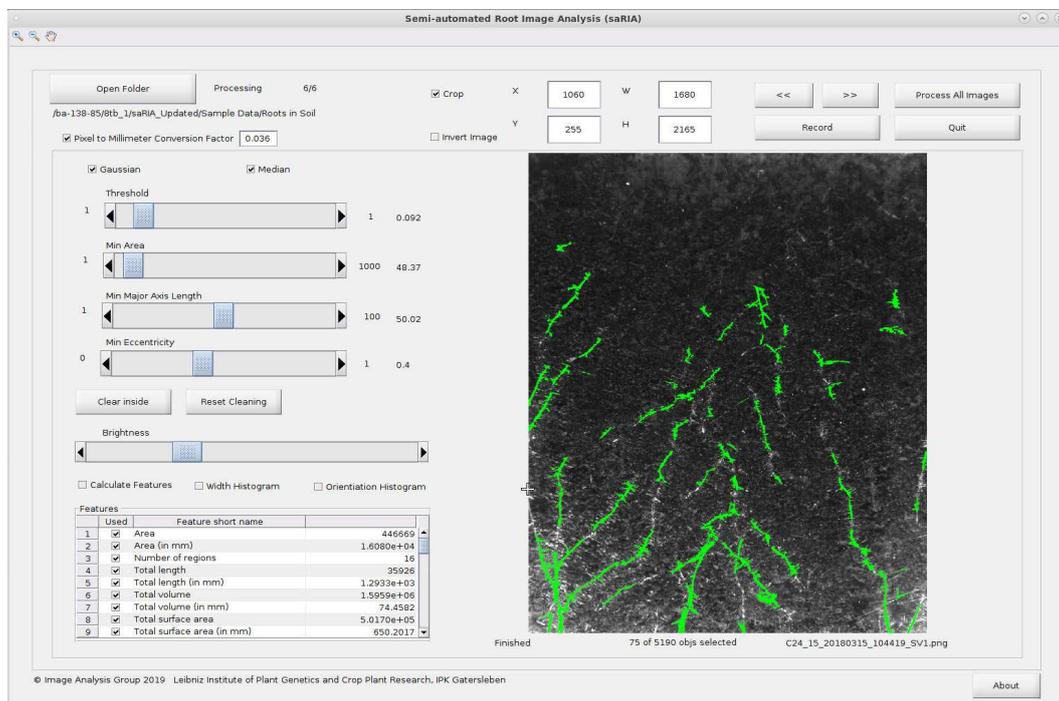


Figure 2.9: The graphical user interface of saRIA. Green coloured pixels indicate segmented root regions of the soil-root image.

washed roots image segmentation using saRIA are included in the Supplementary Information, see Figure S2 and Figure S3. The saRIA software can be applied for analysis of single images or large image datasets to automatically detect and extract multiple root traits. This software is designed for end-users with limited technical knowledge to enable them widely automated analysis of complex soil-root images in an intuitive and transparent manner. The saRIA segmentation, root tracing and trait calculation algorithms require, in average, 5 seconds to process and analyze a 6-megapixel (cropped) image (on Intel(R) Core(TM) i5-4570S CPU @ 2.90GHz, 64-bit quadcore processor with 4GB ram and 500GB HDD) which is significantly faster in comparison to conventional manual segmentation, e.g., SmartRoot. Table 2.2 summarizes the essential features of saRIA vs. other well-established root image analysis tools (SmartRoot, EZ-RHIZO, and WinRHIZO). The major difference between the saRIA and other available tools (Plant Root - roots grown in cloth substrate in custom rhizoboxes, RootReader2D - need high contrast images and RootNav - supports nested root architectures) is that it is capable of segmenting contrast varying disconnected root architectures (in

semi-automated manner with automated trait extraction) in both potting soil and artificial growing media.

In addition to routine analysis of root images, saRIA can be used for rapid generation of ground-truth segmentation data that are highly demanded for advanced machine learning/deep learning techniques.

The study of plant genotype with root phenotype requires a contribution of many groups and utilization of molecular, physiological and imaging techniques. In addition, the performance of phenotype analysis depends on the image quality. The segmentation algorithm currently bundled in saRIA is based on intensity gradient among the pixels. Further extensions of the saRIA segmentation pipeline including advanced machine learning approaches and additional static and dynamic RSA traits like topological data (number of primary and lateral roots, branching angles, lateral density) are planned in the future.

Feature	saRIA	SmartRoot	EZ-RHIZO	WinRHIZO
License	Free,Closed source	Free,Open source	Free,Closed source	Commercial
Platform	Linux,Windows	Cross-Platform, it is an ImageJ Plugin	Windows only	Windows only
Language	MATLAB	Java	C++	x
Root Tracing	Semi-automated	Manual	Manual and Automated	Automated
Medium	Soil,Agar,Washed	Soil,Agar,Washed	Agar	Washed roots,Agar
Database Support	No,but has CSV export support	SQL	SQL	No,data files are saved in ASCII text format

Table 2.2: The feature comparison of saRIA with other software for root segmentation and trait extraction

Acknowledgements

This work was performed within the German Plant-Phenotyping Network (DPPN) which is funded by the German Federal Ministry of Education and Research (BMBF) (project identification number: 031A053).

Author contributions statement

NN, MH, and EG conceived, designed and performed the computational experiments, analyzed the data, wrote the paper, prepared figures, and tables, and reviewed drafts of the paper. CS, RS, and AJ executed the laboratory experiments, acquired image data, and reviewed drafts of the paper. TA co-conceptualized the project and reviewed drafts of the paper.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplementary Material

Supplementary Material

Semi-automated Root Image Analysis (saRIA), N. Narisetti, M. Henke, C. Seiler, R. Shi, A. Junker, T. Altmann, E. Gladilin*.

Table S1. Root Traits: The description of estimated root system architecture traits in saRIA software.

Trait Name	Description
Area	Number of root pixels in the image
Number of Regions	Number of disconnected root objects in the image
Total Length	The sum of major axis length of each root object approximated by fitting ellipse to the root object
Total Volume (V)	<p>The sum of local volume at each root object of skeleton approximated by tubular shape whose average radius is estimated from image</p> $V = \sum_{i=0}^n \pi r_i^2$ <p>Where r_i is the average radius of i^{th} root component in the image.</p>
Total Surface Area (SA)	<p>The sum of surface area at each root object of skeleton approximated by tubular shape whose average radius is estimated from image</p> $SA = \sum_{i=0}^n 2\pi r_i$ <p>Where r_i is the average radius of i^{th} root component in the image.</p>
Specific Root Length	The ratio of total length and total volume of roots in the image.
Number of Branching Points	The total number of branches in the root skeleton
Number of End Points	The total number of end points in the root skeleton
Geometrical X_mean, Y_mean	The mean value of root pixels distribution in horizontal and vertical direction
Geometrical X_median, Y_median	The median value of root pixels distribution in

Geometrical X_std, Y_std	horizontal and vertical direction The standard deviation of root pixels distribution in horizontal and vertical direction
Geometrical X_skew, Y_skew	The skewness of root pixels distribution in horizontal and vertical direction
Geometrical X_kurt, Y_kurt	The kurtosis of root pixels distribution in horizontal and vertical direction
Geometrical X_p25, Y_p25	The 25 percentile of root pixels distribution in horizontal and vertical direction
Geometrical X_p50, Y_p50	The 50 percentile of root pixels distribution in horizontal and vertical direction
Geometrical X_p75, Y_p75	The 75 percentile of root pixels distribution in horizontal and vertical direction
Geometrical X_p99, Y_99	The 99 percentile of root pixels distribution in horizontal and vertical direction
Width mean	Average root diameter
Width median	Median root diameter
Width std	Standard deviation of the root diameter
Width skew	Skewness of root diameter
Width kurt	Kurtosis of root diameter
Width p25	25 percentile of root diameter
Width p50	50 percentile of root diameter
Width p75	75 percentile of root diameter
Width p99	99 percentile of root diameter
Orientation mean	Average root Orientation
Orientation median	Median root Orientation
Orientation std	Standard deviation of the root Orientation
Orientation skew	Skewness of root Orientation
Orientation kurt	Kurtosis of root Orientation
Orientation p25	25 percentile of root Orientation
Orientation p50	50 percentile of root Orientation
Orientation p75	75 percentile of root Orientation
Orientation p99	99 percentile of root Orientation

Table S2. Data set: Data used for traits comparison of saRIA (automatic) Vs Smart root (manual) (See attached excel spread sheet to the supplementary information)

Figure S1: Linear regression fit: An exemplary analysis of linear regression fit with 10 neighbor pixels. Where β_i represents the measured angle between the central pixel (green color blob) and a fitted tangent (gray dotted line).

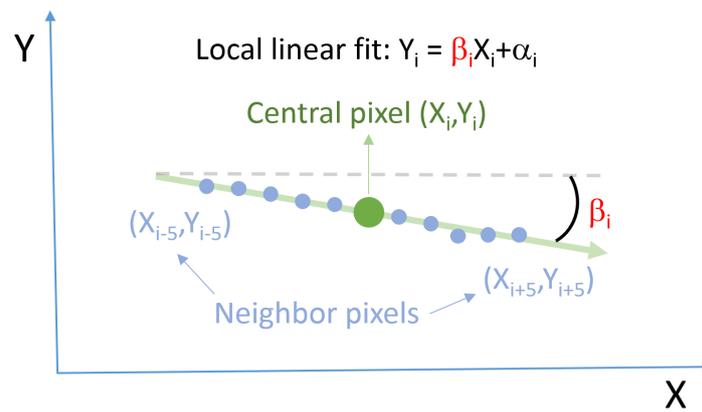


Figure S2: saRIA software: The graphical user interface of saRIA software. Green colour pixels represents the detected roots in the Agar root image.

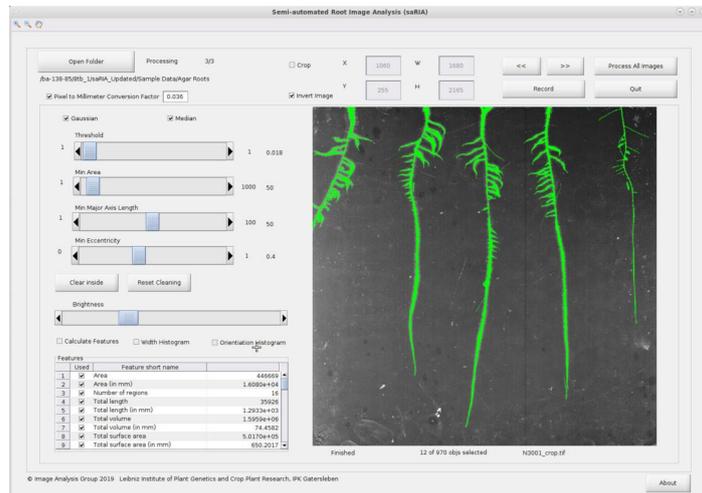
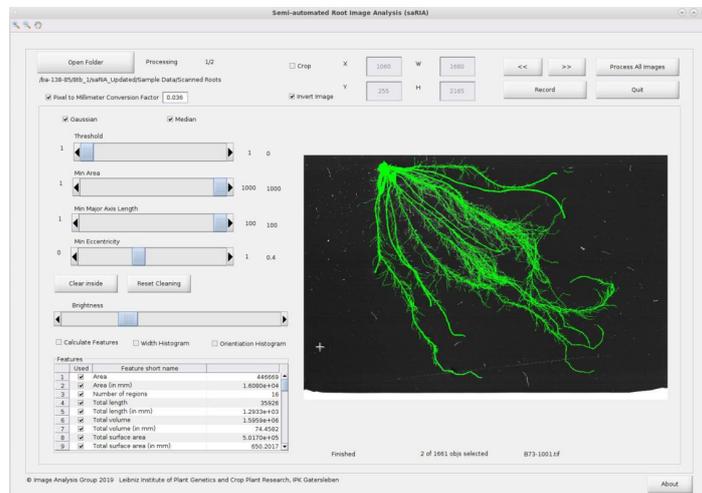


Figure S3: saRIA software: The graphical user interface of saRIA software. Green colour pixels represents the detected roots in the scanned image.



filename	S_Length	S_Surface Area	S_Volume
1826CS006_20180716_100851_SV1	13.39279459	0.53167794	0.00232
1826CS010_20180716_102121_SV1	6.267686365	0.234390449	0.0009
1826CS010_20180716_102127_SV2	7.31497732	0.333042349	0.00162
1826CS014_20180716_102154_SV1	7.972931642	0.304594734	0.00128
1826CS018_20180716_103407_SV1	2.74631221	0.126297701	0.00062
1826CS018_20180716_103413_SV2	0.67884886	0.04944127	3.35E-04
1826CS018_20180716_103418_SV3	1.390074	0.0455144	1.65E-04
1826CS026_20180716_104641_SV1	8.70849059	0.279934674	0.00103
1826CS026_20180716_104647_SV2	0.38969532	0.003146374	1.01E-05
1826CS026_20180716_104656_SV4	2.57417825	0.139765049	0.0007
1826CS030_20180716_104718_SV1	7.784439161	0.284274516	0.00102
1826CS030_20180716_104734_SV4	1.517378	0.118905365	8.02E-04
1826CS034_20180716_104817_SV1	13.16748644	0.638003488	0.00363
1826CS034_20180716_104832_SV2	1.6691192	0.11515787	7.07E-04
1826CS034_20180716_104838_SV3	0.6796204	0.042654432	3.01E-04
1826CS038_20180716_104748_SV1	7.84851047	0.301444444	0.00117
1826CS038_20180716_104802_SV4	2.0990834	0.073185452	0.00027
1826CS042_20180716_103600_SV1	14.53306011	0.574594043	0.0027
1826CS042_20180716_103616_SV4	6.86951359	0.401201668	0.00234
1826CS046_20180716_103529_SV1	3.358901674	0.109038349	0.00038
1826CS046_20180716_103535_SV2	1.0557042	0.05176327	2.28E-04
1826CS058_20180716_100959_SV1	3.220126376	0.09509028	0.00034
1826CS062_20180716_100930_SV1	23.8768452	0.949239339	0.00512
1826CS062_20180716_100937_SV2	3.043874345	0.112148224	0.0005
1826CS066_20180716_101039_SV1	22.7305876	1.051377958	0.00502
1826CS066_20180716_101054_SV4	7.87975261	0.470038265	0.00264
1826CS070_20180716_101110_SV1	5.057206138	0.142136617	0.00054
1826CS070_20180716_101118_SV2	2.57829285	0.162372079	0.00098
1826CS082_20180716_103641_SV1	5.101411038	0.173935792	0.00065
1826CS086_20180716_103708_SV1	17.03917727	0.762231872	0.00379
1826CS086_20180716_103723_SV4	2.01279638	0.10032657	0.00047
1826CS094_20180716_104902_SV1	20.37718286	0.920677838	0.00455
1826CS094_20180716_104907_SV2	4.3607907	0.232455083	0.00132
1826CS098_20180716_105035_SV1	2.828279756	0.075585475	0.00026
1826CS098_20180716_105040_SV2	0.89516711	0.032090169	0.00013
1826CS110_20180716_103736_SV1	15.57383647	0.689428159	0.00305
1826CS114_20180716_102451_SV1	6.26634825	0.214931312	0.00084
1826CS114_20180716_102456_SV2	2.887417166	0.109780067	0.00054
1826CS118_20180716_102421_SV1	19.01537712	0.785252441	0.00327
1826CS122_20180716_101216_SV1	17.06688279	0.788688548	0.00386
1826CS122_20180716_101231_SV4	2.43710809	0.089987174	0.00038
1826CS130_20180716_101246_SV1	17.31861458	0.84067454	0.00485
1826CS130_20180716_101304_SV4	3.61329585	0.173477852	0.00084
1826CS138_20180716_102521_SV1	15.32852511	0.661995578	0.00313
1826CS138_20180716_102527_SV2	7.446346938	0.324695612	0.00177
1826CS138_20180716_102537_SV4	1.38550834	0.053719423	0.00022
1826CS142_20180716_102550_SV1	9.264297598	0.254494767	0.00078
1826CS146_20180716_103831_SV1	33.44295993	1.35432328	0.0072
1826CS146_20180716_103837_SV2	5.13579315	0.252332952	0.0016

1826CS146_20180716_103846_SV4	0.8452682	0.06345203	4.47E-04
1826CS150_20180716_103859_SV1	7.15229178	0.362872573	0.00231
1826CS154_20180716_105003_SV1	4.354021538	0.120289815	0.00037
1826CS166_20180716_105135_SV1	9.466691596	0.37372527	0.00175
1826CS170_20180716_103953_SV1	17.03467918	0.999458132	0.0069
1826CS170_20180716_103958_SV2	5.50082966	0.301744475	0.0016
1826CS174_20180716_103925_SV1	10.37631676	0.341182868	0.00129
1826CS174_20180716_103931_SV2	0.75021	0.03504683	1.66E-04
1826CS178_20180716_102648_SV1	36.19724357	1.482164306	0.00689
1826CS178_20180716_102654_SV2	2.73286984	0.145649495	0.00081
1826CS178_20180716_102704_SV4	3.42089765	0.127407248	0.00053
1826CS182_20180716_102619_SV1	6.26376322	0.294465847	0.00144
1826CS182_20180716_102625_SV2	1.57884188	0.064890727	0.00028
1826CS182_20180716_102635_SV4	4.60554176	0.289414507	0.00216
1826CS186_20180716_101416_SV1	19.00869469	0.728924433	0.00325
1826CS186_20180716_101434_SV4	2.597215513	0.120453879	0.00072
1826CS198_20180716_102034_SV1	7.9601377	0.203017254	0.00065
1826CS210_20180716_104526_SV1	25.88972136	0.95082515	0.00393
1826CS210_20180716_104532_SV2	1.27647059	0.066820473	0.00038
1826CS214_20180716_104553_SV1	9.53293044	0.420915774	0.00207
1826CS218_20180716_105704_SV1	25.77887426	1.122147571	0.00602
1826CS218_20180716_105710_SV2	4.52079126	0.17697617	0.00083
1826CS218_20180716_105718_SV4	1.1229599	0.0711277	4.19E-04
1826CS222_20180716_105733_SV1	6.549464019	0.464155242	0.00371
1826CS230_20180716_105628_SV1	15.16112447	0.582580714	0.00298
1826CS230_20180716_105633_SV2	1.0156255	0.036544386	1.50E-04
1826CS234_20180716_104457_SV1	17.91106481	0.826931707	0.0049
1826CS238_20180716_104429_SV1	3.765152973	0.131268293	0.00065
1826CS242_20180716_103213_SV1	18.47610775	0.825607714	0.00434
1826CS242_20180716_103219_SV2	3.53229104	0.127857258	0.00049
1826CS246_20180716_103143_SV1	23.14317442	1.157591842	0.00639
1826CS246_20180716_103149_SV2	1.43493165	0.051693453	0.00022
1826CS246_20180716_103153_SV3	1.6451689	0.09455021	5.37E-04
1826CS250_20180716_101925_SV1	14.22037917	0.641288169	0.00296
1826CS250_20180716_101940_SV4	2.6434867	0.12775119	5.46E-04
1826CS258_20180716_101756_SV1	23.57651551	1.308475125	0.00852
1826CS258_20180716_101801_SV2	0.6057517	0.029093144	1.41E-04
1826CS262_20180716_101827_SV1	11.60694036	0.493776319	0.00263
1826CS262_20180716_101833_SV2	2.00111963	0.062224443	0.00021
1826CS262_20180716_101842_SV4	12.58472005	0.931494449	0.00743
1826CS270_20180716_103115_SV1	15.10240575	0.556026199	0.0021
1826CS274_20180716_104326_SV1	7.688917736	0.228431554	0.00101
1826CS282_20180716_105535_SV1	23.1953477	0.985072815	0.00493
1826CS282_20180716_105540_SV2	2.31383472	0.124622712	0.00068
1826CS282_20180716_105549_SV4	2.9632268	0.155271646	0.00079
1826CS286_20180716_105602_SV1	17.56742844	0.802130099	0.00415
1826CS290_20180716_105505_SV1	2.54118072	0.090698015	0.00039
1826CS294_20180716_105431_SV1	3.370595211	0.109282365	0.00051
1826CS298_20180716_104258_SV1	12.99082947	0.557608471	0.00264
1826CS302_20180716_104231_SV1	19.33981111	0.79114425	0.00389

1826CS302_20180716_104245_SV4	1.8158042	0.08572725	3.61E-04
1826CS306_20180716_103011_SV1	6.467614329	0.216276528	0.00108
1826CS306_20180716_103016_SV2	1.8106617	0.066894162	0.00027
1826CS314_20180716_101728_SV1	10.46746289	0.392852297	0.00156
1826CS314_20180716_101733_SV2	2.02140939	0.101212548	0.00065
1826CS314_20180716_101742_SV4	5.464604443	0.360325045	0.0023
1826CS322_20180716_101600_SV1	16.43773322	0.989236839	0.00749
1826CS326_20180716_101629_SV1	16.00884094	0.767354254	0.00444
1826CS326_20180716_101635_SV2	4.611843625	0.159553371	0.00065
1826CS326_20180716_101644_SV4	1.85975586	0.093800177	0.00049
1826CS338_20180716_104135_SV1	22.4123341	1.138751828	0.00729
1826CS338_20180716_104149_SV4	4.07600268	0.212386828	0.00111
1826CS342_20180716_104204_SV1	9.35797239	0.456265026	0.00212
1826CS342_20180716_104209_SV2	1.1405361	0.07450396	4.41E-04
1826CS346_20180716_105339_SV1	22.90230309	1.206162774	0.00819
1826CS346_20180716_105344_SV2	5.30397777	0.327155979	0.00192
1826CS346_20180716_105352_SV4	3.4841715	0.237562559	0.00161
1826CS350_20180716_105405_SV1	10.60125155	0.450270101	0.00263
1826CS350_20180716_105410_SV2	3.65058567	0.142798551	0.00064
1826CS354_20180716_105307_SV1	17.09156029	0.633285754	0.00289
1826CS354_20180716_105322_SV4	6.537711025	0.281383522	0.00136
1826CS358_20180716_105236_SV1	8.566186341	0.278752198	0.00102
1826CS358_20180716_105243_SV2	5.00246579	0.237303798	0.00119
1826CS362_20180716_104104_SV1	28.49074313	1.24523764	0.00593
1826CS370_20180716_102810_SV1	5.854270073	0.193435719	0.00071
1826CS378_20180716_101530_SV1	3.894538921	0.100338835	0.00033
1826CS382_20180716_101453_SV1	16.70517549	0.648931896	0.00304

sa_Length	sa_Surface Area	sa_Volume
13.968	0.82568	0.0051249
5.3676	0.27489	0.0016131
7.5888	0.42599	0.0024661
12.0672	0.54914	0.0031709
3.8952	0.23954	0.0014817
4.6908	0.2188	0.0010833
4.5612	0.26802	0.001749
14.2524	0.55078	0.0030586
7.2828	0.3765	0.0021808
6.8004	0.30194	0.0018239
12.9348	0.60832	0.0032981
4.6368	0.24061	0.001713
14.778	0.89086	0.0059265
6.1308	0.34598	0.0020476
3.906	0.29166	0.0018816
16.488	0.942	0.0052866
10.1232	0.48152	0.0024948
16.0992	0.69722	0.0039555
11.8836	0.69951	0.0049528
7.1208	0.43123	0.0026771
7.5672	0.38927	0.0020809
6.2532	0.31837	0.0019848
21.7656	1.1945	0.0070213
9.5472	0.38856	0.0024611
24.75	1.6582	0.010531
11.5056	0.53351	0.0033212
10.6128	0.45769	0.0023966
10.2312	0.47567	0.0027216
11.2284	0.65469	0.0037751
16.1244	0.90123	0.0055968
5.4828	0.30251	0.0017741
23.1156	1.1028	0.0065449
14.7564	0.63826	0.0039029
7.3224	0.33836	0.0018317
3.2832	0.15576	0.00080264
22.2012	0.91573	0.0047928
14.4072	0.58245	0.003468
9.9468	0.53079	0.0029342
24.1596	1.2287	0.0067831
17.3124	0.98726	0.0058309
7.83	0.38129	0.0024219
16.9416	1.0353	0.0070054
8.802	0.54451	0.003885
19.7316	0.85074	0.0053969
11.97	0.56975	0.0037809
4.86	0.23859	0.0013446
9.6228	0.49328	0.0025518
40.5828	1.8292	0.010968
7.0164	0.32946	0.0020807

3.834	0.21055	0.001237
7.7436	0.46946	0.0032013
11.4516	0.44914	0.0022446
15.7608	0.78414	0.0045379
12.546	0.94187	0.0062013
8.6724	0.40076	0.0026739
18.6984	0.74836	0.004153
4.986	0.23939	0.0014065
40.77	2.301	0.014136
14.2452	0.61196	0.0037118
11.8584	0.57684	0.0036686
10.7748	0.54677	0.0035086
5.7456	0.31268	0.0019911
6.336	0.26136	0.001464
18.2052	0.92498	0.0055578
4.3596	0.25586	0.0017769
9.306	0.42344	0.0020348
22.2768	1.1082	0.0063461
3.1464	0.14415	0.00095372
8.6868	0.50251	0.0031783
24.8796	1.4821	0.0091577
7.2216	0.43518	0.0030575
4.5612	0.0024947	0.3105
13.2624	0.67315	0.0047676
17.2476	0.85335	0.0046254
3.4848	0.26832	0.0018302
20.8044	1.0995	0.006308
7.8768	0.44403	0.0031693
17.6904	0.95581	0.0057693
8.5752	0.43706	0.0027241
28.89	1.5413	0.0099498
10.8324	0.46047	0.0025111
5.3352	0.27644	0.0020935
18.3348	1.1424	0.0072331
11.1924	0.55726	0.0031746
26.7876	1.4612	0.0095269
3.8556	0.17859	0.00099997
14.1552	0.83206	0.0049762
4.6944	0.32646	0.0019419
14.8068	0.94696	0.0069502
19.5012	1.0916	0.0064984
13.8636	0.64979	0.003679
29.2608	1.5381	0.0091507
7.1676	0.48206	0.0034807
8.2908	0.44694	0.0030728
16.38	0.95839	0.0058703
4.2732	0.25348	0.0015677
5.1804	0.26235	0.0016532
19.0152	0.977	0.0057926
20.466	1.0901	0.0066466

9.6624	0.41168	0.0024332
9.7272	0.53457	0.0031706
6.0624	0.33797	0.0020439
12.3372	0.69315	0.0041297
7.1784	0.35964	0.0023691
11.952	0.69814	0.0053054
19.1772	1.0792	0.0070472
18.09	0.93852	0.0055979
8.2152	0.44302	0.0026692
4.7592	0.24427	0.0014593
23.3856	1.1559	0.0073077
12.1932	0.60161	0.0036688
11.52	0.56066	0.0032562
3.5028	0.22347	0.0015247
21.1788	1.4141	0.0090078
8.388	0.60615	0.0043689
10.2384	0.65625	0.0044838
13.626	0.70326	0.0039054
18.9504	0.74138	0.0039552
22.5864	0.95366	0.0052474
15.7104	0.82863	0.0055637
12.8088	0.7896	0.0044167
10.1124	0.53876	0.0030974
31.2012	1.4928	0.0084237
12.9132	0.57305	0.0031856
5.346	0.22125	0.0010722
24.9084	1.0277	0.0060848

Notes:

S_length = SmartRoot Length
S_Surface Area = SmartRoot Surface Area
S_Volume = SmartRoot Volume
sa_Length = saRIA Length
sa_Surface Area = saRIA Surface Area
sa_Volume = saRIA Volume

Bibliography

- [1] Jonathan Lynch. Root architecture and plant productivity. *Plant physiology*, 109(1):7, 1995.
- [2] Brenda B Casper and Robert B Jackson. Plant competition underground. *Annual review of ecology and systematics*, 28(1):545–570, 1997.
- [3] A Hodge, D Robinson, BS Griffiths, and AH Fitter. Why plants bother: root proliferation results in increased nitrogen capture from an organic patch when two grasses compete. *Plant, Cell & Environment*, 22(7):811–820, 1999.
- [4] AR Ennos. The scaling of root anchorage. *Journal of Theoretical Biology*, 161(1):61–75, 1993.
- [5] Karl J Niklas and Hanns-Christof Spatz. Allometric theory and the mechanical stability of large trees: proof and conjecture. *American journal of botany*, 93(6):824–828, 2006.
- [6] Stéphane Compant, Marcel GA Van Der Heijden, and Angela Sessitsch. Climate change effects on beneficial plant–microorganism interactions. *FEMS microbiology ecology*, 73(2):197–214, 2010.
- [7] H Charles J Godfray, John R Beddington, Ian R Crute, Lawrence Haddad, David Lawrence, James F Muir, Jules Pretty, Sherman Robinson, Sandy M Thomas, and Camilla Toulmin. Food security: the challenge of feeding 9 billion people. *science*, 327(5967):812–818, 2010.
- [8] JE Malamy. Intrinsic and environmental response pathways that regulate root system architecture. *Plant, cell & environment*, 28(1):67–77, 2005.
- [9] David M Eissenstat. On the relationship between specific root length and the rate of root proliferation: a field study using citrus rootstocks. *New Phytologist*, 118(1):63–68, 1991.

- [10] Sophie de Dorlodot, Brian Forster, Loïc Pagès, Adam Price, Roberto Tuberosa, and Xavier Draye. Root system architecture: opportunities and constraints for genetic improvement of crops. *Trends in plant science*, 12(10):474–481, 2007.
- [11] Adam Huw Price, KA Steele, J Gorham, JM Bridges, BJ Moore, JL Evans, P Richardson, and RGW Jones. Upland rice grown in soil-filled chambers and exposed to contrasting water-deficit regimes: I. root distribution, water use and plant water status. *Field Crops Research*, 76(1):11–24, 2002.
- [12] Yanying Qu, Ping Mu, Hongliang Zhang, Charles Y Chen, Yongming Gao, Yuxiu Tian, Feng Wen, and Zichao Li. Mapping QTLs of root morphological traits at different growth stages in rice. *Genetica*, 133(2):187–200, 2008.
- [13] Sarah M Rich and Michelle Watt. Soil conditions and cereal root system architecture: review and considerations for linking darwin and weaver. *Journal of experimental botany*, 64(5):1193–1208, 2013.
- [14] Guillaume Lobet, Xavier Draye, and Claire Périlleux. An online database for plant image analysis software tools. *Plant methods*, 9(1):38, 2013.
- [15] Samuel Trachsel, Shawn M Kaeppler, Kathleen M Brown, and Jonathan P Lynch. Shovelomics: high throughput phenotyping of maize (*zea mays* l.) root architecture in the field. *Plant and Soil*, 341(1-2):75–87, 2011.
- [16] Anjali S. Iyer-Pascuzzi, Olga Symonova, Yuriy Mileyko, Yueling Hao, Heather Belcher, John Harer, Joshua S. Weitz, and Philip N. Benfey. Imaging and analysis platform for automatic phenotyping and trait ranking of plant root systems. *Plant Physiology*, 152(3):1148–1157, 2010.
- [17] Deo A Heeraman, Jan W Hopmans, and Volker Clausnitzer. Three dimensional imaging of plant roots in situ with X-ray computed tomography. *Plant and soil*, 189(2):167–179, 1997.
- [18] JS Perret, ME Al-Belushi, and M Deadman. Non-destructive visualization and quantification of roots using computed tomography. *Soil Biology and Biochemistry*, 39(2):391–399, 2007.
- [19] Peter J Gregory, DJ Hutchison, Derek B Read, Paul M Jenneson, Walter B Gilboy, and Edward J Morton. Non-invasive imaging of roots

- with high resolution X-ray micro-tomography. In *Roots: the dynamic interface between plants and the Earth*, pages 351–359. Springer, 2003.
- [20] Saoirse R Tracy, Jeremy A Roberts, Colin R Black, Ann McNeill, Rob Davidson, and Sacha J Mooney. The X-factor: visualizing undisturbed root architecture in soils using X-ray computed tomography. *Journal of experimental botany*, 61(2):311–313, 2010.
- [21] Louise van der Weerd, Mireille MAE Claessens, Tom Ruttink, Frank J Vergeldt, Tjeerd J Schaafsma, and Henk Van As. Quantitative NMR microscopy of osmotic stress responses in maize and pearl millet. *Journal of Experimental Botany*, 52(365):2333–2343, 2001.
- [22] Henk Van As. Intact plant MRI for the study of cell water relations, membrane permeability, cell-to-cell and long distance water transport. *Journal of Experimental Botany*, 58(4):743–756, 2006.
- [23] Henk Van As, Tom Scheenen, and Frank J Vergeldt. MRI of intact plants. *Photosynthesis Research*, 102(2-3):213, 2009.
- [24] Suqin Fang, Xiaolong Yan, and Hong Liao. 3D reconstruction and dynamic modeling of root architecture in situ and its application to crop phosphorus research. *The Plant Journal*, 60(6):1096–1108, 2009.
- [25] Guang Zeng, Stanley T Birchfield, and Christina E Wells. Automatic discrimination of fine roots in minirhizotron images. *New Phytologist*, 177(2):549–557, 2008.
- [26] Rongli Shi, Astrid Junker, Christiane Seiler, and Thomas Altmann. Phenotyping roots in darkness: disturbance-free root imaging with near infrared illumination. *Functional Plant Biology*, 45(4):400–411, 2018.
- [27] Patrick Armengaud, Kevin Zambaux, Adrian Hills, Ronan Sulpice, Richard J Pattison, Michael R Blatt, and Anna Amtmann. EZ-Rhizo: integrated software for the fast and accurate measurement of root system architecture. *The Plant Journal*, 57(5):945–956, 2009.
- [28] Jordon Pace, Nigel Lee, Hsiang Sing Naik, Baskar Ganapathysubramanian, and Thomas Lübberstedt. Analysis of maize (*zea mays* l.) seedling roots with the high-throughput image analysis tool ARIA (Automatic Root Image Analysis). *PLoS One*, 9(9):e108255, 2014.

- [29] Daniel Leitner, Bernd Felderer, Peter Vontobel, and Andrea Schnepf. Recovering root system traits using image analysis exemplified by two-dimensional neutron radiography images of lupine. *Plant Physiology*, 164(1):24–35, 2014.
- [30] Jacques Le Bot, Valérie Serra, José Fabre, Xavier Draye, Stéphane Adamowicz, and Loïc Pagès. DART: a software to analyse root system architecture and development from captured images. *Plant and Soil*, 326(1-2):261–273, 2010.
- [31] J-L Arsenault, S Poulcur, C Messier, and R Guay. WinRHIZO: a root-measuring system with a unique overlap correction method. *HortScience*, 30(4):906D–906, 1995.
- [32] Guillaume Lobet, Loïc Pagès, and Xavier Draye. A novel image-analysis toolbox enabling quantitative analysis of root system architecture. *Plant physiology*, 157(1):29–39, 2011.
- [33] Stefan Mairhofer, Craig Sturrock, Darren Wells, Malcolm J. Bennett, Sacha Mooney, and Tony Pridmore. On the evaluation of methods for the recovery of plant root systems from X-ray computed tomography images. *Functional Plant Biology*, 42:460, 01 2015.
- [34] Richard J. Flavel, Chris N. Guppy, Sheikh M. R. Rabbi, and Iain M. Young. An image processing and analysis tool for identifying and analysing complex plant root systems in 3D soil using non-destructive analysis: Root1. *PLOS ONE*, 12(5):1–18, 05 2017.
- [35] Tao Wang, Mina Rostamza, Zhihang Song, Liangju Wang, G. McNickle, Anjali S. Iyer-Pascuzzi, Zhengjun Qiu, and Jian Jin. SegRoot: a high throughput segmentation method for root image analysis. *Computers and Electronics in Agriculture*, 162:845 – 854, 2019.
- [36] Abraham George Smith, Jens Petersen, Raghavendra Selvan, and Camilla Ruø Rasmussen. Segmentation of Roots in Soil with U-Net. *arXiv e-prints*, page arXiv:1902.11050, Feb 2019.
- [37] Guohao Yu, Alina Zare, Hudanyun Sheng, Roser Matamala, Joel Reyes-Cabrera, Felix B. Frischi, and Thomas E. Juenger. Root Identification in Minirhizotron Imagery with Multiple Instance Learning. *arXiv e-prints*, page arXiv:1903.03207, Mar 2019.
- [38] MathWorks. *MATLAB and Statistics Toolbox Release 2018b*, 2018.

- [39] Derek Bradley and Gerhard Roth. Adaptive thresholding using the integral image. *Journal of graphics tools*, 12(2):13–21, 2007.
- [40] Kelly H Zou, Simon K Warfield, Aditya Bharatha, Clare M C Tempany, Michael R Kaus, Steven J Haker, William M Wells, Ferenc A Jolesz, and Ron Kikinis. Statistical validation of image segmentation quality based on a spatial overlap index. *Academic radiology*, 11(2):178–189, 2004.
- [41] Alain Pierret, Santimaitree Gonkhamdee, Christophe Jourdan, and Jean-Luc Maeght. IJ_rhizo: an open-source software to measure scanned images of root samples. *Plant and Soil*, 373(1):531–539, Dec 2013.
- [42] Taras Galkovskyi, Yuriy Mileyko, Alexander Bucksch, Brad Moore, Olga Symonova, Charles A Price, Christopher N Topp, Anjali S Iyer-Pascuzzi, Paul R Zurek, Suqin Fang, et al. GiA Roots: software for the high throughput analysis of plant root system architecture. *BMC plant biology*, 12(1):116, 2012.
- [43] Michael P. Pound, Andrew P. French, Jonathan A. Atkinson, Darren M. Wells, Malcolm J. Bennett, and Tony Pridmore. RootNav: navigating images of complex root architectures. *Plant Physiology*, 162(4):1802–1814, 2013.

3. Fully-automated Root Image Analysis (faRIA)

Paper 2: Narisetti, N., Henke, M., Seiler, C. et al. Fully-automated root image analysis (faRIA). *Sci Rep* 11, 16047 (2021). <https://doi.org/10.1038/s41598-021-95480-y>

Impact Factor: 4.38 (according to SCImago Journal Rank, January 2024)

Abstract

High-throughput root phenotyping in the soil became an indispensable quantitative tool for the assessment of effects of climatic factors and molecular perturbation on plant root morphology, development and function. To efficiently analyse a large amount of structurally complex soil-root images advanced methods for automated image segmentation are required. Due to often unavoidable overlap between the intensity of fore- and background regions simple thresholding methods are, generally, not suitable for the segmentation of root regions. Higher-level cognitive models such as convolutional neural networks (CNN) provide capabilities for segmenting roots from heterogeneous and noisy background structures, however, they require a representative set of manually segmented (ground truth) images. Here, we present a GUI-based tool for fully automated quantitative analysis of root images (faRIA) using a pre-trained CNN model, which relies on an extension of the U-net architecture. The developed CNN framework was designed to efficiently segment root structures of different size, shape and optical contrast using low budget hardware systems. The CNN model was trained on a set of 6465 masks derived from 182 manually segmented near-infrared (NIR) maize root images. Our experimental results show that the proposed approach achieves a Dice coefficient of 0.87 and outperforms existing tools (e.g., SegRoot) with Dice coefficient of 0.67 by application not only to NIR but also to other imaging modalities and plant species such as barley and arabidopsis soil-root images from LED-rhizotron and UV imaging systems, respectively. In summary, the developed software framework enables users to efficiently analyse soil-root images in an automated manner (i.e. without manual interaction with data and/or parameter tuning) providing quantitative plant scientists with a powerful analytical tool.

3.1 Introduction

Image based high-throughput phenotyping of roots is one of the emerging disciplines in plant phenomics. It aims to extract the plant morphological and physiological properties in a non-destructive manner to study the plant performance under given conditions [1]. Traditional approaches to root phenotyping have relied on destructive and artificial grown mediums such as liquids or gels [2, 3]. However, the root growth is known to be dependent on physical conditions [4] and such studies have shown a non-typical response of the roots in soil [5, 6].

More recently, non-destructive methods such as X-ray computed tomography [7, 8], nuclear magnetic resonance (NMR) microscopy [9] and laser scanning [10] provide unique insights into 3D organization of living root architecture, however, their throughput capabilities are presently rather limited. Moreover, minirhizotrons [11, 12] and rhizotron systems [13, 14] have gained popularity to enable non-invasive imaging of roots in a soil environment. However, the minirhizotrons require a repeated photographing of roots through a transparent surface of below ground observation tubes [15]. In contrast, rhizotron systems contain rectangular glass pots which requires a single photographing of roots [16]. Recently, near-infrared (NIR) imaging of roots growing along transparent pots were presented in our previous works [17, 18]. These systems contain special low pass filters to block root exposure to visible light and the images were taken by NIR camera under suitable illumination.

Due to high level of optical soil heterogeneity, soil-root images exhibit a relatively low contrast between back- and foreground structures. Consequently, at the local scale root and soil pixels cannot be distinguished on the basis of their intensity values only. Several root image solutions were suggested in the past, however, most of them were designed for a specific imaging system [19, 20, 21, 22, 23]. Examples of general-purpose semi-automated tools include GiA Roots [24], IJ-Rhizo [25] as well as our previously published saRIA software [26]. All these tool rely on thresholding and morphological filtering techniques to segment the roots from background. Other root phenotyping solutions like SmartRoot [27, 28] require manual segmentation by placing multiple landmarks along the roots that are subsequently interconnected to the root skeleton. All the above software solutions are time consuming, have limited throughput capabilities, and require expertise in parameter tuning.

To overcome the limitations of existing methods, automated root image segmentation solutions are required for high-throughput root image segmentation and phenotyping. In the last five years, deep learning gained high attention especially in computer vision applications, because of the ability to directly extract and train relevant multi-level features from data without prior knowledge and human effort in feature design. Convolutional neural networks (CNNs) are a class of deep learning approaches that have shown to outperform traditional methods in many applications of the computer vision that are associated with higher level cognitive abilities [29]. CNNs have been shown to outperform conventional approaches when applied to traditionally difficult tasks of image analysis including pattern detection and object segmentation in biomedical images [30, 31], traffic scenes [32] and remote sensing [33]. In recent years, they were also used for high-throughput plant phenotyping such as detection of wheat roots grown in germination paper [34], segmentation of roots from soil in X-ray tomography [35] and segmentation of spikes in wheat plants [36]. However, most of these works present exemplary application and/or computational frameworks that can hardly be handled by end-users without advanced programming skills.

The focus of this work is on semantic segmentation of soil-root images by which root pixels are automatically segmented from soil regions. For this kind of approach, CNNs often use encoder-decoder architecture. Till date, several papers have been published on this type of CNN architecture for biomedical [30, 31] and areal applications [32, 33]. Moreover, this type of architectures are constantly improving by cascading or fusing the CNNs in biomedical [37, 38] and remote sensing applications [39].

Application of CNNs to automated image analysis and plant phenotyping became an emerging trend in quantitative plant sciences in the recent years [40]. However, reliable software tools suitable for a particular plant type are rarely available due to the large variability of optical plant appearance, differences between experimental setups [35, 40], and the absence of labelled ground truth data [41, 42]. Consequently, only a few software tools for high-throughput plant image analysis and phenotyping are presently known.

Previously published state of the art encoder-decoder CNN solutions for root image segmentation include RootNav 2.0 [43], SegRoot [44] and RootNet [45]. Among those, RootNav 2.0 and RootNet tools were primarily developed for particular experimental setups such as roots grown on germination paper with high contrast between root and (blue) background pixels, and, thus, cannot be expected to perform accurately by application to other imaging

modalities such as noisy soil-root images in this work.

Among the above mentioned tools, SegRoot appears to be the most suitable one for soil-root image segmentation as it is previously shown to be capable of segmenting roots from soil background in minirhizotrons systems. Moreover, the architecture of SegRoot is somewhat similar to U-net and it transfers the location of feature maps to decoder for image segmentation. However, this approach failed to detect fine, blurry and low contrast roots, which, in turn, compromises the accuracy of resulting phenotypic traits such as estimated root biomass and other geometric features. To overcome these limitations, here, we adopted a U-net [30] based encoder-decoder architecture which transfers both location and pixel information of the feature maps to the decoder. Also, it is especially useful when large amount of manually annotated data is challenging, such as often the case in biomedical applications.

The aim of this work was to develop an efficient and handy tool for fully automated root image segmentation and quantification using a pre-trained deep CNN framework which could be used in a straightforward manner even by unskilled users. Although, our approach relies on supervised model training, for the end-users such a model-based image analysis is performed in a fully automated manner (i.e. without interaction with data and/or parameter tuning) in contrast to purely manual or semi-automated image segmentation approaches where such interactions are required. Consequently, we termed this approach fully-automated root image analysis (faRIA). The main contributions of this work include:

- Development of a CNN approach to automated root image segmentation based on the U-net architecture from [30],
- Training and application of the CNN model for efficient segmentation of root structures of different size, shape and optical contrast on low budget hardware systems using image masking approach,
- Evaluation and comparison of our CNN model vs. other state-of-the-art tools for root image analysis using the Dice similarity metrics,
- Evaluation of our CNN framework performance on images of different root imaging modalities,
- Development of a GUI based front-end for efficient handling of the algorithmic framework suitable also for IT-unskilled users.

The paper is structured as follows: first, we describe the methodological framework of proposed U-net based deep learning algorithm and performance matrices for soil-root segmentation. Then, a brief experimental setup consist of data preparation, training and prediction procedure are discussed. Followed by, the results of experimental investigation are presented including a comparison of faRIA performance to other image segmentation tools, performance on resized images and robustness by application to other image modalities and plant species. In Discussion, we summarize the results of an evaluation study using faRIA image segmentation and present its GUI implementation for efficient application in high-throughput root phenotyping.

3.2 Methods & Computational Setup

3.2.1 Deep CNN Model for Root Image Segmentation

The proposed CNN architecture is derived from the original U-net [30] which provides a versatile framework for semantic image segmentation consisting of encoder and corresponding decoder units. Our CNN model has a depth of 3 which is less than original U-net depth of 4 due to the smaller input image size. Further, in our approach the batch normalization [46] is applied after each convolutional layer in contrast to the original U-net architecture where it was not the case. The motivation behind the batch normalization is it is known to make model performance more faster and stable [47, 46]. Furthermore, the original U-net [30] used Dropout layer which we avoided because in some cases the combination of batch normalization and dropout layers can cause worse results [48]. Also, kernel size of the convolutional layers was set larger in our approach than in the original U-net to improve the continuity in segmentation of roots [49]. The details of the convolutional parameters in comparison to the original U-net are summarized in Table 3.1.

Motivated by the encoder-decoder architecture of U-net, a network framework for soil-root image segmentation was constructed, see Figure 3.1. In particular, our network was designed to be trained on patches of input images in original resolution. This was introduced in order to enable model training using larger amount of ground truth data on consumer GPUs while preserving high-frequency image information which otherwise would be lost either by restricting the training set to maximum possible capacity of GPU RAM or by image downscaling. Furthermore, training of CNN on image patches instead of full-size images is known to be more advantageous for learning local features [50]. Therefore, the architecture was designed in such a way that it has input and output layers of the size 256x256. In what fol-

Table 3.1: Convolutional parameters of the original U-net and proposed modifications.

Convolutional parameters	Original U-net	Proposed modifications
Kernel size	3x3	7x7
Transposed kernel size	2x2	3x3
Stride	1x1	2x2
Padding	unpadded	padding with zeros
Depth	4	3
Number of filters	(64, 128, 256, 512, 1028)	(16, 32, 64, 128)

lows, the details of network encoder and decoder layers are described.

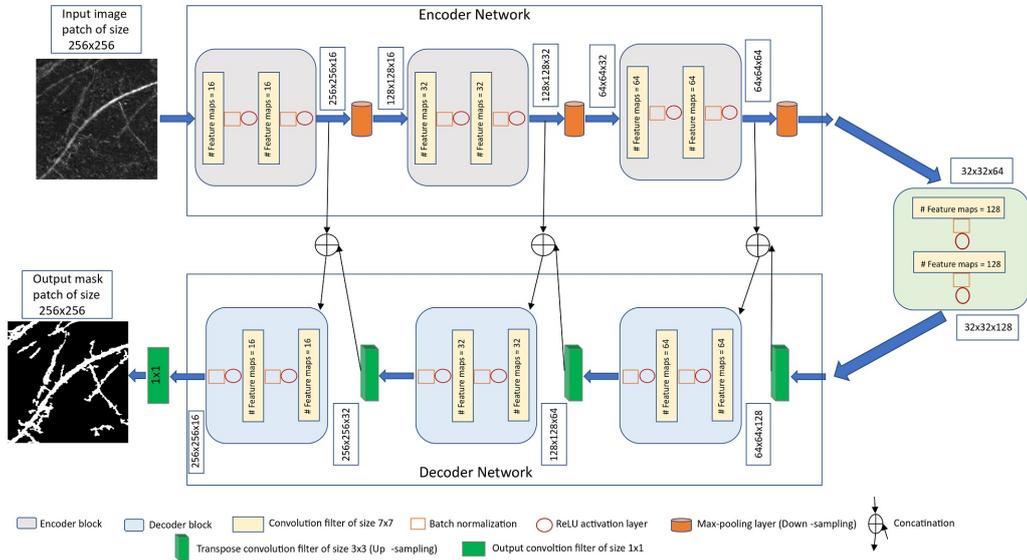


Figure 3.1: The proposed U-net architecture for soil-root image segmentation.

Encoder network: The encoder network consists of 3 encoder blocks. The

first encoder block takes the image patches of size 256×256 as input and produces corresponding feature maps of size $(256 \times 256 \times 16)$ as output. Then the feature maps are forwarded to the second and third encoder blocks to generate further feature maps for the root detection. Each encoder block consists of two convolutional layers to learn feature maps at respective levels, where each convolutional layer consists of 7×7 convolution filter followed by batch normalization [46] and a non-linear activation function called Rectified Linear Unit (ReLU) [51]. Here, batch normalization improves the network performance and stability by normalizing the feature maps at respective levels [46]. Followed by each encoder block, max-pooling operation using general window size of 2×2 [50, 52] is applied for down-sampling the feature maps by half of its original size. This results in aggregate features are generated more efficiently. All three encoders are repeated with varying depth of 16, 32 and 64 to detect diverse root features respectively. The details of each encoder block and corresponding max-pool layers are given in Table 3.2.

Followed by encoder network, a bridge encoder block without max-pooling layer is applied. This results in 128 feature maps of each size 32×32 are generated.

Decoder network: The output from the bridge encoder ($32 \times 32 \times 128$) is upsampled using 3×3 transpose convolution with same padding and stride 2. This means size of feature maps ($32 \times 32 \times 128$) were double to ($64 \times 64 \times 128$) by applying filter of size 3×3 to all input elements and border elements were computed using zero padding. Then the resulting feature map is concatenated with the corresponding encoder feature maps. This results in feature maps of size $(64 \times 64 \times 256)$ are generated. Then it is passed through a convolutional layers like encoder block but having decreasing channel depth of 64. This process is repeated for remaining decoder blocks with decreasing channel depth of 32 and 16. The details of each decoder block and corresponding transpose layer outputs are given in Table 3.3. Finally, the output of the final decoder block is fed into a convolutional layer of size $1 \times 1 \times 1$ with "Softmax" activation function [53] to classify each pixel as root or non-root at the patch level. The output of proposed architecture is a predicted mask of size 256×256 like the input image patch a shown in Figure 3.1.

3.2.2 Performance Metrics

To evaluate the performance of the proposed U-net model during training and testing stage, Dice coefficient (DC) [54] is used. It measures the area of intersection between the model and ground truth segmentation and its

Table 3.2: Details of all encoder blocks and corresponding max-pool layer output

Encoder Block #	Input to encoder block	Conv-olution filter size	Number of feature maps	Output of encoder block	Input to max-pool	Max-pool output
Block 1	256x256	7x7	16	256x256x16	256x256x16	128x128x16
Block 2	128x128	7x7	32	128x128x32	128x128x32	64x64x32
Block 1	64x64	7x7	64	64x64x64	64x64x64	32x32x64

Table 3.3: Details of all decoder blocks and corresponding transpose convolutional layers

Decoder Block #	Input to trans-posed convolu-tion	Output of trans-posed convolu-tion	Number of de-coder blocks	Conv-olution filter size	Number of fea-ture maps	Output of de-coder block
Block 1	32x32x128	64x64x128	64x64x128	7x7	128	64x64x64
Block 2	64x64x64	128x128x64	128x128x64	7x7	64	128x128x32
Block 2	128x128x32	256x256x32	256x256x32	7x7	32	256x256x16

value ranges from 0 to 1, where 1 corresponds to 100% perfect and 0 to false segmentation. The Dice coefficient is defined as:

$$DC = \frac{2 * (P \cap G)}{P \cup G} = \frac{2 * \sum_i^N P_i G_i}{\sum_i^N P_i + \sum_i^N G_i}, \quad (3.1)$$

where P and G are predicted and ground truth binary images respectively. P_i and G_i are output values 0 and 1 of pixel i in predicted and ground truth binary image respectively. Also, the above equation can be re-written as following:

$$DC = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}. \quad (3.2)$$

From Equation 3.2 it follows that the model would likely overestimate soil pixels and underestimate root pixels in the segmented image, because root images typically contain significantly more background pixels than root pixels. In that case, precision defines the ratio of correctly predicted root pixels to the number of pixels predicted to be root and recall is the ratio of correctly predicted root pixels to the number of actual root pixels in the image.

3.2.3 Data and Image Annotation

Near-infrared (NIR) images of Maize plant roots grown in soil were captured by using IPK plant phenotyping system for large plants [17]. Images are taken by one side-view 12MP monochrome camera (UI-5200SE-M-GL, IDS) with chip sensitive in NIR portion of electromagnetic spectrum and suitable distortion-free lens (V1228-MPY). Also, it includes homogeneous infrared LED light source (850 nm) and filters preventing reflections during image acquisition. In brief, plants were grown in rhizopots [342x350 mm (WxL)] filled with the potting substrate (Potgrond P, Klassmann).

200 greyscale root images of Maize plants acquired with the IPK plant phenotyping system were selected for the ground truth segmentation. This labelling task is performed by agronomists using our previously published software for semi-automated root image analysis (saRIA) [26] which provides an efficient graphical user interface for tuning parameters of image segmentation including intensity threshold, morphology and noise removal to generate an accurate segmentation of roots in soil. The images acquired with the above imaging system have resolution of 2345x2665. A detailed root annotation with saRIA took approximately 5-10 minutes per image depending on the amount of root pixels in the image. Figure 3.2 shows an example of IPK plant phenotyping system images and their corresponding binary segmentation using saRIA. This binary mask contains all roots as foreground in white and the remaining pixels as background in black.

To enable application of the proposed model to a broad range of root imaging modalities, the model originally developed for NIR root image segmentation was applied to LED-based rhizotron and ultraviolet (UV) imaging systems [18, 26]. In fact, such approach is feasible because root structures in both image modalities exhibit large similarities. The rhizotron system contains a root camera (Allied Vision Prosilica GT 6600) and uses white LED illumination to image the roots growing in soil along plexiglass plates. The UV system contains two monochrome UV-sensitive cameras (UI-5490SE-

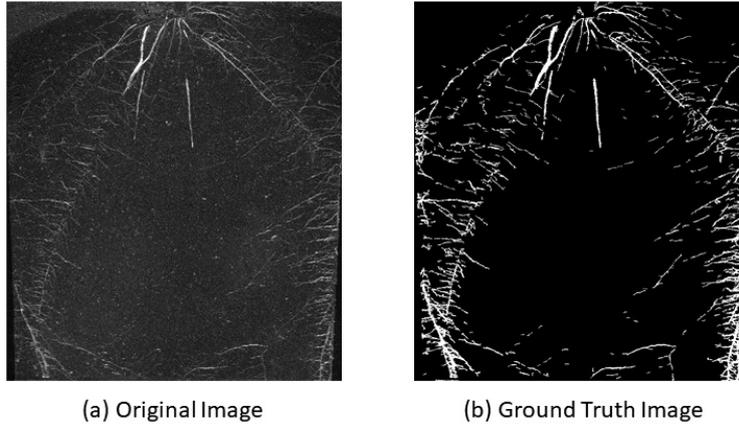


Figure 3.2: Exemplary root image from IPK plant phenotyping system: (a) Maize plant roots 28 days after sowing, (b) corresponding root segmentation from soil performed using saRIA.

M-GL, IDS) with two sets of LED illumination panels (UV, 380 nm) in a custom-made imaging box. It is suitable for capturing small plants in transparent pots of size [77x77x97mm (WxLxH)] filled with the potting substrate (Potgrond P, Klasmann). This system allows non-invasive acquisition of root images in darkness [18].

3.2.4 Training

The proposed U-net model was developed under Python 3.6.1 using TensorFlow [55] library with Keras API [56, 57]. Image processing functions like cropping and morphological functions (dilation, erosion) were implemented using PIL, Numpy [58] and Scikit-Image [59] packages. Then the model was trained on Linux operating system (Intel(R) Xeon(R) Gold 6130 CPU @ 2.10GHz) with NVIDIA Tesla P100-PCIE-16GB graphic card.

Images analysed in this work contain both thin and fine root structures that may have only one or few pixels in width. To preserve such fine structures the binary masks were dilated similar to strategy applied in SegRoot [44]. Originally 2345x2665 sized root images of maize plants are analysed step-wise using the 256x256 crop masks. Thereby, the original image edges were padded with zeros so that both its width and height are divisible by 256. Hence, original image size is increased to 2560x2816 with zero-padding. Then each image is partitioned into 110 non-overlapping 256x256 crop masks

and approximately 20000 crop masks are generated for all images. However, 2/3 of those cropped masks contain only background structures that contribute to training network the background appearance. To avoid potential imbalance between plant and non-plant training masks, only cropped regions with both root and background pixels information of the size 256x256 were selected from 182 original images. Then each cropped image is normalized in the range of [0, 1] for feature consistency in the CNN network.

Subsequently, the data set was partitioned into training and validation sets in the ratio of 85:15. The training set is used to optimize the proposed model with Adam optimizer [60] in such way that the weight parameters improve the model segmentation performance. Also, the initial weights of the networks were defined randomly as proposed by *Krizhevsky et al.* [61] with the mean 0 and the standard deviation of 0.05. Here, the model training was initialized for maximum of 200 epochs with 16 number of convolutional channel features and batch size of 128 as per system constraints. Loss functions quantify the unhappiness of our network during training and it defines the difference between predicted output and ground truth generated by saRIA. The result of loss function can be improved by updating weights of the network in an iterative manner. Here, more commonly used "binary cross-entropy loss" function [50] is used to predict binary class label (i.e., roots and non-roots) at each patch level. This function compares each pixel prediction (0: non-root, 1: root) to the corresponding ground truth pixel and averages all pixels loss for computing total loss of the image. Therefore, each pixel contributes to the overall objective loss function. Then the learning rate of the Adam optimizer [60] was estimated from a range of reasonable values (0.00001, 0.0001, 0.001, 0.1, 1 and 10) while monitoring the training and validation Dice coefficient of the model.

3.2.5 Prediction

As stated in image annotation subsection, the images from IPK plant phenotyping system have the original resolution of 2345x2665, while the proposed U-net model requires input images of the size 256x256. In the preprocessing stage, zero padding is applied to test images similar as it was done in the training stage. Then non-overlapping 256x256 masks were generated. The model does predictions on these 256x256 masks that are then combined to one single output image. Finally, the zero padded pixels were removed and the segmented image with resolution identical to the original input image was generated. This complete process is dynamic and automatized in the prediction stage as shown in Figure 3.3. Since the output layer is given by

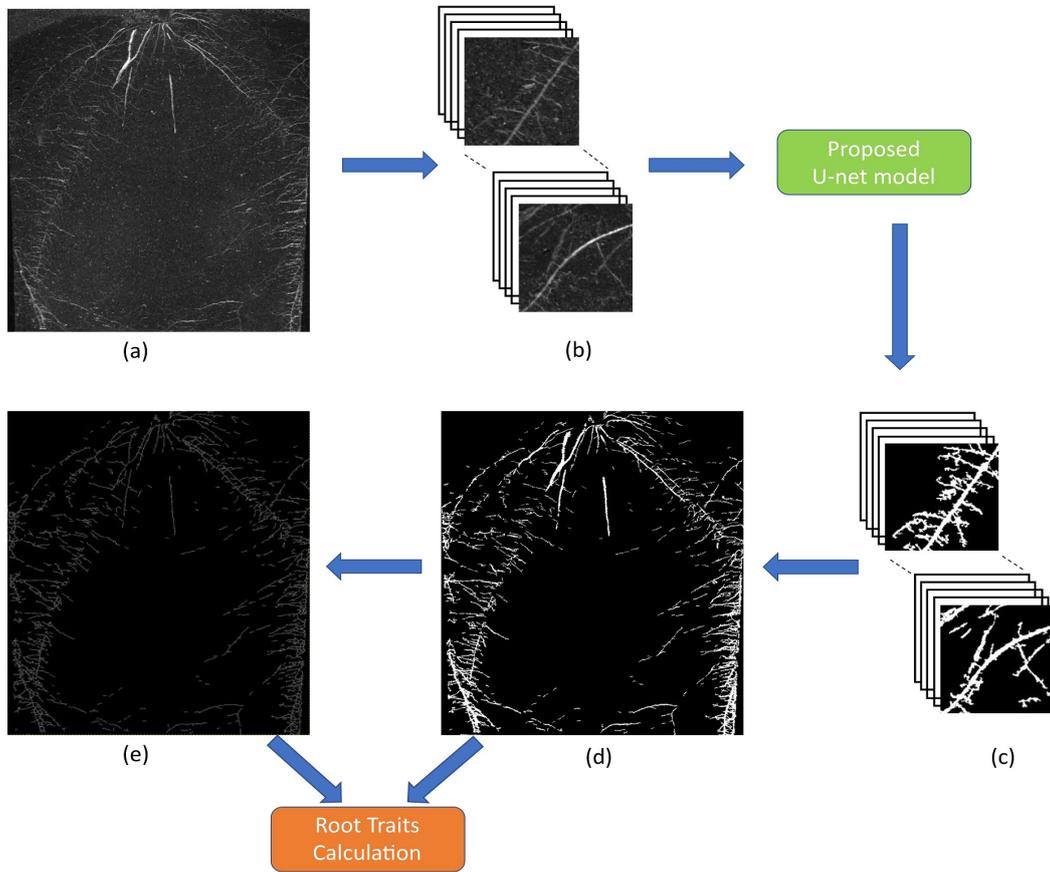


Figure 3.3: Workflow of the pipeline for image processing and segmentation in faRIA. Green and orange colour boxes represents the operations of image segmentation and trait calculation: (a) original image, (b) original image patches of size 256x256, (c) segmented image patches of size 256x256, (d) binary segmentation of original image, (e) binary skeleton of (d).

the Sigmoid activation function, the predicted segmentation is a probability map with values ranging between 0 and 1. Hence the generated probability map was converted to a binary image using threshold T . Here, the root pixels with a relatively high $T \geq 0.9$ is chosen to avoid misclassification for the soil-root image segmentation. After fully automated segmentation, the proposed model performs phenotyping of segmented root structures similar to saRIA [26].

In practice, the end-users prefer to have an easy-to-use software solution including the Graphical User Interface (GUI). Therefore, a user-friendly

GUI front-end was developed under the MATLAB 2019b environment [62] to comfortably operate the complex algorithmic framework of faRIA software. Figure 3.3 shows the complete workflow involved in faRIA for automatic root segmentation and trait extraction. For import of deep learning models trained under Python the MATLAB interoperability routine *importKerasNetwork* [62] was used. According to specification of this function, the U-net models trained in Python were exported in the so-called h5 file format, which is supported by the recent versions of MATLAB including 2019b.

In addition to 256 cropped masks, the proposed U-net model was extended to train on full images. This model has an input and output images of size 1024x1024 as per our system constraints. So that the original and ground truth images were resized to 1024x1024 using bi-linear interpolation method [63]. Also, the model consist of an additional encoder and decoder blocks with convolution mask of size 5x5 in their respective networks. Therefore, encoder network generates the feature maps from size 1024x1024x1 to 32x32x128 and inverse size in the decoder network. To distinguish both networks, the proposed U-net model on 256 and 1024 masks are named as faRIA:256 and faRIA:1024, respectively.

3.3 Results

3.3.1 Training and Validation of faRIA

As discussed above, the training and validation of faRIA:256 model was performed on totally 6465 image patches in the ratio of 85:15 between train and test images, respectively. The performance of the trained model is analysed using binary cross-entropy loss, Dice coefficient, precision and recall at each epoch during learning stage of the network. Figure 3.4 shows the training and validation performance of the faRIA:256 over 200 epochs. It turned out that the training loss (Figure 3.4a) was minimized and platen the curve near to zero after epoch number 140. Simultaneously, training DC, precision and recall were maximized and achieved more than 90% of the accuracy from epoch number 100. But generalized performance of the model is measured using validation parameters. Figure 3.4b explains that the proposed model achieved maximum validation Dice coefficient of 0.874 and minimum validation loss of 0.033 at epoch number 71.

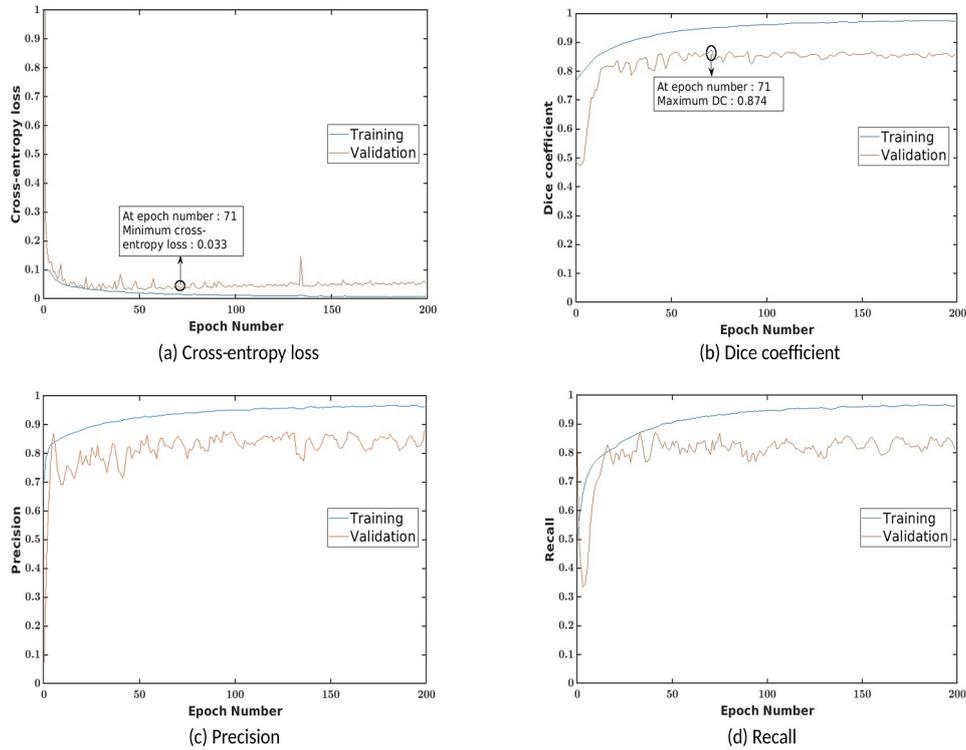


Figure 3.4: Training and validation performance of the faRIA:256 model over 200 epochs. X- and Y-axes represent the epoch number and performance measure, respectively.

3.3.2 Evaluation of faRIA vs. SegRoot

For comparing the performance of faRIA:256 model with existing tools, SegRoot [44] was trained on the same image data set. For this purpose, the SegRoot model was trained on 256x256 image blocks for 200 epochs with best practical parameters of depth 5 and width 8 as suggested in Wang *et al.* [44]. In addition, to validate the performance of proposed model on full image instead of 256x256 blocks (faRIA:256), faRIA:1024 was proposed. The faRIA:1024 model was trained for 200 epochs with training configurations similar to faRIA:256. Table 3.4 and 3.5 show the training parameters and performance measures of the faRIA:256 with respect to SegRoot and faRIA:1024.

Followed by training performance, an exemplary performance of above three models on test image was performed, see in Figure 3.5. Thereby, the

Table 3.4: Training parameters of SegRoot, faRIA:1024 and faRIA:256 over 200 epochs

Training parameter	SegRoot-8-5	faRIA:1024	faRIA:256
Learning rate	0.01	0.001	0.001
Batch size	128	3	128
Epochs	200	200	200

Table 3.5: Training performance of SegRoot, faRIA:1024 and faRIA:256 over 200 epochs

Validation measure	SegRoot-8-5	faRIA:1024	faRIA:256
Cross-entropy loss	0.374	0.043	0.033
Dice coefficient	0.666	0.888	0.874
Precision	0.652	0.901	0.849
Recall	0.735	0.824	0.846

faRIA:256 model showed the DC of 0.83 whereas SegRoot and faRIA:1024 achieved 0.42 and 0.44 respectively. Also, the presence of marginal artefacts in faRIA:1024 and faRIA:256 compared to ground truth are shown in Figure 3.6.

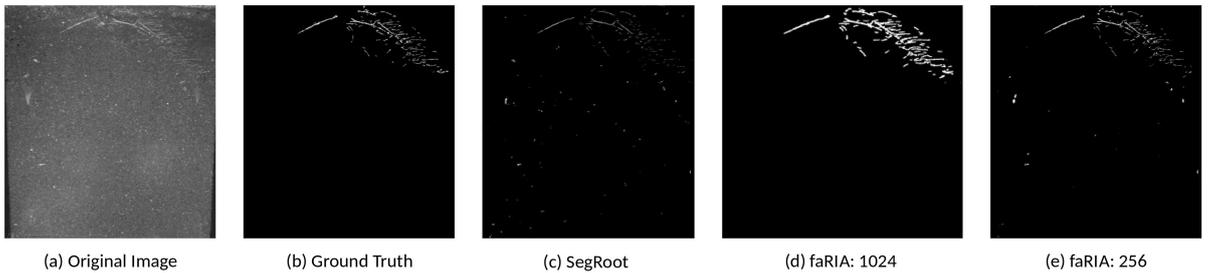


Figure 3.5: Segmentation performance: (a) original image, (b) ground truth segmentation by saRIA, (c) SegRoot with DC: 0.42, (d) faRIA:1024 with DC: 0.44, (e) faRIA:256 with DC: 0.83.

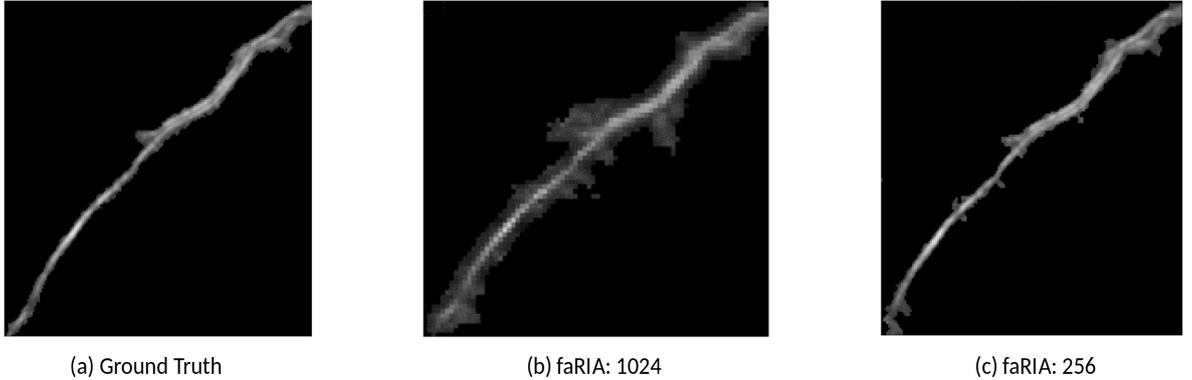


Figure 3.6: Segmentation artefacts: (a) original root structure, (b) faRIA:1024: noisy segmentation of (a) at root edges, (c) faRIA:256: noise-free segmentation of (a).

3.3.3 Segmentation of Further Image Modalities

The faRIA:256 model originally trained on maize plant roots from IPK plant phenotyping system is applied to LED-based rhizotron and UV imaging systems for the root segmentation from soil. Figure 3.7 and 3.9 shows the DC of faRIA:256 model over 40 barley and 30 arabidopsis root images from rhizotron and UV imaging system and achieved mean DC of 0.85 and 0.68 respectively. An exemplary segmentation of rhizotron (image number 4 in Figure 3.7) and UV image (image number 6 in Figure 3.9) are shown in Figure 3.8 a, b, c, e and 3.10 a, b, c, e respectively. Here, the faRIA:256 model resulted DC of 0.87 and 0.79 for rhizotron and UV image compared to the ground truth generated by saRIA respectively. In addition, the performance of the SegRoot on same rhizotron and UV image compared to the ground truth is shown in the Figure 3.8 d, f and 3.10 d, f respectively. Here, false negative (green) and false positive (pink) pixels represents the undetected and falsely classified root pixels in the predicted segmentation compared to the ground truth.

3.3.4 Evaluation of Phenotypic Traits vs. saRIA

In addition to the segmentation performance, phenotyping characterization obtained with faRIA are also evaluated in comparison to saRIA. Here, correlation coefficient of determination R^2 and significance level p-value are used to measure the percent of the faRIA calculated traits that are close to the

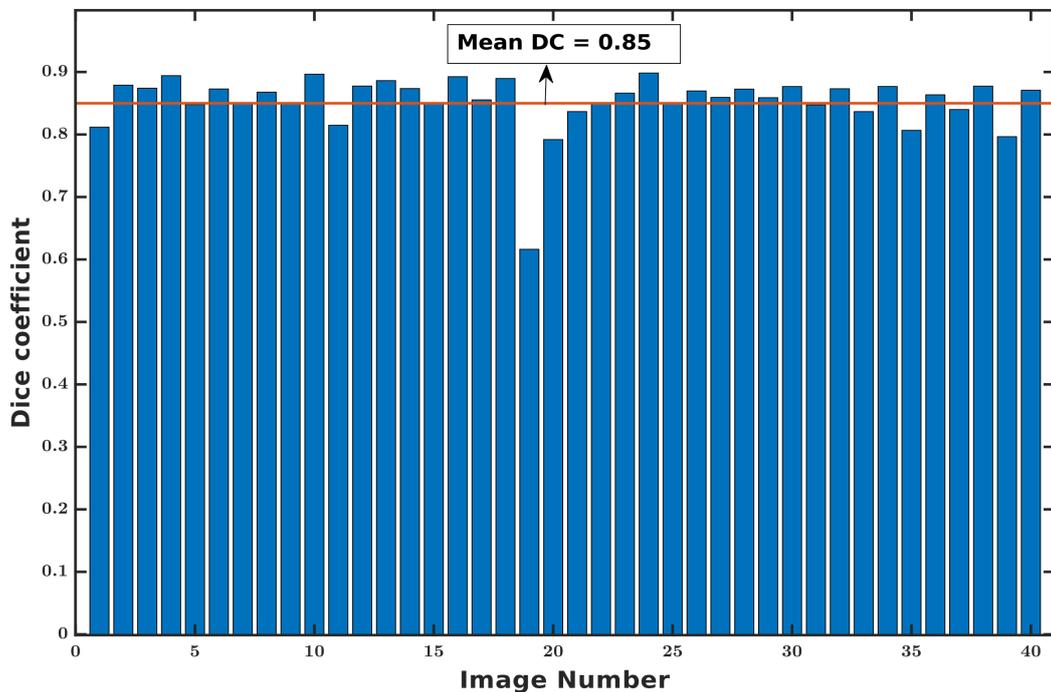


Figure 3.7: Dice coefficient of faRIA:256 over 40 barley soil-root images from rhizotron imaging system. The orange line represents the mean DC value.

ground-truth (from saRIA) and model validation respectively. Figure 3.11 shows the correlation between the saRIA (x-axis) and faRIA (y-axis) outputs for four traits where each point denotes one particular image out of 40 Barley root images from rhizotron imaging system. Out of 75 traits, only four important traits for root biomass calculation are presented for faRIA evaluation. They are total root area, total root length, total root surface area and total root volume. Further information on definition of traits is included in the Supplementary Information, see Table S1. Figure 3.11 shows that correlations between traits calculated with saRIA and faRIA are highly significant and exhibit R^2 values greater than 0.98, 0.97, 0.98 and 0.98 and p-values $1.59e-40$, $5.01e-38$, $7.63e-42$, and $5.13e-42$, respectively.

3.3.5 Graphical User Interface and Runtime

Figure 3.12 shows the GUI of faRIA software which is freely available as a pre-compiled executable program from <https://ag-ba.ipk-gatersleben.de/faria.html>. In addition to fully automated image segmentation, faRIA calculates 75 root

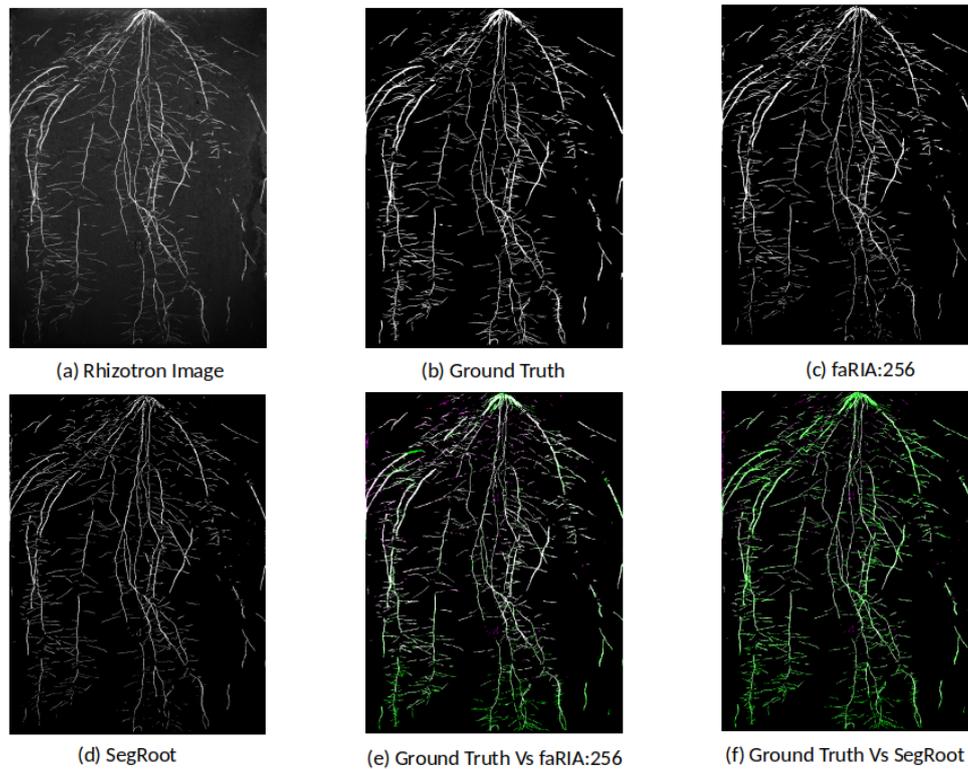


Figure 3.8: Applicability of faRIA on rhizotron soil-root images: (a) image of Barley plant roots at lateral stage, (b) ground truth segmentation generated by saRIA, (c) predicted segmentation using faRIA:256, (d) predicted segmentation using SegRoot, (e) overlay of faRIA prediction on ground truth with $DC = 0.87$, (f) overlay of SegRoot prediction on ground truth with $DC = 0.73$. Green, pink and white colour pixels represents false negatives, false positives and correctly segmented pixels in the predicted image with respect to ground truth respectively.

traits that are categorized into 12 feature groups named area (number of root pixels), number of disconnected root objects, total length, surface area, volume, number of branching and ending points, statistical distribution (mean, median, standard deviation, skewness, kurtosis, percentile and bootstrap) of root geometry in horizontal and vertical direction, width, orientation and convex-hull. In the present release, the phenotyping module of faRIA is identical to our saRIA software [26]. Further information on definition of traits is included in the Supplementary Information, see Table S1.

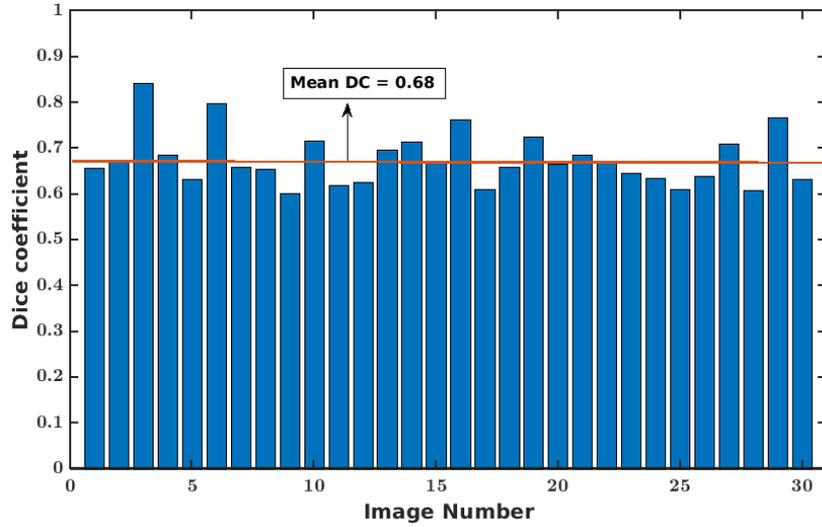


Figure 3.9: Dice coefficient of faRIA:256 over 30 Arabidopsis soil-root images from UV imaging system. The orange line represents the mean DC value.

The faRIA software provides users with an option to select faRIA:256 or faRIA:1024 model depending on image quality, time and accuracy. The faRIA software can analyse a single image or large image data set to automatically detect and extract multiple root traits. Regarding timing performance, the faRIA segmentation, root tracing and trait calculation all together take, in average, 80 seconds using faRIA:256 and 15 seconds using faRIA:1024 models to process and analyse a 6-megapixel (cropped) image on a system with Intel(R) Xeon(R) Gold 6130 CPU @2.10GHz. Therefore, faRIA:1024 can process at least 3 times faster than faRIA:256 for root image analysis.

3.4 Discussion & Conclusion

Our experimental results on different plant species from different imaging systems have demonstrated a remarkable accuracy of an adopted U-net model for fully automated soil-root image segmentation. During the training stage, the faRIA:256 model achieved nearly zero loss and $\geq 95\%$ of accuracy measured by the Dice coefficient (DC) crossover 200 epochs, see Figure 3.4. By application to the test images, the best performance was found at the epoch number 71 with the maximum DC of 0.874 and minimum loss of 0.033. For larger number of epochs, validation error was just marginally higher. However, the precision and recall are contrasting each other at low DC epochs,

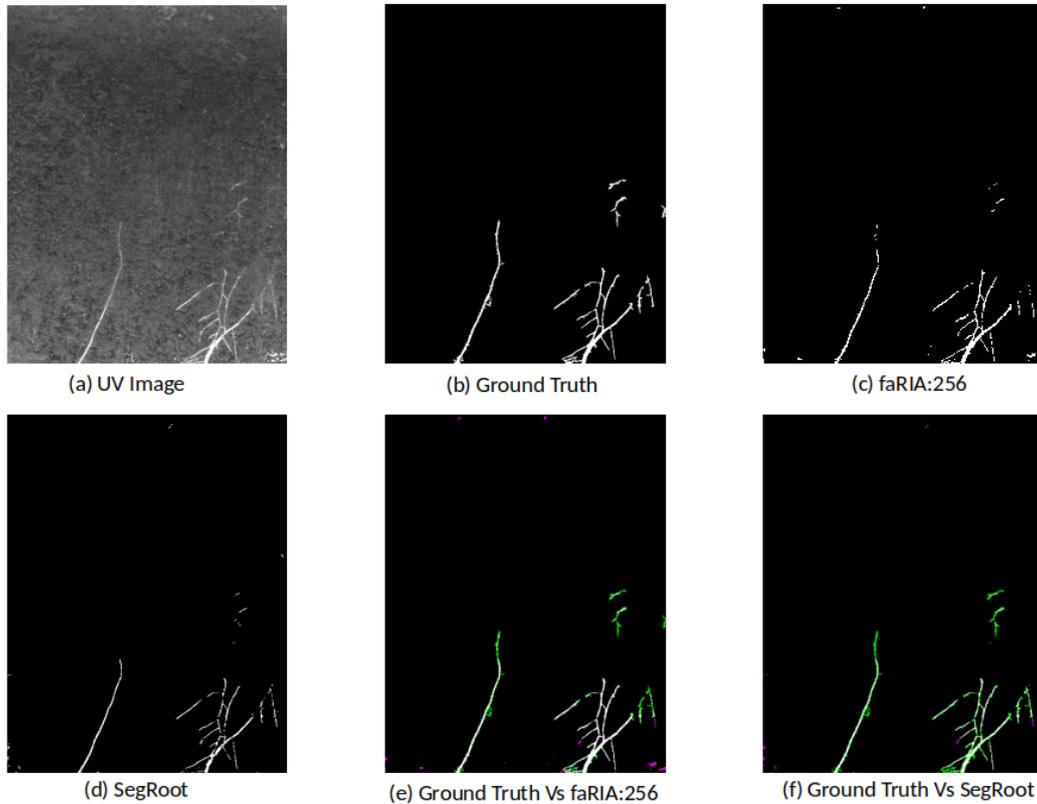


Figure 3.10: Applicability of faRIA on UV soil-root images: (a) image of Arabidopsis plant roots from UV imaging system, (b) ground truth segmentation generated by saRIA, (c) predicted segmentation using faRIA:256, (d) predicted segmentation using SegRoot, (e) overlay of faRIA prediction on ground truth with $DC = 0.80$, (f) overlay of SegRoot prediction on ground truth with $DC = 0.67$. Green, pink and white colour pixels represents false negatives, false positives and correctly segmented pixels in the predicted image with respect to ground truth respectively.

and both achieved maximum at epoch number 71. Therefore, the network weights and optimization parameters at epoch number 71 are adopted as the best model for soil-root image segmentation.

The performance of the faRIA:256 model was compared with the SegRoot. From the summary in Table 3.5, it is evident that faRIA:256 is significantly outperforming the SegRoot on our data set improving the cross-entropy loss by the factor 10 and DC by 20%, respectively. We draw this results back

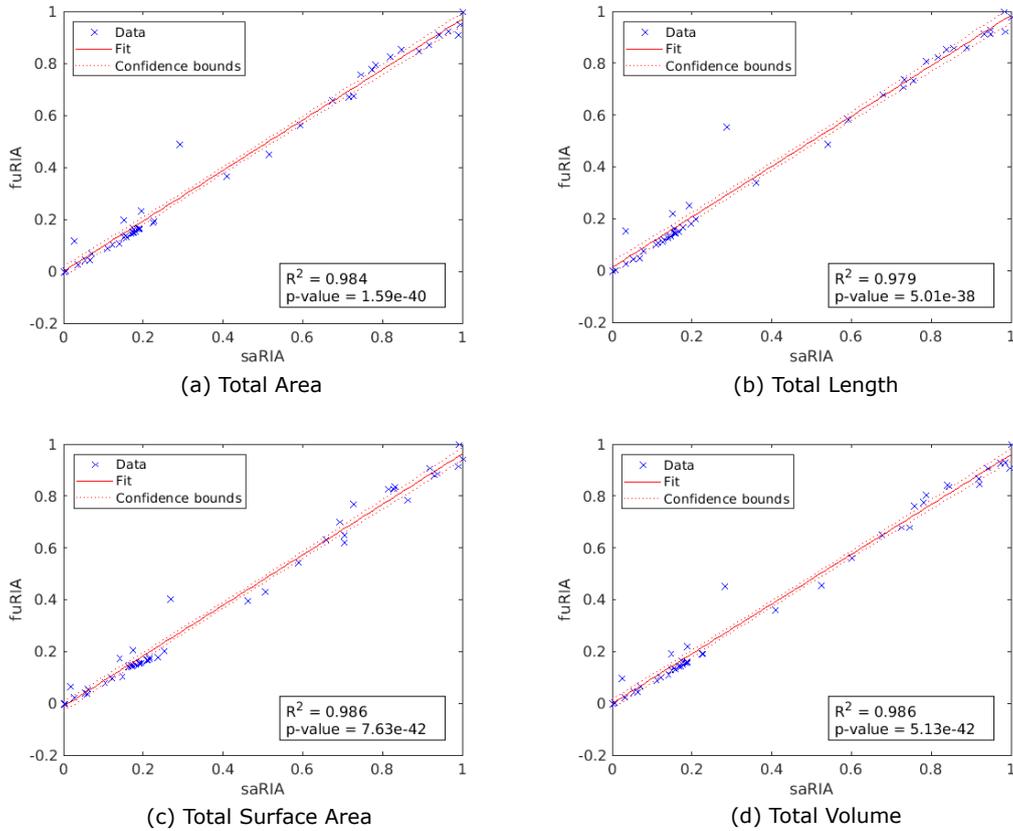


Figure 3.11: Correlation between root traits calculated using semi-automated saRIA (x-axis) and fully-automated faRIA (y-axis) image segmentation. Each point represents a trait value estimated from one of 40 soil-root images from rhizotron imaging system. The red colour solid line and dotted lines represent a fitted curve and 95% confidence bounds, respectively. The R^2 value indicates good conformity between saRIA and faRIA results of image segmentation and trait calculation.

to the fact that the SegRoot model transfers only max-pooling indices (i.e., location of feature maps) from encoder to decoder for feature concatenation and reconstruction, whereas our U-net model transfers complete feature map information (i.e., both location and pixel values) to the decoder. This leads to detection of both primary and secondary low contrast roots with the improved DC in comparison to the SegRoot, see Figure 3.5. However, more information required for U-net makes the decoder path expensive and requires more memory (9.47 MB) than the SegRoot (1.49 MB).

In addition to the faRIA:256 model, which was trained on 256x256 patches of original large root images, the performance of proposed U-net architecture was reformulated on full images and validated with images downscaled to the size of 1024x1024 due to our hardware limitations using the faRIA:1024 model. While both faRIA:1024 and faRIA:256 models demonstrated a comparable accuracy in the training stage, faRIA:256 exhibits more balanced performance between precision and recall than faRIA:1024. This imbalance is caused by the pixels of intermediate intensity on the boundary between the soil and root regions that correspond to average values calculated by downscaling. Pixels of intermediate intensities lead to false positive detection (Figure 3.5b). In particular, it is the case by segmentation of thin root structures in downscaled images using the faRIA:1024 model.

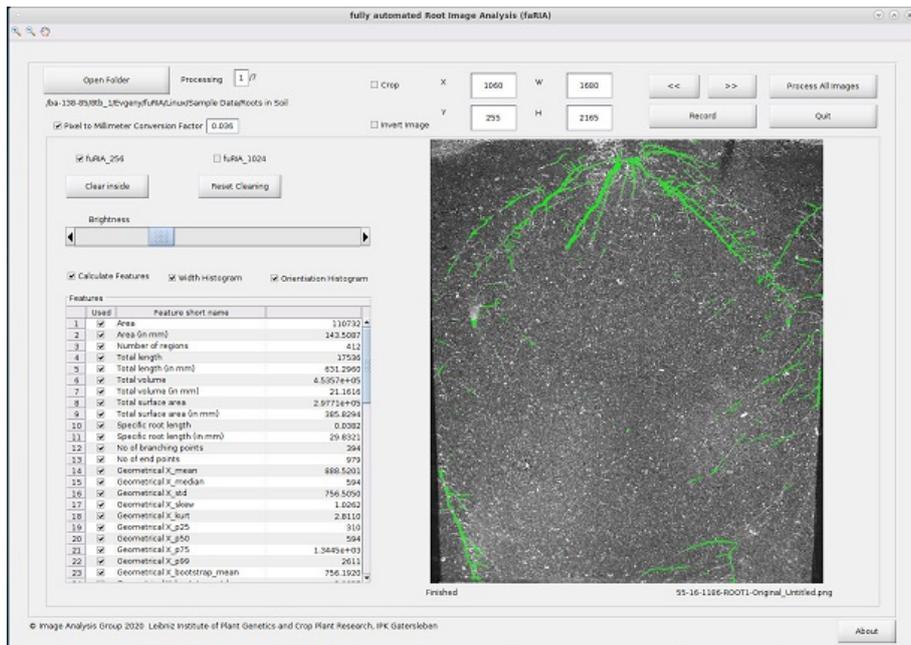


Figure 3.12: Graphical user interface of faRIA. Green colour pixels represent root regions automatically segmented by the U-net model.

Since roots and background regions exhibit similar structural properties in images of different modalities and plant species, our model originally trained on NIR maize roots images could also be applied to barley and arabidopsis roots from LED-rhizotron and UV imaging systems, respectively. For rhizotron images it achieved the minimum accuracy of 80% for all images with

exception of the image number 19 in Figure 3.7. The overall mean DC=0.85 indicates a fairly accurate segmentation of rhizotron images. The exceptional image with the number 19 exhibit low DC due to the presence of high intensity noise similar to root structures. Moreover, our model preserves the root thickness and continuity in the secondary roots compared to the SegRoot as shown in Figure 3.8 e, f. This results in DC of rhizotron image 0.87 is higher than the SegRoot 0.73.

The application of faRIA on UV images, the accuracy of the faRIA:256 model ranged between 60-83% with the mean DC=0.7, see Figure 3.9. A relatively low DC for some UV images is due to the presence of diverse artefacts including low contrast between the root architecture and heterogeneous soil regions, in-homogeneous scene illumination (i.e., vertical intensity gradient). This results in inaccurate segmentation (pink colour pixels) of low contrast structures and false detection of high intensity background structures as shown in Figure 3.10. However, faRIA:256 achieved the continuity in the root segmentation along the contrast varying root structures with DC of 0.80 (Figure 3.10 e) whereas SegRoot results in discontinues root structures with DC of 0.67 (Figure 3.10 f). Therefore, approximately 80% of the root pixels were correctly detected by faRIA:256 compared to the ground truth. Further examples of NIR, rhizotron and UV root image segmentation for juvenile or adult plants are in the Supplementary Information (see Figures S1-S6).

Furthermore, a direct comparison between phenotypic traits calculated with semi-automated (saRIA) and fully automated (faRIA) approaches shows a highly significant correlation which indicates that root image segmentation and phenotyping using faRIA as practically as good as human-supervised one.

Further, investigations with extended and/or augmented image data are required to improve the accuracy of segmentation of other root images that were not included in the original training set. On the other hand, it cannot be excluded that training of dedicated models with a narrow focus on a particular type of imaging modality and image structures could be a more reliable strategy to achieve more accurate results.

In conclusion, automated segmentation and analysis of a large amount of structurally heterogeneous and noisy soil-root images is a challenging task which solution is highly demanded in quantitative plant science. Here, we present an efficient GUI-based software tool for fully automated soil-root image segmentation which relies on the U-net CNN architecture trained on a set

of 6465 masks derived from 182 manually segmented soil-root images. The proposed algorithmic framework is capable to efficiently segment root structures of different size, shape and contrast with higher accuracy of DC=0.87 in comparison to the state-of-the-art solutions (SegRoot: DC=0.67). Our experimental results showed that the model trained with representative patches of root and background structures enables consideration of a larger amount ground truth data than original full-size images. Thereby, the faRIA:256 model trained on smaller size masks outperforms the larger mask model (faRIA1024) with respect to the overall precision and recall by comparison with ground truth data. In addition to NIR maize root images that were originally used for CNN model training, the faRIA tool can also be applied to other imaging modalities and plants species that exhibit similar structural properties of root and background regions. In addition to root image segmentation, faRIA calculates a number of useful phenotypic traits that in our experimental studies were shown to exhibit a significant correlation ($R^2 = 0.98$) with the ground truth traits. While the present CNN framework was predominantly trained with regular soil-root images, further investigations are required to address such challenging problems as segmentation of roots overlaid with a large scale noise (for example, due to water condensation) or filling artificial gaps in the root system that occur due to inhomogeneous scene illumination. Possible approaches to addressing these problems include, for example, appropriate augmentation of the training data set and/or alternative CNN models.

Acknowledgements

This work was performed within the German Plant-Phenotyping Network (DPPN) which is funded by the German Federal Ministry of Education and Research (BMBF) (project identification number: 031A053). M.H. was supported from European Regional Development Fund-Project "SINGING PLANT" (No. CZ.02.1.01/0.0/0.0/16_026/0008446).

Author contributions statement

N.N., M.H., and E.G. conceived, designed and performed the computational experiments, analysed the data, wrote the paper, prepared figures, and tables, and reviewed drafts of the paper. C.S. and A.J. executed the laboratory experiments, acquired image data, and reviewed drafts of the paper. J.O. supervised the computational data analysis and reviewed drafts of the paper.

T.A. co-conceptualized the project and reviewed drafts of the paper.

Declaration

All the protocols involving plants adhered to the ethical guidelines for plant usage were followed while conducting the experiments.

Supplementary Material

Supplementary Material

Fully-automated Root Image Analysis (faRIA), N. Narisetti*, M. Henke, C. Seiler, A. Junker, J. Ostermann, T. Altmann, E. Gladilin*.

Table S1. Root Traits: The description of estimated root system architecture traits in faRIA software.

Trait Name	Description
Area	Number of root pixels in the image
Number of Regions	Number of disconnected root objects in the image
Total Length	The sum of major axis length of each root object approximated by fitting ellipse to the root object
Total Volume (V)	<p>The sum of local volume at each root object of skeleton approximated by tubular shape whose average radius is estimated from image</p> $V = \sum_{i=0}^n \pi r_i^2$ <p>Where r_i is the average radius of i^{th} root component in the image.</p>
Total Surface Area (SA)	<p>The sum of surface area at each root object of skeleton approximated by tubular shape whose average radius is estimated from image</p> $SA = \sum_{i=0}^n 2\pi r_i$ <p>Where r_i is the average radius of i^{th} root component in the image.</p>
Specific Root Length	The ratio of total length and total volume of roots in the image.
Number of Branching Points	The total number of branches in the root skeleton
Number of End Points	The total number of end points in the root skeleton
Geometrical X_mean, Y_mean	The mean value of root pixels distribution in horizontal and vertical direction
Geometrical X_median, Y_median	The median value of root pixels distribution in horizontal and vertical direction
Geometrical X_std, Y_std	The standard deviation of root pixels distribution in horizontal and vertical direction

Geometrical X_skew, Y_skew	The skewness of root pixels distribution in horizontal and vertical direction
Geometrical X_kurt, Y_kurt	The kurtosis of root pixels distribution in horizontal and vertical direction
Geometrical X_p25, Y_p25	The 25 percentile of root pixels distribution in horizontal and vertical direction
Geometrical X_p50, Y_p50	The 50 percentile of root pixels distribution in horizontal and vertical direction
Geometrical X_p75, Y_p75	The 75 percentile of root pixels distribution in horizontal and vertical direction
Geometrical X_p99, Y_p99	The 99 percentile of root pixels distribution in horizontal and vertical direction
Geometrical X_bootstrap_mean, Y_bootstrap_mean	The mean value of bootstrapping root pixels distribution in horizontal and vertical direction
Geometrical X_bootstrap_stdev, Y_bootstrap_stdev	The median value of bootstrapping root pixels distribution in horizontal and vertical direction
Width mean	Average root diameter
Width median	Median root diameter
Width std	Standard deviation of the root diameter
Width skew	Skewness of root diameter
Width kurt	Kurtosis of root diameter
Width p25	25 percentile of root diameter
Width p51	50 percentile of root diameter
Width p75	75 percentile of root diameter
Width p99	99 percentile of root diameter
Width bootstrap_mean	The mean value of bootstrapping root pixels width
Width bootstrap_median	The median value of bootstrapping root pixels width
Orientation mean	Average root Orientation
Orientation median	Median root Orientation
Orientation std	Standard deviation of the root Orientation
Orientation skew	Skewness of root Orientation
Orientation kurt	Kurtosis of root Orientation
Orientation p25	25 percentile of root Orientation
Orientation p50	50 percentile of root Orientation
Orientation p75	75 percentile of root Orientation
Orientation p99	99 percentile of root Orientation
Orientation bootstrap_mean	The mean value of bootstrapping root pixels orientation

Orientation bootstrap_median	The median value of bootstrapping root pixels orientation
SeedAngle mean	Mean orientation of root pixels with respect to root originating seed point in RSA
SeedAngle median	Median orientation of root pixels with respect to root originating seed point in RSA
SeedAngle skew	Skewness of root pixels with respect to root originating seed point in RSA
SeedAngle kurt	Kurtosis of root pixels with respect to root originating seed point in RSA
SeedAngle p25	25 percentile of root pixels with respect to root originating seed point in RSA
SeedAngle p50	50 percentile of root pixels with respect to root originating seed point in RSA
SeedAngle p75	75 percentile of root pixels with respect to root originating seed point in RSA
SeedAngle p99	99 percentile of root pixels with respect to root originating seed point in RSA
SeedAngle bootstrap_mean	The mean value of bootstrapping root pixels seed orientation
SeedAngle bootstrap_stdev	The median value of bootstrapping root pixels orientation
ConvexHull area	Area of convex-hull of RSA
ConvexHull width	Width of convex-hull of RSA
ConvexHull height	Height of convex-hull of RSA
ConvexHull specific_area	The ratio of convex-hull area and actual area of RSA

Figure S1. Example of NIR root image segmentation for juvenile Maize plant using the faRIA:256 model.

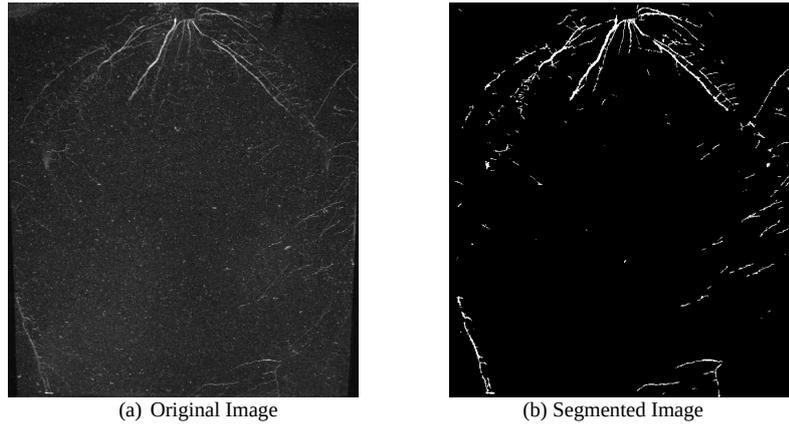


Figure S2. Example of NIR root image segmentation for adult Maize plant using the faRIA:256 model.

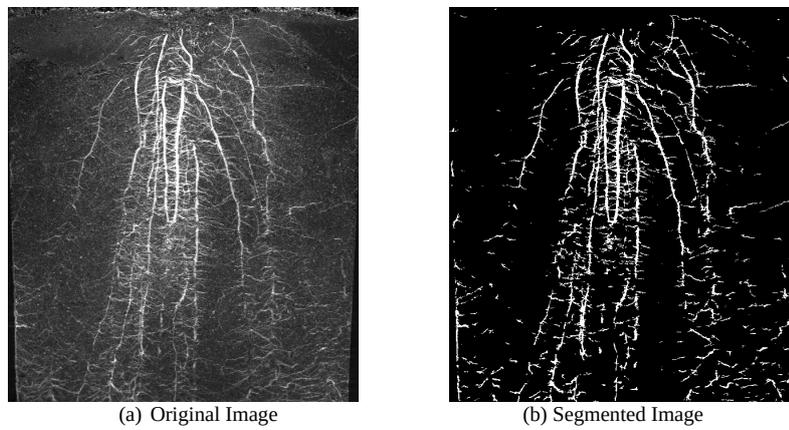


Figure S3. Example of Rhizotron root image segmentation for juvenile Barley plant using the faRIA:256 model.

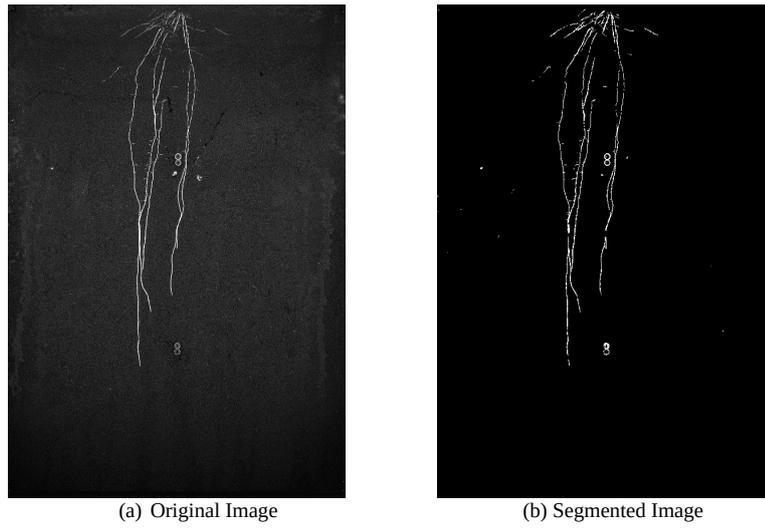


Figure S4. Example of Rhizotron root image segmentation for adult Barley plant using the faRIA:256 model.

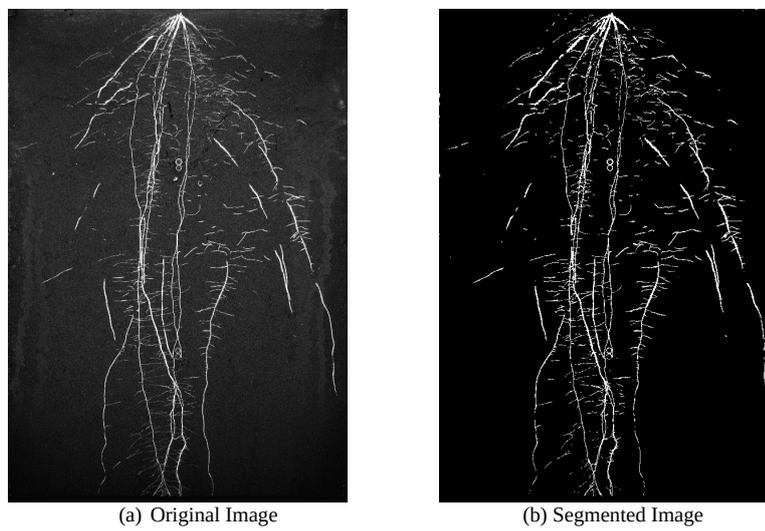


Figure S5. Example of UV root image segmentation for juvenile Arabidopsis plant using the faRIA:256 model.

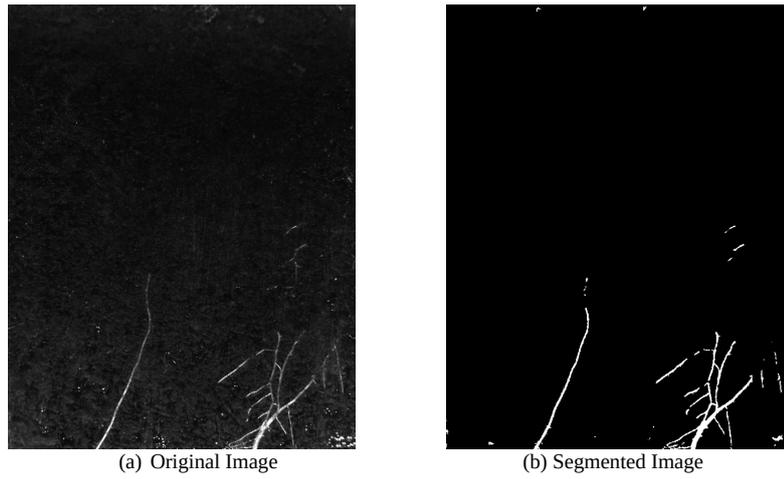
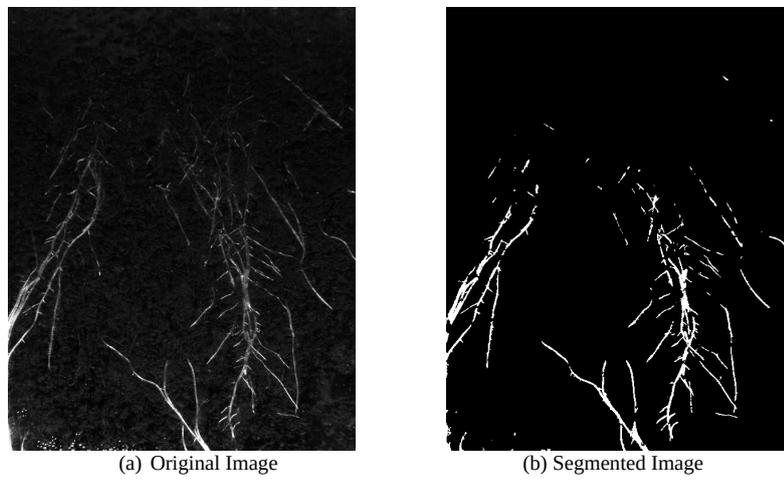


Figure S6. Example of UV root image segmentation for adult Arabidopsis plant using the faRIA:256 model.



Bibliography

- [1] Jonathan Lynch. Root architecture and plant productivity. *Plant physiology*, 109(1):7, 1995.
- [2] Anjali S. Iyer-Pascuzzi, Olga Symonova, Yuriy Mileyko, Yueling Hao, Heather Belcher, John Harer, Joshua S. Weitz, and Philip N. Benfey. Imaging and analysis platform for automatic phenotyping and trait ranking of plant root systems. *Plant Physiology*, 152(3):1148–1157, 2010.
- [3] Samuel Trachsel, Shawn M Kaeppler, Kathleen M Brown, and Jonathan P Lynch. Shovelomics: high throughput phenotyping of maize (*zea mays* l.) root architecture in the field. *Plant and soil*, 341(1-2):75–87, 2011.
- [4] AG Bengough and CE Mullins. Penetrometer resistance, root penetration resistance and root elongation rate in two sandy loam soils. *Plant and Soil*, 131(1):59–66, 1991.
- [5] T Wojciechowski, MJ Gooding, L Ramsay, and PJ Gregory. The effects of dwarfing genes on seedling root growth of wheat. *Journal of Experimental Botany*, 60(9):2565–2573, 2009.
- [6] M Watt, S Moosavi, Shaun Cameron Cunningham, JA Kirkegaard, GJ Rebetzke, and RA Richards. A rapid, controlled-environment seedling root screen for wheat correlates well with rooting depths at vegetative, but not reproductive, stages at two field sites. *Annals of Botany*, 112(2):447–455, 2013.
- [7] JS Perret, ME Al-Belushi, and M Deadman. Non-destructive visualization and quantification of roots using computed tomography. *Soil Biology and Biochemistry*, 39(2):391–399, 2007.
- [8] Saoirse R Tracy, Jeremy A Roberts, Colin R Black, Ann McNeill, Rob Davidson, and Sacha J Mooney. The x-factor: visualizing undisturbed

- root architecture in soils using x-ray computed tomography. *Journal of experimental botany*, 61(2):311–313, 2010.
- [9] Louise van der Weerd, Mireille MAE Claessens, Tom Ruttink, Frank J Vergeldt, Tjeerd J Schaafsma, and Henk Van As. Quantitative nmr microscopy of osmotic stress responses in maize and pearl millet. *Journal of Experimental Botany*, 52(365):2333–2343, 2001.
- [10] Suqin Fang, Xiaolong Yan, and Hong Liao. 3d reconstruction and dynamic modeling of root architecture in situ and its application to crop phosphorus research. *The Plant Journal*, 60(6):1096–1108, 2009.
- [11] Guang Zeng, Stanley T Birchfield, and Christina E Wells. Automatic discrimination of fine roots in minirhizotron images. *New Phytologist*, 177(2):549–557, 2008.
- [12] Mark G Johnson, David T Tingey, Donald L Phillips, and Marjorie J Storm. Advancing fine root research with minirhizotrons. *Environmental and Experimental Botany*, 45(3):263–289, 2001.
- [13] SC Van de Geijn, J Vos, J Groenwold, J Goudriaan, and PA Leffelaar. The wageningen rhizolab—a facility to study soil-root-shoot-atmosphere interactions in crops. *Plant and Soil*, 161(2):275–287, 1994.
- [14] Morris G Huck and Howard M Taylor. The rhizotron as a tool for root research. In *Advances in agronomy*, volume 35, pages 1–35. Elsevier, 1982.
- [15] Amram Eshel and Tom Beekman. *Plant roots: the hidden half*. CRC press, 2013.
- [16] Kerstin A Nagel, Alexander Putz, Frank Gilmer, Kathrin Heinz, Andreas Fischbach, Johannes Pfeifer, Marc Faget, Stephan Blossfeld, Michaela Ernst, Chryssa Dimaki, et al. Growscreen-rhizo is a novel phenotyping robot enabling simultaneous measurements of root and shoot growth for plants grown in soil-filled rhizotrons. *Functional Plant Biology*, 39(11):891–904, 2012.
- [17] Astrid Junker, Moses M. Muraya, Kathleen Weigelt-Fischer, Fernando Arana-Ceballos, Christian Klukas, Albrecht E. Melchinger, Rhonda C. Meyer, David Riewe, and Thomas Altmann. Optimizing experimental procedures for quantitative evaluation of crop plant performance in high throughput phenotyping systems. *Frontiers in Plant Science*, 5:770, 2015.

- [18] Rongli Shi, Astrid Junker, Christiane Seiler, and Thomas Altmann. Phenotyping roots in darkness: disturbance-free root imaging with near infrared illumination. *Functional Plant Biology*, 45(4):400–411, 2018.
- [19] Patrick Armengaud, Kevin Zambaux, Adrian Hills, Ronan Sulpice, Richard J Pattison, Michael R Blatt, and Anna Amtmann. Ez-rhizo: integrated software for the fast and accurate measurement of root system architecture. *The Plant Journal*, 57(5):945–956, 2009.
- [20] Jordon Pace, Nigel Lee, Hsiang Sing Naik, Baskar Ganapathysubramanian, and Thomas Lübberstedt. Analysis of maize (*zea mays* l.) seedling roots with the high-throughput image analysis tool aria (automatic root image analysis). *PLoS One*, 9(9):e108255, 2014.
- [21] Jacques Le Bot, Valérie Serra, José Fabre, Xavier Draye, Stéphane Adamowicz, and Loïc Pagès. Dart: a software to analyse root system architecture and development from captured images. *Plant and Soil*, 326(1-2):261–273, 2010.
- [22] J-L Arsenault, S Poulcur, C Messier, and R Guay. Winrhizo™, a root-measuring system with a unique overlap correction method. *HortScience*, 30(4):906D–906, 1995.
- [23] Thibaut Bontpart, Cristobal Concha, Mario Valerio Giuffrida, Ingrid Robertson, Kassahun Admkie, Tulu Degefu, Nigusie Girma, Kassahun Tesfaye, Teklehaimanot Haileselassie, Asnake Fikre, Masresha Fetene, Sotirios A. Tsiftaris, and Peter Doerner. Affordable and robust phenotyping framework to analyse root system architecture of soil-grown plants. *The Plant Journal*, 103(6):2330–2343, 2020.
- [24] Taras Galkovskyi, Yuriy Mileyko, Alexander Bucksch, Brad Moore, Olga Symonova, Charles A Price, Christopher N Topp, Anjali S Iyer-Pascuzzi, Paul R Zurek, Suqin Fang, et al. Gia roots: software for the high throughput analysis of plant root system architecture. *BMC plant biology*, 12(1):116, 2012.
- [25] Alain Pierret, Santimaitree Gonkhamdee, Christophe Jourdan, and Jean-Luc Maeght. Ij_rhizo: an open-source software to measure scanned images of root samples. *Plant and soil*, 373(1-2):531–539, 2013.
- [26] Narendra Narisetti, Michael Henke, Christiane Seiler, Rongli Shi, Astrid Junker, Thomas Altmann, and Evgeny Gladilin. Semi-automated root image analysis (saria). *Scientific reports*, 9(1):1–10, 2019.

- [27] Guillaume Lobet, Loïc Pagès, and Xavier Draye. A novel image-analysis toolbox enabling quantitative analysis of root system architecture. *Plant physiology*, 157(1):29–39, 2011.
- [28] Jinhai Cai, Zhanghui Zeng, Jason N Connor, Chun Yuan Huang, Vanessa Melino, Pankaj Kumar, and Stanley J Miklavcic. Rootgraph: a graphic optimization tool for automated image analysis of plant roots. *Journal of experimental botany*, 66(21):6551–6562, 2015.
- [29] Liang Zheng, Yi Yang, and Qi Tian. Sift meets cnn: A decade survey of instance retrieval. *IEEE transactions on pattern analysis and machine intelligence*, 40(5):1224–1244, 2017.
- [30] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.
- [31] W Bai, M Sinclair, G Tarroni, O Oktay, M Rajchl, G Vaillant, AM Lee, N Aung, E Lukaschuk, MM Sanghvi, F Zemrak, K Fung, JM Paiva, V Carapella, YJ Kim, H Suzuki, B Kainz, PM Matthews, SE Petersen, SK Piechnik, S Neubauer, B Glocker, and D Rueckert. Human-level cmr image analysis with deep fully convolutional networks. *ArXiv*, page 1710.09289, 05 2018.
- [32] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(12):2481–2495, 2017.
- [33] Dimitrios Marmanis, Jan D Wegner, Silvano Galliani, Konrad Schindler, Mihai Datcu, and Uwe Stilla. Semantic segmentation of aerial images with an ensemble of cnss. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2016*, 3:473–480, 2016.
- [34] Michael P Pound, Jonathan A Atkinson, Alexandra J Townsend, Michael H Wilson, Marcus Griffiths, Aaron S Jackson, Adrian Bulat, Georgios Tzimiropoulos, Darren M Wells, Erik H Murchie, et al. Deep machine learning provides state-of-the-art performance in image-based plant phenotyping. *Gigascience*, 6(10):gix083, 2017.
- [35] Clément Douarre, Richard Schielein, Carole Frindel, Stefan Gerth, and David Rousseau. Transfer learning from synthetic data applied to soil–

- root segmentation in x-ray tomography images. *Journal of Imaging*, 4(5):65, 2018.
- [36] Tanuj Misra, Alka Arora, Sudeep Marwaha, Viswanathan Chinnusamy, Atmakuri Ramakrishna Rao, Rajni Jain, Rabi Narayan Sahoo, Mrinmoy Ray, Sudhir Kumar, Dhandapani Raju, et al. Spikesegnet-a deep learning approach utilizing encoder-decoder network with hourglass for spike segmentation and counting in wheat plant from visual imaging. *Plant methods*, 16(1):1–20, 2020.
- [37] Renzhen Wang, Shilei Cao, Kai Ma, Yefeng Zheng, and Deyu Meng. Pairwise learning for medical image segmentation. *Medical Image Analysis*, 67:101876, 2021.
- [38] Neerav Karani, Ertunc Erdil, Krishna Chaitanya, and Ender Konukoglu. Test-time adaptable neural networks for robust medical image segmentation. *Medical Image Analysis*, 68:101907, 2021.
- [39] Salman Khan, Khan Muhammad, Tanveer Hussain, Javier Del Ser, Fabio Cuzzolin, Siddhartha Bhattacharyya, Zahid Akhtar, and Victor Hugo C. de Albuquerque. Deepsmoke: Deep learning model for smoke detection and segmentation in outdoor environments. *Expert Systems with Applications*, page 115125, 2021.
- [40] Yu Jiang and Changying Li. Convolutional neural networks for image-based high-throughput plant phenotyping: a review. *Plant Phenomics*, 2020, 2020.
- [41] Yezi Zhu, Marc Aoun, Marcel Krijn, Joaquin Vanschoren, and High Tech Campus. Data augmentation using conditional generative adversarial networks for leaf counting in arabidopsis plants. In *BMVC*, page 324, 2018.
- [42] Junjie Chen and Xinghua Shi. A sparse convolutional predictor with denoising autoencoders for phenotype prediction. In *Proceedings of the 10th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics*, pages 217–222, 2019.
- [43] Robail Yasrab, Jonathan A Atkinson, Darren M Wells, Andrew P French, Tony P Pridmore, and Michael P Pound. RootNav 2.0: Deep learning for automatic navigation of complex plant root architectures. *GigaScience*, 8(11), 11 2019. giz123.

- [44] Tao Wang, Mina Rostamza, Zhihang Song, Liangju Wang, G McNickle, Anjali S Iyer-Pascuzzi, Zhengjun Qiu, and Jian Jin. Segroot: a high throughput segmentation method for root image analysis. *Computers and Electronics in Agriculture*, 162:845–854, 2019.
- [45] Robail Yasrab, Michael P Pound, Andrew P French, and Tony P Pridmore. Rootnet: A convolutional neural networks for complex plant root phenotyping from high-definition datasets. *bioRxiv*, 2020.
- [46] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*, 2015.
- [47] Shibani Santurkar, Dimitris Tsipras, Andrew Ilyas, and Aleksander Madry. How does batch normalization help optimization?, 2019.
- [48] Xiang Li, Shuo Chen, Xiaolin Hu, and Jian Yang. Understanding the disharmony between dropout and batch normalization by variance shift, 2018.
- [49] Chao Peng, Xiangyu Zhang, Gang Yu, Guiming Luo, and Jian Sun. Large kernel matters – improve semantic segmentation by global convolutional network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.
- [50] R. R. Jha, G. Jaswal, D. Gupta, S. Saini, and A. Nigam. Pixisegnet: pixel-level iris segmentation network using convolutional encoder–decoder with stacked hourglass bottleneck. *IET Biometrics*, 9(1):11–24, 2020.
- [51] Forest Agostinelli, Matthew Hoffman, Peter Sadowski, and Pierre Baldi. Learning activation functions to improve deep neural networks. *arXiv preprint arXiv:1412.6830*, 2014.
- [52] Limin Wang, Sheng Guo, Weilin Huang, and Yu Qiao. Places205-vggnet models for scene recognition. *arXiv preprint arXiv:1508.01667*, 2015.
- [53] Rob A Dunne and Norm A Campbell. On the pairing of the softmax activation and cross-entropy penalty functions and the derivation of the softmax activation function. In *Proc. 8th Aust. Conf. on the Neural Networks, Melbourne*, volume 181, page 185. Citeseer, 1997.
- [54] Kelly H Zou, Simon K Warfield, Aditya Bharatha, Clare MC Tempany, Michael R Kaus, Steven J Haker, William M Wells III, Ferenc A Jolesz,

- and Ron Kikinis. Statistical validation of image segmentation quality based on a spatial overlap index1: scientific reports. *Academic radiology*, 11(2):178–189, 2004.
- [55] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Craig Chen Zhifeng Citro, Greg S Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467*, 2016.
- [56] Chunwei Tian, Yong Xu, and Wangmeng Zuo. Image denoising using deep cnn with batch renormalization. *Neural Networks*, 121:461–473, 2020.
- [57] Chunwei Tian, Lunke Fei, Wenxian Zheng, Yong Xu, Wangmeng Zuo, and Chia-Wen Lin. Deep learning on image denoising: An overview. *Neural Networks*, 131:251–275, 2020.
- [58] Stéfan van der Walt, S Chris Colbert, and Gael Varoquaux. The numpy array: a structure for efficient numerical computation. *Computing in science & engineering*, 13(2):22–30, 2011.
- [59] Stefan Van der Walt, Johannes L Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D Warner, Neil Yager, Emmanuelle Gouillart, and Tony Yu. scikit-image: image processing in python. *PeerJ*, 2:e453, 2014.
- [60] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [61] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.
- [62] Mathworks, matlab and statistics toolbox release 2019b, 2019. The MathWorks, Natick, MA, USA.
- [63] Alan C. Bovik. Chapter 3 - basic gray level image processing. In Al Bovik, editor, *The Essential Guide to Image Processing*, pages 43 – 68. Academic Press, Boston, 2009.

4. Deep Learning Based Greenhouse Image Segmentation and Shoot Phenotyping (DeepShoot)

Paper 3: Narisetti N, Henke M, Neumann K, Stolzenburg F, Altmann T and Gladilin E (2022) Deep Learning Based Greenhouse Image Segmentation and Shoot Phenotyping (DeepShoot). *Front. Plant Sci.* 13:906410. doi: 10.3389/fpls.2022.906410

Impact Factor: 6.6 (according to Clarivate releases the Journal Citation Report (JCR 2023))

Abstract

Background Automated analysis of large image data is highly demanded in high-throughput plant phenotyping. Due to large variability in optical plant appearance and experimental setups, advanced machine and deep learning techniques are required for automated detection and segmentation of plant structures in complex optical scenes.

Methods Here, we present a GUI-based software tool (DeepShoot) for efficient, fully automated segmentation and quantitative analysis of greenhouse-grown shoots which is based on pre-trained U-net deep learning models of arabidopsis, maize and wheat plant appearance in different rotational side- and top-views.

Results Our experimental results show that the developed algorithmic framework performs automated segmentation of side- and top-view images of different shoots acquired at different developmental stages using different phenotyping facilities with an average accuracy of more than 90% and outperforms shallow as well as conventional and encoder backbone networks in cross-validation tests with respect to both precision and performance time.

Conclusion The DeepShoot tool presented in this work provides an efficient solution for automated segmentation and phenotypic characterization of greenhouse-grown plant shoots suitable also for end-users without advanced IT skills. Primarily trained on images of three selected plants, this tool can be applied to images of other plant species exhibiting similar optical properties.

4.1 Introduction

Image-based high-throughput plant phenotyping became a method of choice in quantitative plant sciences aiming to reveal casual links between phenotypic and genomic plant traits under varying environmental conditions [1]. The ultimate goal is to make assessment of plant phenotypic traits data as efficient and scalable as genomic screening [2, 3]. However, efficient and accurate processing and analysis of large image data from different optical set-ups represents a challenging task constituting one of the major bottle necks in the pipeline of phenome-genome correlation [4].

The first critical step by quantitative analysis of plant image data is the image segmentation, which aims to classify all image pixels into two or more distinctive classes, e.g., foreground (plant) and background (non-plant) regions. Due to a number of natural and technical factors, segmentation of plant structures from background regions renders a challenging task. Inhomogeneous illumination, shadows, occlusions, reflections and dynamic optical appearance of growing plants complicate definition of invariant criterions for detection of different parts (e.g., leaves, flowers, fruits, spikes) of different plant types (e.g., arabidopsis, maize, wheat) at different developmental stages (e.g., juvenile, adult) in different views (e.g., top or multiple side views) [5]. Consequently, conventional methods that are typically based on some suitable image features and tailored to a particular data cannot be extended to new data in a straight forward manner. For example, one such popular approach to unsupervised image segmentation is based on analysis of differences between plant-containing and 'empty' reference images [6]. Thereby, it is assumed that background intensity/colours remain unchanged after plants were moved into photo chamber. However, due to shadows and reflections both background and plant regions change their optical appearance. Moreover, these changes are dynamically progressing in course of plant development. Consequently, an 'empty' background image does not provide an ideal reference for straightforward segmentation of plant structures. Mapping of RGB image onto alternative colour spaces such as HSV and/or L^*a^*b is known to be useful for separability of fore- and background colours [7, 8, 9]. However, it cannot completely solve the problem of overlapping plant and background colours.

To overcome the above limitations of uni-modal image analysis, a registration-classification approach to plant image segmentation was suggested in our previous work [10], which relies on pre-segmentation of plant regions in image modalities with higher fore-/background contrast, such as fluorescence

images, followed by their co-registration with low-contrast image modalities (e.g., visible light or near-infrared images). Since segmentation masks derived from one image modality does not perfectly match to another image modality, classification of plant and marginal background structures in masked image regions has to be subsequently performed using pre-trained intensity/colour models. In some rare cases of severe plant movements due to relocation of carriers from one photo chamber to another one, substantial differences between plant contours in two different image modalities can occur. Although, the registration-classification showed relatively high accuracy of final segmentation results, the principle requirement of high-contrast multimodal data as well as occasional movement artefacts limit its application to experiments where only one single image modality (typically visible light images) is acquired.

Numerous further supervised approaches to intensity-/colour-based plant image segmentation was proposed in the past. In [11], automated segmentation of arabidopsis top-view images using a super pixel- and random forest classification-based algorithm was presented. In this approach, pre-labelled masks were used to segment each plant from the multi-tray experiment. However, as many other colour-based models it is limited to a particular experimental setup and plant type. More recently, [12] proposed a neural network based shallow learning method for the segmentation of side view visible light images. This approach classifies each pixel based on neighbourhood pixel information of the trained ground truth data and outperforms conventional thresholding methods.

All the above state of the art techniques requires reference images, presence of particular image features and expertise in manual parameter tuning for each image to be segmented. Consequently, conventional supervised techniques are typically trained on and applied to particular types of plants, experimental set-ups and illumination scenes. However, high-throughput phenotyping of thousands and millions of plant images demands a fully automated, efficient and accurate segmentation algorithms with higher order cognitive abilities that can tolerate variation in scene illumination and plant/background colours.

In recent times, convolutional neural networks (CNNs) gained high attention especially in computer vision applications, because of the ability to directly extract and train relevant multi-level features from data without prior knowledge and human effort in feature design. CNNs have been shown to outperform conventional approaches when applied to many traditionally

difficult tasks of image analysis including pattern detection and object segmentation in biomedical images [13, 14], traffic scenes [15] and remote sensing [16]. In recent years, they were also used for high-throughput plant phenotyping such as detection of wheat roots grown in germination paper [17], segmentation of roots from soil in X-ray tomography [18] and segmentation of spikes in wheat plants [19]. However, most of these works present exemplary applications and/or computational frameworks that can hardly be handled by end-users without advanced programming skills.

The aim of this work is to develop an efficient and handy tool for automated shoot image segmentation and quantification for different plant types using a pre-trained deep CNN framework which could be used in a straight forward manner even by unskilled users. The GUI software tool (**DeepShot**) developed for this purpose relies on the U-net segmentation model from [13] which was trained on ground truth images of three different plants (arabidopsis, maize and barley) acquired from two different views (side, top) in different stages of their development. The article is structured as follows. First, we present our methodological framework including the proposed U-net based framework for shoot image segmentation, ground truth data generation as well as training and evaluation procedures. Then, the results of experimental investigations are presented including model performance by application to segmentation of test shoot images vs. alternative state-of-the-art solutions.

4.2 Materials and Methods

4.2.1 Image Data

The deep learning-based shoot image analysis tool (DeepShoot) is designed for automated segmentation and quantification of visible light (VIS) images of arabidopsis, maize and barley shoots acquired from greenhouse phenotyping experiments using LemnaTec-Scanalyzer3D high throughput plant phenotyping platforms (LemnaTec GmbH, Aachen, Germany). Figure 4.1 shows examples of arabidopsis, maize and barely images from three different LemnaTec phenotyping platforms tailored to screening of large, mid-size and small plants. All three phenotypic platforms have different designs of photo chambers, illumination, colours of background walls and camera resolutions ranging between 1-6 Mpx.



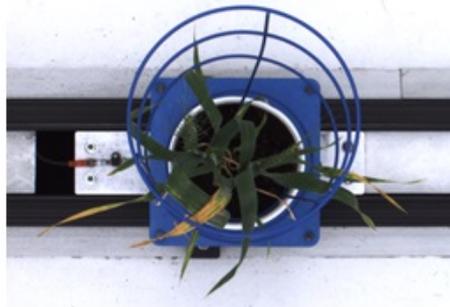
(a) Arabidopsis side view



(b) Arabidopsis top view



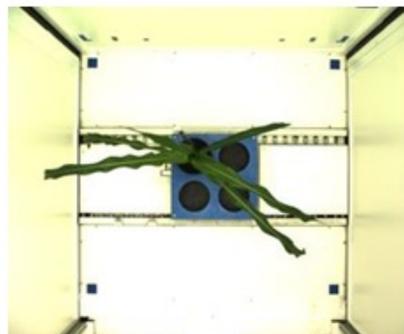
(c) Barley side view



(d) Barley top view



(e) Maize side view



(f) Maize top view

Figure 4.1: Examples of side- and top-view images of arabidopsis (a, b), barley (c, d) and maize (e, f) plants acquired with three different plant phenotyping platforms.

4.2.2 Ground Truth Generation

For training of CNN segmentation models, a representative set of ground truth images with an accurate annotation of fore- and background image regions is required. In this work, generation of ground truth images of different greenhouse cultured plants was performed using the GUI software tool (kmSeg) [9], which allows to efficiently annotate image regions by manual selection of pre-calculated k-means colour clusters corresponding to targeted plant structures. Background structures that exhibit similar colour fingerprint as plant regions and, thus, could not be separated by colour clustering are excluded or subsequently removed using manual region masking and cleaning likewise provided with the kmSeg tool. Semi-automated segmentation of a typical greenhouse image using kmSeg takes between 1-5 minutes depending on colour composition and structural complexity of a given plant shoot image.

4.2.3 Image Segmentation using CNN

The proposed CNN model is derived from the original encoder-decoder architecture of U-net [13], which provides a versatile framework for the semantic segmentation. In our model, batch normalization [20] is applied after each convolution layer in contrast to the original U-net. Because batch normalization improves the network performance and stability by normalizing the feature maps at respective levels [20, 21]. Furthermore, original U-net used dropout layers to remove outliers in the feature maps. But we avoided this layer because the combination of batch normalization and dropout layers can cause worse results [22]. Also, to improve the segmentation quality on largely connected patterns, a larger kernel size is considered in our approach compared to the original U-net [23]. Finally, our CNN model has less depth (of 3) compared to the original U-net depth of 4 due to the smaller input image size. The detailed comparison of convolutional parameters with respect to original U-net are summarized in Table 4.1.

Under consideration of the above suggested modifications, the U-net framework was adapted to the task of multimodal shoot image segmentation, see Figure 4.2. This network is designed in such way that training and testing are performed on patches of input images in original resolution. The advantage of this image masking approach is that it enables model training using large amount of ground truth data on consumer GPUs without losing high frequency information due to image downscaling. Furthermore, training of CNNs on image patches is more advantageous for learning local features

Table 4.1: Convolutional parameters of the original U-net and proposed modifications.

Convolutional parameters	Original U-net	Proposed modifications
Kernel size	3x3	7x7
Transposed kernel size	2x2	3x3
Stride	1x1	2x2
Padding	unpadded	padding with zeros
Depth	4	3
Number of filters	(64, 128, 256, 512, 1028)	(16, 32, 64, 128)

than full-size images [24]. Therefore, the input and output layers of the network are designed to operate on images of the size 256x256. Further details of network encoder and decoder layers are described below.

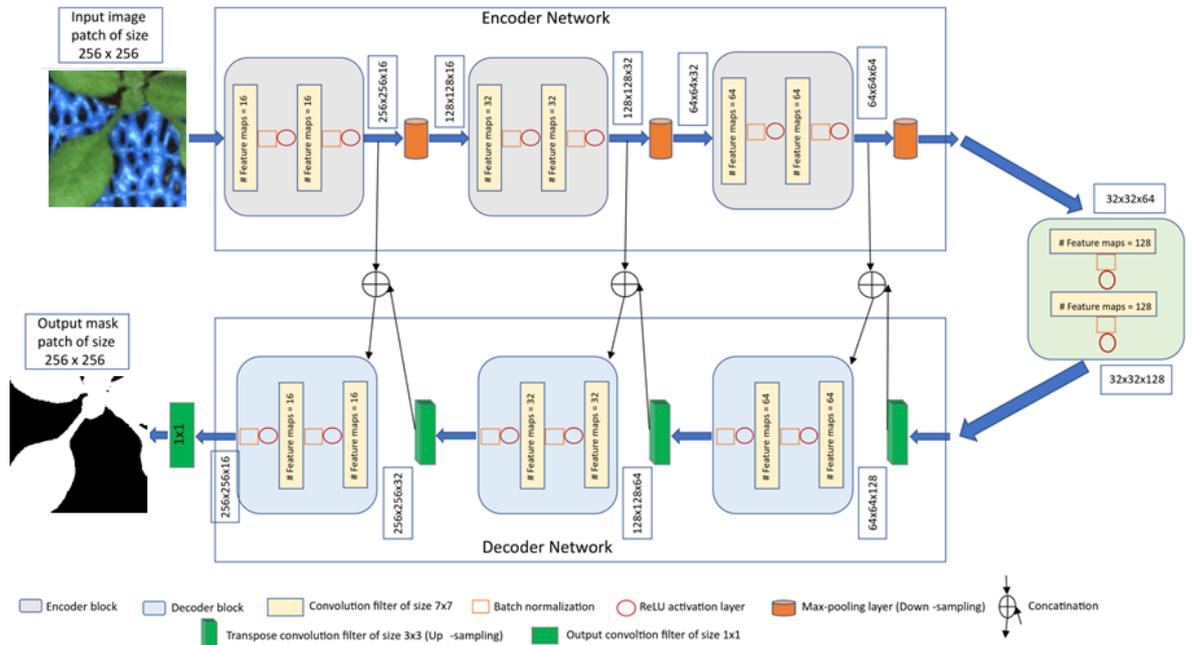


Figure 4.2: The proposed U-net architecture for shoot image segmentation.

Encoder network: The encoder network consists of 3 encoder blocks. The

first encoder block takes the image patches of size 256x256 as input and produces corresponding feature maps of size (256x256x16) as output. Then, the feature maps are forwarded to the second and third encoder blocks to generate further feature maps for the plant pixel detection. Each encoder block consists of two convolutional layers to learn feature maps at respective levels, where each convolutional layer consists of 7x7 convolution filter followed by batch normalization [20] and a non-linear activation function called Rectified Linear Unit (ReLU) [25]. Followed by each encoder block, max-pooling operation using general window size of 2x2 [24, 26] is applied for down-sampling the feature maps by half of its original size. The above steps enable a more efficient aggregation of image features. All three encoders are repeated with varying depth of 16, 32 and 64 to detect diverse plant features respectively.

Followed by encoder network, a bridge encoder block without max-pooling layer is applied. This results in 128 feature maps of the size 32x32.

Decoder network: The output from the bridge encoder (32x32x128) is up-sampled using 3x3 transpose convolution with the same padding and stride 2. This means size of feature maps (32x32x128) were doubled to (64x64x128) by applying filter of size 3x3 to all input elements and border elements were computed using zero padding. Then the resulting feature map is concatenated with the corresponding encoder feature maps. This results in feature maps of size (64x64x256) that are subsequently passed through a convolutional layer like encoder block but having decreasing channel depth of 64. This process is repeated for remaining decoder blocks with decreasing channel depth of 32 and 16. Finally, the output of the final decoder block is fed into a convolutional layer of size 1x1x1 with "Softmax" activation function [27] to classify each pixel as plant or non-plant at the patch level. The output of proposed architecture is a predicted mask of size 256x256 like the input image patch a shown in Figure 4.2.

4.2.4 Performance Metrics

To evaluate the performance of the proposed U-net model during training and testing stage, Dice coefficient (DC) [28] is used. It measures the area of intersection between the model and ground truth segmentation and its value ranges from 0 to 1, where 1 corresponds to 100% perfect and 0 to false segmentation. The Dice coefficient is defined as:

$$DC = \frac{2 * (P \cap G)}{P \cup G} = \frac{2 * \sum_i^N P_i G_i}{\sum_i^N P_i + \sum_i^N G_i} \quad (4.1)$$

where P and G are predicted and ground truth binary images respectively. P_i and G_i are output values 0 and 1 of pixel i in predicted and ground truth binary image respectively.

4.2.5 Computational Implementation

Training

The proposed U-net architecture was developed under Python 3.8 using TensorFlow [29] with Keras API. In addition, image processing operations like reading, cropping and training data preparation were done using PIL, Numpy [30] and Scikit-Image [31] packages. Then the proposed model was trained on GPU machine with Linux operating system (Intel(R) Core (TM) i7-10700K CPU @ 3.80GHz) and NVIDIA RTX 3090 - 24GB graphic card. As discussed above, the model is designed in such way that training will be performed on patches of original image. So, to generate non-overlapping patches of size 256x256, original images were padded with zeros at the image edges so that its width and height are divisible by 256. Out of these non-overlapping patches, both plant and background masks are considered in equal proportion to avoid potential imbalance between plant and non-plant training masks. Then each cropped mask is normalized in the range of [0, 1] for feature consistency in the CNN network. The overview of prepared training data of arabidopsis, barley and maize and their growth stages are described in Table 4.2 and 4.3 respectively. Regarding information on growth stages, an approximately equal number of images from different developmental stages (early, mid and late developmental phases) were analysed in this study.

Table 4.2: Overview of training data including images of arabidopsis, barley and maize plants.

Plant	Original resolution	Side view		Top view	
		#of images	#of masks	#of images	#of masks
Arabidopsis	2056x2454	197	17730	193	17730
Barley	1234x1624	100	3395	100	1908
Maize	2056x2454	100	3669	55	1036

Table 4.3: Overview of growth stages of training data from arabidopsis, barley and maize plants.

Plant/View	# of images	Growth stage		
		Early	Mid	Late
Arabidopsis/Side	197	62	65	70
Arabidopsis/Top	193	62	65	66
Barley/Side	100	30	35	35
Barley/Top	100	30	35	35
Maize/Side	100	30	35	35
Maize/Top	50	20	20	15

Subsequently, based on our experience and previous works [32, 33], the above prepared data sets were partitioned into training and validation in the ratio of 85:15 respectively. The initial weights of the proposed model were defined randomly with zero mean and standard deviation of 0.05 as proposed by *Krizhevsky et al.* [34]. Then the model was optimized with Adam optimizer [35] to improve the segmentation performance on training data sets. Binary cross-entropy loss function [36] is used to measure the unhappiness of the model during training and it defines the difference between predicted output and ground truth generated by the kmSeg tool as described above. This function compares each pixel prediction (0: non-plant, 1: plant) with ground truth pixel and averages all pixels loss for computing total loss of the image. Therefore, each pixel contributes to the overall objective loss function. Then the model was trained for 100 epochs with 16 number of convolutional channel features and batch size of 128 as per system constraints. The learning rate alters the magnitude of the updates to the model weights during each iteration and initialized with 0.001. Then a learning rate scheduler was used to dynamically reduce the learning rate by a factor of 0.2 if the validation loss is not improved in the next 5 iterations. Because a large learning rate will update weights too large which result performance of model will oscillate over training epochs. Whereas a too small learning rate may never converge and get stuck on suboptimal solution [37]. Here, note that all data sets (arabidopsis, barley and maize) were trained in similar way with same parameter configuration.

Prediction

As stated above, original shoot images have variational resolution where as proposed model requires input images of the size 256x256. So, during the prediction stage, original images are padded with zeros then non-overlapping 256x256 masks were generated similar as it was done in the training stage. The model does predictions on these 256x256 masks then they are combined to a single output image as shown in Figure 4.3. This process is dynamic, that means any image with resolution greater than 256x256 can be segmented in an automated manner.

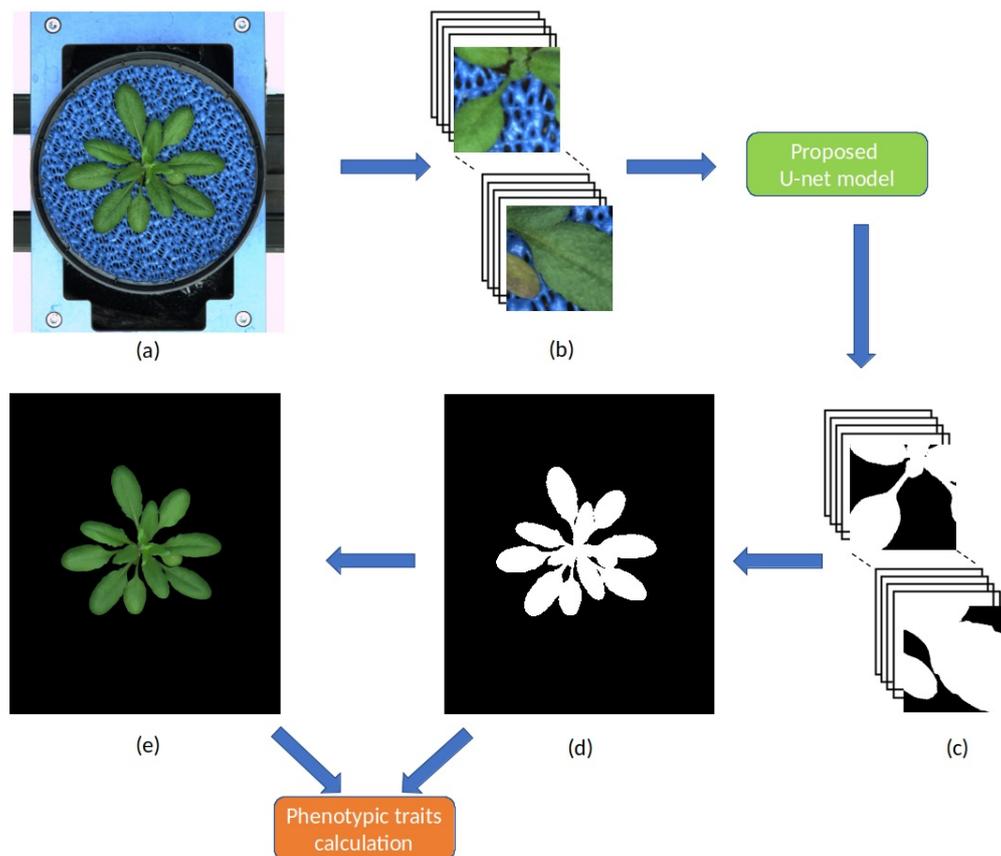


Figure 4.3: Workflow of the pipeline for image processing and segmentation in the DeepShoot tool. Green and orange colour boxes represents the operations of image segmentation and trait calculation: (a) original image, (b) original image patches of size 256x256, (c) segmented image patches of size 256x256, (d) binary segmentation of original image, (e) RGB colour space of (d).

Since the output layer of the model is a Sigmoid activation function, the predicted segmentation is a probability map with values ranging in between 0 and 1. Hence, this probability map is converted to a binary image using threshold 'T'. Here, relatively, plant pixels will have high probabilities compared to the background pixels. Therefore, $T \geq 0.6$ is chosen to consider all high probability pixels as plant pixels in the final segmentation. After fully automated segmentation, phenotypic traits of segmented plant structures were calculated in the final step.

Graphical User Interface

In practice, end-users prefer to have an easy-to-use software solution with a Graphical User Interface (GUI). Therefore, a user-friendly GUI front-end was developed under the MATLAB 2021a environment [38] to comfortably operate the complex algorithmic framework of shoot segmentation software. Figure 4.3 shows the complete workflow involved in shoot tool for automated plant segmentation and trait extraction. For import of deep learning models trained under Python the MATLAB interoperability routine *importKerasNetwork* [38] was used. According to specification of this function, the U-net models trained in Python were exported in the so-called h5 file format, which is supported by the recent versions of MATLAB including 2021a.

4.2.6 Method Comparison

The performance of our proposed model is compared with the recently published shallow learning based neural network (NN) by Adams et al. [12] which was developed and evaluated for the same application as ours, namely, segmentation of greenhouse shoot images. This algorithm classifies each pixel based on 3x3 neighbourhood information from red, green and blue channels using fully-connected neural networks. In this study, the same NN model architecture was retrained on our image data set with a large number of neighbourhoods features 5,939,562 and higher batch size of 4096 compared to the original study of 51,353 and 1024 respectively.

In addition, proposed encoder backbone of the U-net architecture is compared with different encoder backbones include vgg19 [39], resnet50 [40] and xception [41]. These models were trained on same image data set with similar training configuration except increased number of filters (64, 128, 256, 512) as discussed in subsection 4.2.5.

4.3 Results

4.3.1 Training and Validation

As described above, the proposed network was trained and validated on six different data sets including arabidopsis, barley and maize images acquired from three different plant phenotyping facilities. Thereby, each these three data sets were subdivided into training and validation sets in the ratio of 85:15, respectively. The model performance is analysed using binary cross-entropy loss (CE loss) and Dice coefficient at each epoch during the learning stage of the network. Because of dynamic optical appearance of growing plants, segmentation of shoot regions in side view images represents a more difficult task. This results in discontinuous shoot structures in segmented images. Therefore, it is important to give equal weights to errors related to both background and plant pixels in this study using Dice coefficient. Figure 4.4 shows the training and validation performance of the proposed model on six different data sets over 100 epochs. It shows that the training loss of six models was minimized and platen the curve after epoch number 60. Simultaneously, training DC was maximized and achieved more than 90% of the accuracy for all models by the end of the training epochs. In turn, generalized performance of the model is measured using validation measurements. Likewise training performance, validation DC also achieved more than 90% accuracy with low value of loss for all models at the end of the epochs. The brief overview of training and validation measurements are shown in Table 4.4.

Table 4.4: Training and validation performance of the shoot models.

Plant	Training		Validation	
	CE loss	DC	CE loss	DC
Arabidopsis side view	0.00075	0.9821	0.00083	0.9707
Arabidopsis top view	0.00297	0.9907	0.00345	0.9846
Barley side view	0.01172	0.9737	0.01300	0.9589
Barley top view	0.02986	0.9593	0.03366	0.9551
Maize side view	0.00687	0.9834	0.00906	0.9731
Maize top view	0.05433	0.9742	0.05673	0.9671

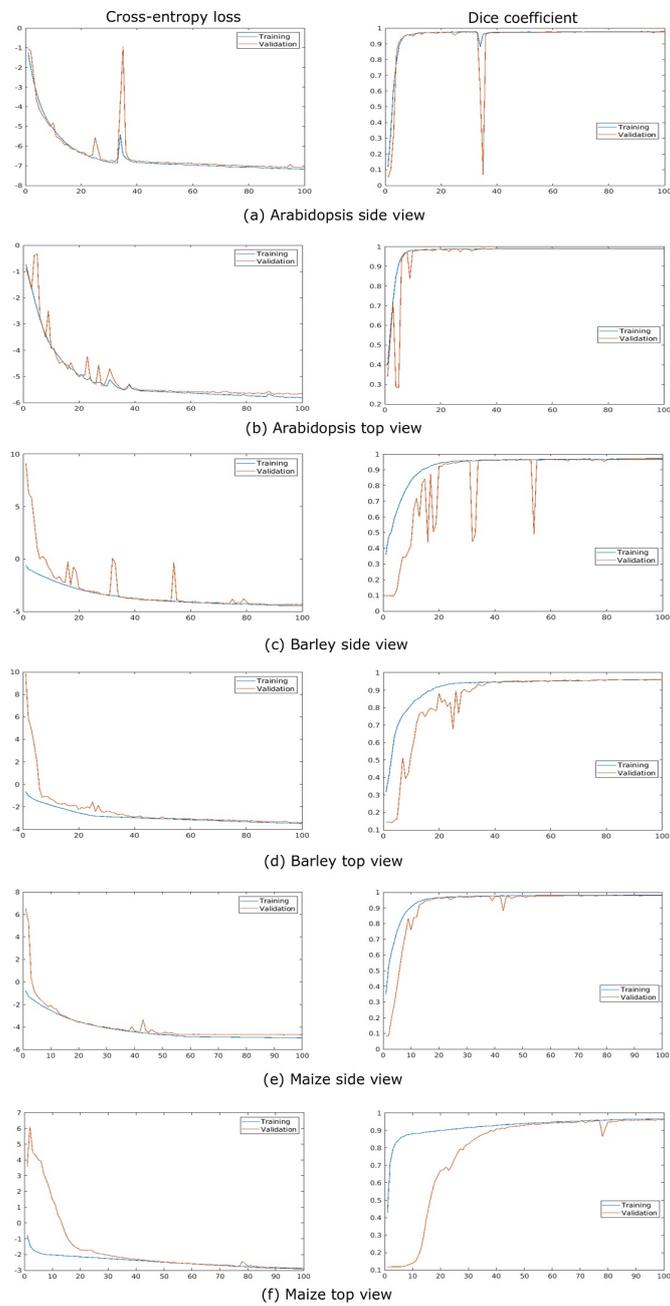


Figure 4.4: Training and validation performance of the shoot model over 100 epochs. X- and Y-axes represent the epoch number and performance measure, respectively. For visualization purpose, logarithmic cross-entropy values are plotted for all models.

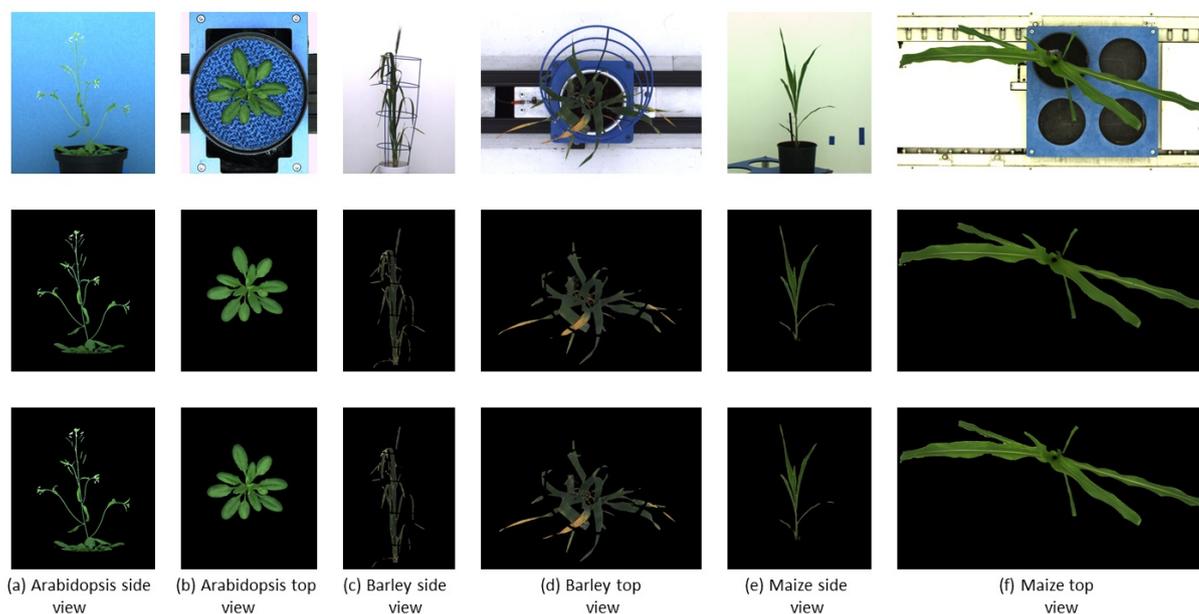


Figure 4.5: Segmentation performance: first, second and third row represents the original RGB image, ground truth segmentation by the kmSeg tool and predicted segmentation by shoot tool, respectively. The DC of each image as following: (a) 0.9117 (b) 0.9876 (c) 0.9384 (d) 0.9617 (e) 0.9709 (f) 0.9843.

In addition to the training performance, the exemplary segmentation of all models on test images is shown in Figure 4.5. It turns out that all models performed with a relatively higher DC of 0.95 except of arabidopsis side-view model which has the DC of 0.9117 compared to the ground truth generated by the kmSeg tool. Furthermore, trained models were tested on variational data sets from arabidopsis top-view like stress and multi-tray experiments as shown in Figure 4.6. Here, the model resulted in DC of 0.9664 and 0.9873 for stress and multi-tray experiments image compared to the ground truth respectively.

4.3.2 Evaluation on the Reference Dataset

To measure the performance of the model on unseen data, our CNN model trained on arabidopsis top-view images from LemnaTec-Scanalyzer3D was applied to the set of arabidopsis top-view from Scharr et al. [42]. This dataset was frequently used for CNN model training and evaluation in a number of previous works within the scope of CVPPP competitions (<https://www.plant-phenotyping.org/CVPPP2018>, <https://www.plant-phenotyping.org/>

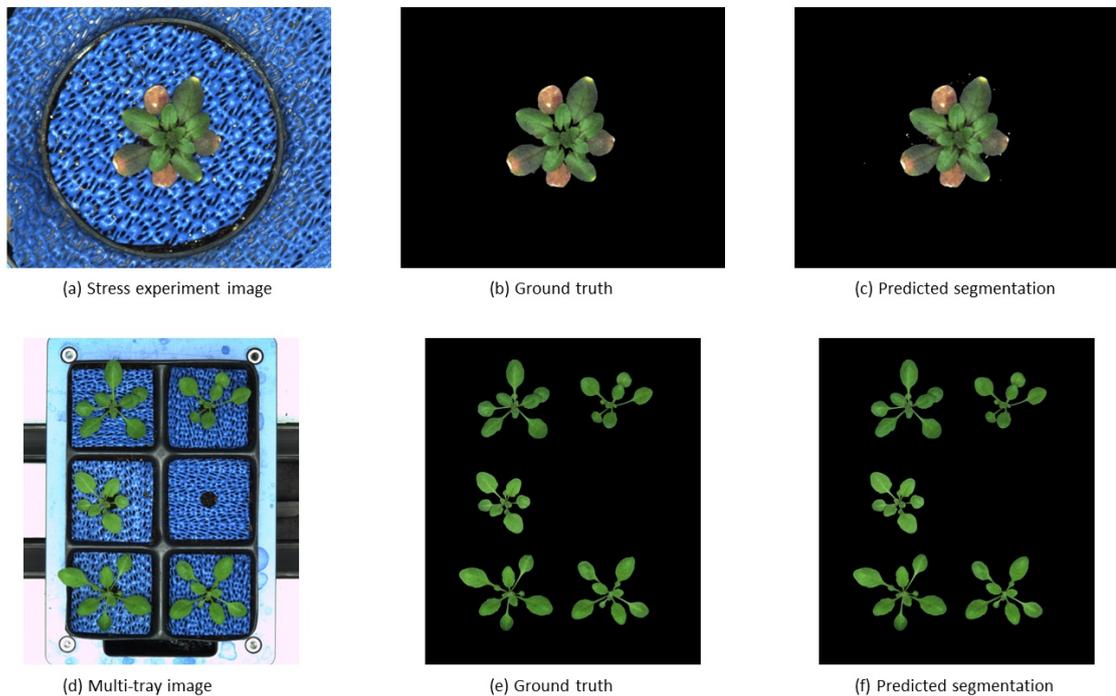


Figure 4.6: Segmentation performance on variational data sets: DC of stress and multi-tray experiments are 0.9664 and 0.9873, respectively.

CVPPP2019 <https://www.plant-phenotyping.org/CVPPP2020>). However, here it is only used for cross-validation of our model trained on images from our phenotyping facility. Figure 4.7 shows the mean DC of single and multi-tray experiments from Scharr et al. data set. The model resulted in the mean DC of 0.93 over 100 images and 0.95 over 27 images for single and multi-tray experiments respectively. Examples of segmentation of single-tray images from the references data set are shown in Figure 4.8.

4.3.3 Evaluation of DeepShoot vs. Alternative Solutions

The proposed U-net was compared with the recently published shallow learning based neural network (NN) by Adams et al. [12] which was originally developed and evaluated for shoot side view image segmentation. Figure 4.9 shows the comparative analysis of 17, 25 and 20 side view images of Arabidopsis, barley and maize plants, respectively. It briefs that proposed U-net outperforms with $DC > 0.9$ for all images where as neural networks predic-

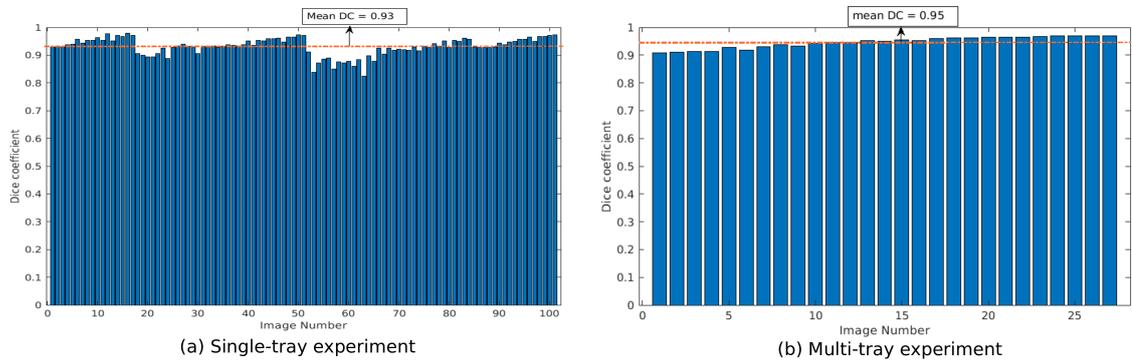


Figure 4.7: Evaluation of image segmentation on the references data set from Scharr et al.: Dice coefficient of arabidopsis top-view model over 100 and 27 images for single- and multi-tray experiments, respectively. The dotted orange line represents the mean DC value.

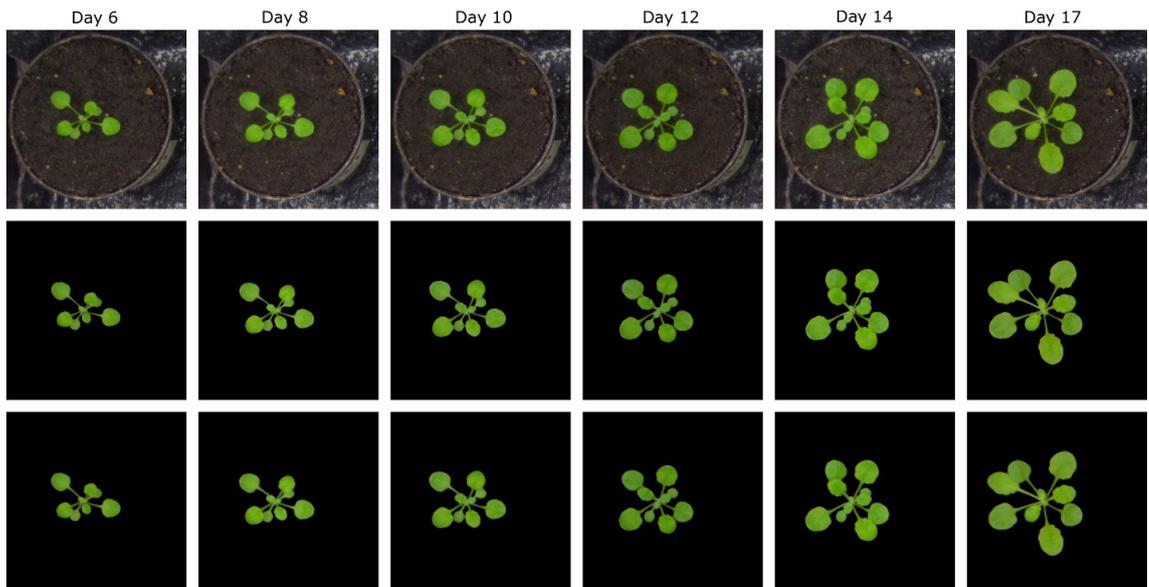


Figure 4.8: Examples of segmentation of arabidopsis top-view images from Scharr et al. All images were segmented with DC over 0.9.

tions have DC in between 0.5 and 0.8. An exemplary segmentation of three plants using neural network and proposed U-net with respect to ground truth is shown in Figure 4.10. Also, the computational time of both segmentation models required for the prediction on Intel(R) Xeon(R) Gold CPU @2.10 GHz with 20 CPU cores is listed in Table 4.5.

Furthermore, a comparison of different encoder backbones (vgg19, resnet50

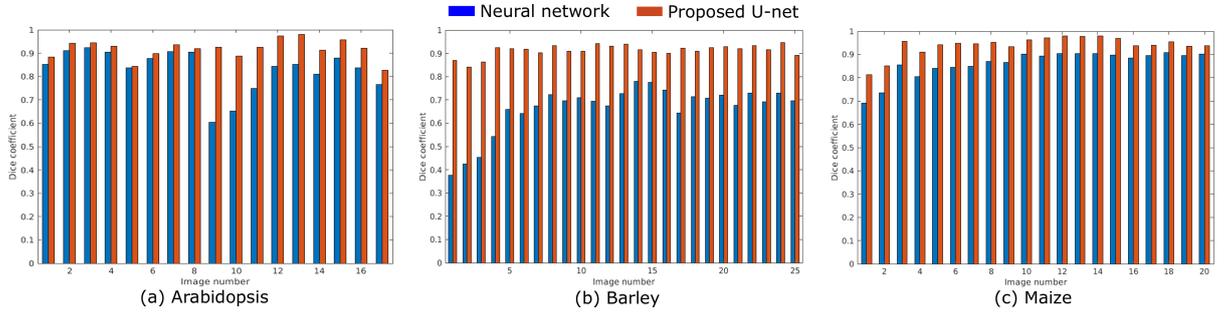


Figure 4.9: Performance of neural network (blue) and proposed U-net (orange) segmentation models on (a) 17 arabidopsis (b) 25 barley (c) 20 maize side view images.

Table 4.5: Computational time of shoot segmentation algorithms in seconds per image on a system with Intel(R) Xeon(R) Gold 6130 CPU @2.10GHz with 20 CPU cores.

Algorithm	Arabidopsis (2056x2454)	Barley (1234x1624)	Maize (2056x2454)
Neural networks	20410	13100	20410
Proposed U-net	7.2	3.05	7.2

and xception) of the U-net architecture was performed. Figure 4.11 shows the performance of alternative U-net backbones by training on arabidopsis top view images. It shows that both resnet50 and xception networks have higher validation loss (> 0.004) and it increases over number of iterations. On the other hand, vgg19 and proposed U-net are promising comparable performance with lower validation loss of 0.0033. In addition, the complexity of alternative U-net models with different encoder backbones on arabidopsis top view images is shown in Table 4.6.

4.3.4 DeepShoot GUI Tool

Figure 4.12 shows the GUI of DeepShoot software which is freely available as a precompiled executable program from <https://ag-ba.ipk-gatersleben.de/ds.html>. In addition to automated image segmentation, DeepShoooot calculates 35 shoot traits that are categorized into 4 feature groups (i.e. area, bounding

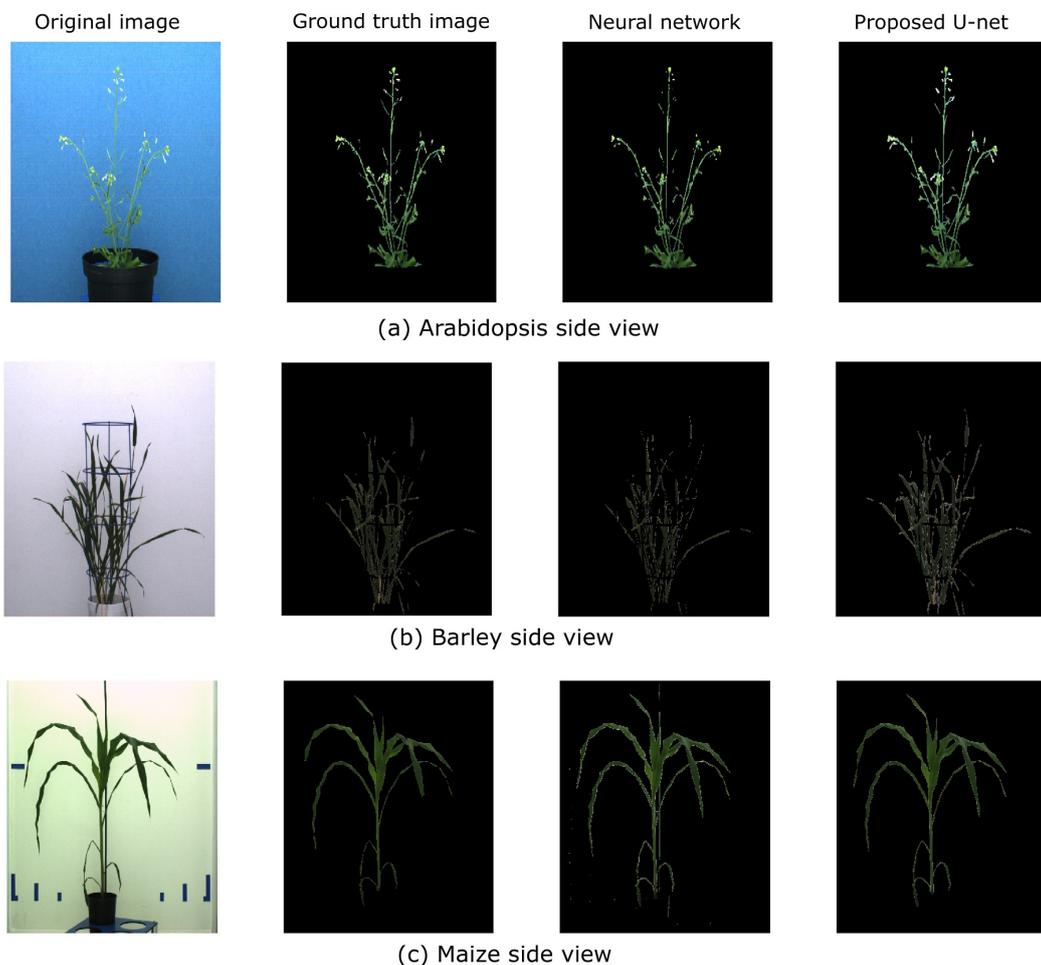


Figure 4.10: Evaluation of neural network segmentation with respect to the proposed U-net on arabidopsis, barley and maize side view images. (a) NN DC: 0.7824, DeepShoot DC: 0.8342 (b) NN DC: 0.6973, DeepShoot DC: 0.8924 (c) NN DC: 0.8746, DeepShoot DC: 0.9360.

box traits, convex-hull area and statistical colour features). Further information on definition of traits can be found in the Supplementary Table S1 accompanying this paper.

In order to restrict the analysis to the region of interest (ROI), users can define a custom ROI as a rectangle or polynomial shape using crop or *Clear outside* buttons of the DeepShoot GUI. The DeepShoot tool can be applied for analysis single images in a step-by-step manner or for automated processing of all images in a selected folder. Regarding DeepShoot time performance, image segmentation and traits calculation all together take in

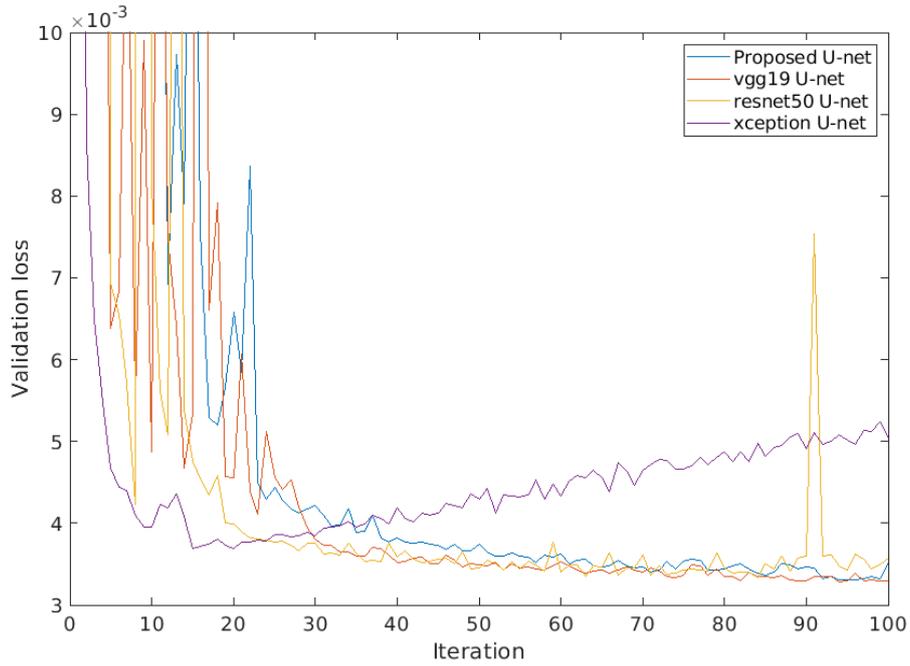


Figure 4.11: Loss performance of alternative U-net models with different encoder backbones on arabidopsis top view images.

Table 4.6: Complexity of alternative U-net models with different encoder backbones on arabidopsis top view images.

Parameter	Proposed U-net	vgg19	resnet50	xception
#of trainable parameters	2,484,721	24,780,993	15,053,121	5,583,065

average 18.5 seconds to process and analyse a 5-megapixel image on a system with Intel(R) Xeon(R) Gold 6130 CPU @2.10GHz with 20 CPU cores.

4.4 Discussion & Conclusion

Automated processing and quantitative analysis of a large amount of phenotypic image data represents a critical point determining the efficiency and accuracy of trait computation. The deep learning-based tool for automated

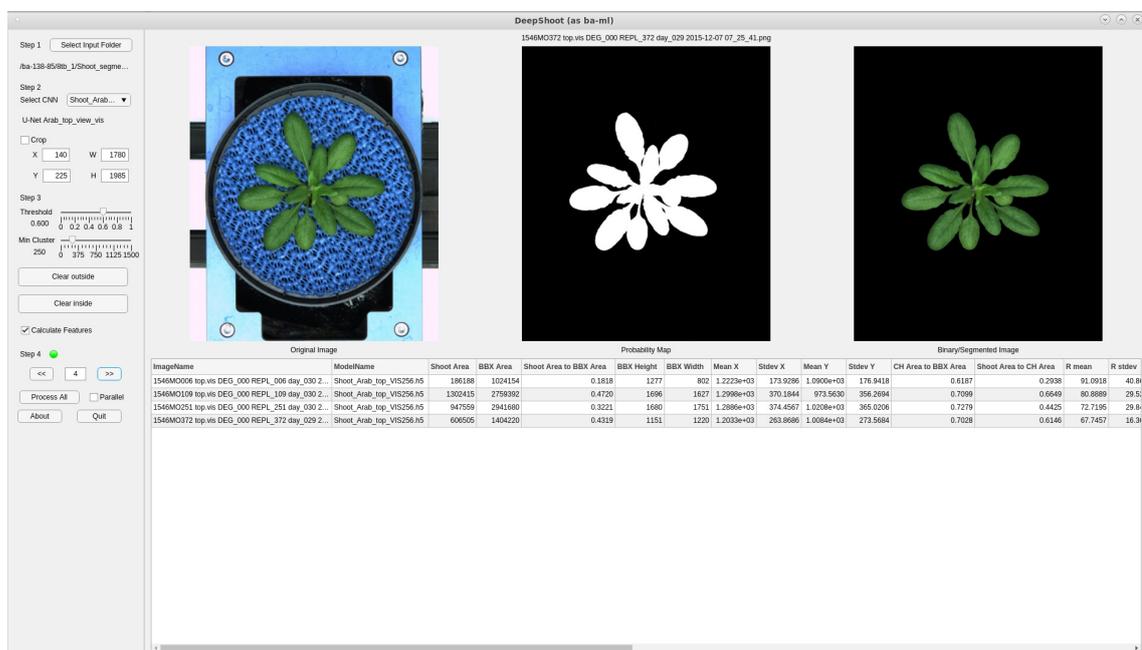


Figure 4.12: Graphical User Interface of the DeepShoot tool: left, middle and right images represent the original image, predicted probability map and predicted colour image, respectively.

shoot image segmentation and phenotypic analysis developed in this work aims to address this challenging task. Our experimental tests on three different plant types (arabidopsis, barley and maize) and two different views (side and top) showed that performance of the model during the training is improved over the number of iterations. On the other hand, model trained (for all plants) before iteration number 40 was under performed and showed worse performance for model validation. However, due to the dynamic reduction in learning rate by a factor of 0.2 a stable performance with more than 90% Dice coefficient for all shoot models was achieved. Additional information on impact of learning rate can be found in Supplementary Figure S1 accompanied with this paper. Moreover, arabidopsis and maize models achieved low CE loss values where as barley models have slightly higher CE loss values due to the variational leaves like yellow and brown colour leaves. This is reflected in lower DC of barley side- and top-view test images (0.9384 and 0.9617) compared to the arabidopsis top-view and maize models (> 0.97). Also, the trained model exhibited a low value of DC (0.9117) for arabidopsis side view test image compared to the other models due to the low contrast of secondary stems which have intensity similar to the background pixels.

In addition, the trained arabidopsis top-view model is validated on reference data sets including examples of stressed and multi-tray experiments. Our experimental results showed that model achieved a remarkably high DC of 0.9664 (stressed plants) and 0.9873 (multi-tray images) on these unseen data. However, small background noisy objects which have intensity and patterns similar to the leaves require additional application morphological operations (e.g., min cluster size) that are also available with the DeepShoot GUI tool. Furthermore, the model achieved very high DC (> 0.9) especially on untrained images with different background from Scharr et al. dataset. Overall our results indicate that the CNN model trained a particular set of images can also be applied to unseen data exhibiting similar plant shoot patterns but different background regions.

The performance of proposed U-net was compared with the shallow learning neural networks. Thereby, it was shown that most of the arabidopsis and maize images have relatively low discrepancy between predicted DC of both algorithms, because these images contain, mostly, high contrast green colour pixels for the target structures. In contrast, the shallow neural network exhibited a significantly lower DC on barley images. We draw the observation back to the fact that barley plants have more variable colour fingerprint including brown and yellow leaves. This shows that neural network is only capable of segmenting high contrast shoot structures whereas the U-net model is capable of segmenting both high contrast and colour-altering shoot structures. Because CNN frameworks are capable of generating multi-level features including neighbourhood information, colour, spatial patterns and textural features compared to shallow learning methods where only neighbourhood information was calculated. Therefore, rich in information makes DeepShoot outperforms the shallow networks. Furthermore, tests of computational performance of the shallow neural network vs. the proposed U-net model demonstrated the superior performance of the latter. In summary, the DeepShoot tool enables users to perform segmentation and analysis of plant shoot images faster and more accurate in comparison to the shallow neural network.

Furthermore, performance of proposed U-net model is compared with vgg19, resnet50 and xception encoder backbones. Thereby, it was observed that lower depth architecture vgg19 achieved better results in comparison to deep depth architectures such as resnet50 and xception that tend to overfitting. This can be attributed to higher complexity of these multi-layer networks that generate too many redundant features. However, vgg19 model still contains a large number of convolution layers with trainable weights

which makes it 10 times larger in size than our proposed U-net. Therefore, our proposed model achieves optimum results at the lower level of complexity which enables to perform high-throughput plant phenotyping on both lower and higher hardware configuration systems in real time.

It is known that U-net captures not only colour but also spatial pattern information. From this perspective, one can expect larger segmentation errors by application of DeepShoot to optical scenes strongly deviating from plant and background structures used by our model training. Nevertheless, our tests with unseen shoot images indicated that the present CNN framework can also be applied to analysis of quite different optical scenes or field-like images as long as the target plant structures are optically somewhat similar to images used in our training sets. Users are free to try and evaluate the performance of provided segmentation models on their particular images. From that perspective there are no other restrictions as the requirement of RGB image with the size $\geq 256 \times 256$.

Moreover, segmentation of thin or twisted leaves, flowers as well as shadowed or light-reflecting regions (such as metallic surfaces) is more prone to misclassification, which in turn may lead to fracturing of targeted structures or false-segmented background regions. Nevertheless, improvements of model accuracy and generability can be certainly expected by extending the training set of ground truth images with more and more variable data, in particular, more examples of stressed/aged phenotypes exhibiting non-green colours, e.g., brown, yellow, red leaves. Furthermore, the tool can be extended by an automated detection of the plant type and the camera view (side or top) that have to be manually selected in the present implementation from the list of pre-trained CNN models. Finally, further investigations are required to quantitatively assess and to compare different model architectures as well as performance of binary vs. multi-class segmentation models.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

N.N., M.H. and E.G. conceived, designed and performed the computational experiments, analysed the data, wrote the paper, prepared figures, and tables, and reviewed drafts of the paper. K.N. executed the laboratory experiments, acquired image data, and reviewed drafts of the paper. T.A. co-conceptualized the study and reviewed the manuscript.

Funding

This work was performed within the German Plant-Phenotyping Network (DPPN) which is funded by the German Federal Ministry of Education and Research (BMBF) (project identification number: 031A053). M.H. was supported from European Regional Development Fund-Project "SINGING PLANT" (No. CZ.02.1.01/0.0/0.0/16_026/0008446).

Supplementary Material

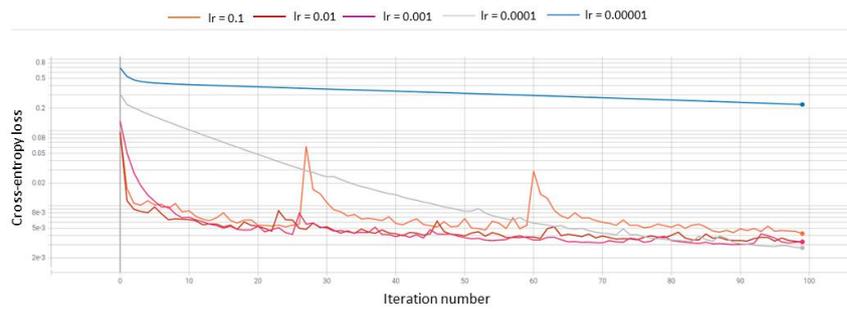
Supplementary Material:

Deep learning based greenhouse image segmentation and shoot phenotyping (DeepShoot),
N. Narisetti*, M. Henke, K. Neumann, T. Altmann, E. Gladilin*.

Table S1. Shoot Traits: The description of estimated shoot phenotypic traits in DeepShoot software.

Feature name	Description
Shoot Area	Total number of pixels in the segmented shoot image
BBX Area	Total number of pixels in the bounding box of the segmented shoot image
Shoot Area to BBX Area	Ratio of actual shoot area to bounding box area
BBX Height	Bounding box height of the segmented shoot image
BBX Width	Bounding box width of the segmented shoot image
Mean X	Mean of the geometrical distribution of segmented shoot pixels in horizontal direction
Stdev X	Standard deviation of the geometrical distribution of segmented shoot pixels in horizontal direction
Mean Y	Mean of the geometrical distribution of segmented shoot pixels in vertical direction
Stdev Y	Standard deviation of the geometrical distribution of segmented shoot pixels in vertical direction
CH Area to BBX Area	Ratio of convex-hull area to the bonding box area of segmented shoot image
Shoot Area to CH Area	Ratio of actual number of segmented pixels to convex-hull area
R mean, G mean, B mean	Mean value of red, green and blue (RGB) channels of segmented shoot image respectively
R stdev, G stdev, B stdev	Standard deviation of red, green and blue (RGB) channels of segmented shoot image respectively
H mean, S mean, V mean	Mean value of hue, saturation and value (HSV) channels of segmented shoot image respectively
H stdev, S stdev, V stdev	Standard deviation of hue, saturation and value (HSV) channels of segmented shoot image respectively
L mean, a mean, b mean	Mean value of CIE Lab channels of segmented shoot image respectively
L stdev, a stdev, b stdev	Standard deviation of CIE Lab channels of segmented shoot image respectively
R mean, G mean, B mean	Mean value of red, green and blue (RGB) channel of segmented shoot image respectively
R stdev, G stdev, B stdev	Standard deviation of red, green and blue (RGB) channel of segmented shoot image respectively
hny_H mean, hny_S mean, hny_Y mean	Mean value of hue, saturation and luminance (HSY) channel of segmented shoot image respectively
hny_H stdev, hny_S stdev, hny_Y stdev	Standard deviation of hue, saturation and luminance (HSY) channel of segmented shoot image respectively

Figure S1: Impact of learning rate (lr) on model training. X-axis represents number of iteration model trained and y-axis represents calculated binary cross entropy loss.



Note: Model with learning rate 0.001 represents optimally trained model with exponential reduction in loss at initial iterations and stabilized loss over the iterations.

Bibliography

- [1] Lei Li, Qin Zhang, and Danfeng Huang. A review of imaging techniques for plant phenotyping. *Sensors*, 14(11):20078–20111, 2014.
- [2] Nathan D Miller, Brian M Parks, and Edgar P Spalding. Computer-vision analysis of seedling responses to light and gravity. *The Plant Journal*, 52(2):374–381, 2007.
- [3] Noah Fahlgren, Malia A Gehan, and Ivan Baxter. Lights, camera, action: high-throughput plant phenotyping is ready for a close-up. *Current opinion in plant biology*, 24:93–99, 2015.
- [4] Massimo Minervini, Hanno Scharr, and Sotirios A Tsafaris. Image analysis: the new bottleneck in plant phenotyping [applications corner]. *IEEE signal processing magazine*, 32(4):126–131, 2015.
- [5] Michael Henke, Astrid Junker, Kerstin Neumann, Thomas Altmann, and Evgeny Gladilin. Comparison and extension of three methods for automated registration of multimodal plant images. *Plant methods*, 15(1):1–15, 2019.
- [6] Sruti Das Choudhury, Srinidhi Bashyam, Yumou Qiu, Ashok Samal, and Tala Awada. Holistic and component plant phenotyping using temporal image sequence. *Plant methods*, 14(1):1–21, 2018.
- [7] Jean-Michel Pape and Christian Klukas. Utilizing machine learning approaches to improve the prediction of leaf counts and individual leaf segmentation of rosette plant images. *Proceedings of the Computer Vision Problems in Plant Phenotyping (CVPPP)*, pages 1–12, 2015.
- [8] Isabelle Philipp and Thomas Rath. Improving plant discrimination in image processing by use of different colour space transformations. *Computers and electronics in agriculture*, 35(1):1–15, 2002.

- [9] Michael Henke, Kerstin Neumann, Thomas Altmann, and Evgeny Gladilin. Semi-automated ground truth segmentation and phenotyping of plant structures using k-means clustering of eigen-colors (kmseg). *Agriculture*, 11(11), 2021.
- [10] Michael Henke, Astrid Junker, Kerstin Neumann, Thomas Altmann, and Evgeny Gladilin. A two-step registration-classification approach to automated segmentation of multimodal images for high-throughput greenhouse plant phenotyping. *Plant Methods*, 16(1):1–10, 2020.
- [11] Unseok Lee, Sungyul Chang, Gian Anantrio Putra, Hyungseok Kim, and Dong Hwan Kim. An automated, high-throughput plant phenotyping system using machine learning-based plant segmentation and image analysis. *PloS one*, 13(4):e0196615, 2018.
- [12] Jason Adams, Yumou Qiu, Yuhang Xu, and James C Schnable. Plant segmentation by supervised machine learning methods. *The Plant Phenome Journal*, 3(1):e20001, 2020.
- [13] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.
- [14] Wenjia Bai, Matthew Sinclair, Giacomo Tarroni, Ozan Oktay, Martin Rajchl, Ghislain Vaillant, Aaron M Lee, Nay Aung, Elena Lukaschuk, Mihir M Sanghvi, et al. Automated cardiovascular magnetic resonance image analysis with fully convolutional networks. *Journal of Cardiovascular Magnetic Resonance*, 20(1):65, 2018.
- [15] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(12):2481–2495, 2017.
- [16] Dimitrios Marmanis, Jan D Wegner, Silvano Galliani, Konrad Schindler, Mihai Datcu, and Uwe Stilla. Semantic segmentation of aerial images with an ensemble of cnss. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2016, 3:473–480, 2016.
- [17] Michael P Pound, Jonathan A Atkinson, Alexandra J Townsend, Michael H Wilson, Marcus Griffiths, Aaron S Jackson, Adrian Bulat, Georgios Tzimiropoulos, Darren M Wells, Erik H Murchie, et al. Deep

- machine learning provides state-of-the-art performance in image-based plant phenotyping. *Gigascience*, 6(10):gix083, 2017.
- [18] Clément Douarre, Richard Schielein, Carole Frindel, Stefan Gerth, and David Rousseau. Transfer learning from synthetic data applied to soil-root segmentation in x-ray tomography images. *Journal of Imaging*, 4(5):65, 2018.
- [19] Tanuj Misra, Alka Arora, Sudeep Marwaha, Viswanathan Chinnusamy, Atmakuri Ramakrishna Rao, Rajni Jain, Rabi Narayan Sahoo, Mrinmoy Ray, Sudhir Kumar, Dhandapani Raju, et al. Spikesegnet-a deep learning approach utilizing encoder-decoder network with hourglass for spike segmentation and counting in wheat plant from visual imaging. *Plant methods*, 16(1):1–20, 2020.
- [20] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*, 2015.
- [21] Shibani Santurkar, Dimitris Tsipras, Andrew Ilyas, and Aleksander Madry. How does batch normalization help optimization?, 2019.
- [22] Xiang Li, Shuo Chen, Xiaolin Hu, and Jian Yang. Understanding the disharmony between dropout and batch normalization by variance shift, 2018.
- [23] Chao Peng, Xiangyu Zhang, Gang Yu, Guiming Luo, and Jian Sun. Large kernel matters – improve semantic segmentation by global convolutional network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.
- [24] Ranjeet Ranjan Jha, Gaurav Jaswal, Divij Gupta, Shreshth Saini, and Aditya Nigam. Pixisegnet: Pixel-level iris segmentation network using convolutional encoder–decoder with stacked hourglass bottleneck. *IET Biometrics*, 9(1):11–24, 2019.
- [25] Forest Agostinelli, Matthew Hoffman, Peter Sadowski, and Pierre Baldi. Learning activation functions to improve deep neural networks. *arXiv preprint arXiv:1412.6830*, 2014.
- [26] Limin Wang, Sheng Guo, Weilin Huang, and Yu Qiao. Places205-vggnet models for scene recognition. *arXiv preprint arXiv:1508.01667*, 2015.

- [27] Rob A Dunne and Norm A Campbell. On the pairing of the softmax activation and cross-entropy penalty functions and the derivation of the softmax activation function. In *Proc. 8th Aust. Conf. on the Neural Networks, Melbourne*, volume 181, page 185. Citeseer, 1997.
- [28] Kelly H Zou, Simon K Warfield, Aditya Bharatha, Clare MC Tempany, Michael R Kaus, Steven J Haker, William M Wells III, Ferenc A Jolesz, and Ron Kikinis. Statistical validation of image segmentation quality based on a spatial overlap index1: scientific reports. *Academic radiology*, 11(2):178–189, 2004.
- [29] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Craig Chen Zhifeng Citro, Greg S Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467*, 2016.
- [30] Stéfan van der Walt, S Chris Colbert, and Gael Varoquaux. The numpy array: a structure for efficient numerical computation. *Computing in science & engineering*, 13(2):22–30, 2011.
- [31] Stefan Van der Walt, Johannes L Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D Warner, Neil Yager, Emmanuelle Gouillart, and Tony Yu. scikit-image: image processing in python. *PeerJ*, 2:e453, 2014.
- [32] Alessandro Crimi, Spyridon Bakas, Hugo Kuijf, Bjoern Menze, and Mauricio Reyes. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: Third International Workshop, BrainLes 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 14, 2017, Revised Selected Papers*, volume 10670. Springer, 2018.
- [33] V. Roshan Joseph. Optimal ratio for data splitting. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, n/a(n/a), 2022.
- [34] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.
- [35] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

- [36] R. R. Jha, G. Jaswal, D. Gupta, S. Saini, and A. Nigam. Pixiseg-net: pixel-level iris segmentation network using convolutional encoder–decoder with stacked hourglass bottleneck. *IET Biometrics*, 9(1):11–24, 2020.
- [37] Yoshua Bengio. Practical recommendations for gradient-based training of deep architectures. In *Neural networks: Tricks of the trade*, pages 437–478. Springer, 2012.
- [38] Mathworks, matlab and statistics toolbox release 2021a, 2021. The MathWorks, Natick, MA, USA.
- [39] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [40] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- [41] François Chollet. Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1251–1258, 2017.
- [42] Hanno Scharr, Massimo Minervini, Andrew P French, Christian Klukas, David M Kramer, Xiaoming Liu, Imanol Luengo, Jean-Michel Pape, Gerrit Polder, Danijela Vukadinovic, et al. Leaf segmentation in plant phenotyping: a collation study. *Machine vision and applications*, 27(4):585–606, 2016.

5. Automated Spike Detection in Diverse European Wheat Plants Using Textural Features and the Frangi Filter in 2D Greenhouse Images

Paper 4: Narisetti N, Neumann K, Röder MS and Gladilin E (2020) Automated Spike Detection in Diverse European Wheat Plants Using Textural Features and the Frangi Filter in 2D Greenhouse Images. *Front. Plant Sci.* 11:666. doi: 10.3389/fpls.2020.00666

Impact Factor: 6.6 (according to Clarivate releases the Journal Citation Report (JCR 2023))

Abstract

Spike is one of the crop yield organs in wheat plants. Determination of the phenological stages, including heading time point (HTP), and area of spike from non-invasive phenotyping images provides the necessary information for the inference of growth-related traits. The algorithm previously developed by Qiongyan et al. for spike detection in 2-D images turns out to be less accurate when applied to the European cultivars that produce many more leaves. Therefore, we here present an improved and extended method where (i) wavelet amplitude is used as an input to the Laws texture energy-based neural network instead of original grayscale images and (ii) non-spike structures (e.g., leaves) are subsequently suppressed by combining the result of the neural network prediction with a Frangi-filtered image. Using this two-step approach, a 98.6% overall accuracy of neural network segmentation based on direct comparison with ground-truth data could be achieved. Moreover, the comparative error rate in spike HTP detection and growth correlation among the ground truth, the algorithm developed by Qiongyan et al., and the proposed algorithm are discussed in this paper. The proposed algorithm was also capable of significantly reducing the error rate of the HTP detection by 75% and improving the accuracy of spike area estimation by 50% in comparison with the Qiongyan et al. method. With these algorithmic improvements, HTP detection on a diverse set of 369 plants was performed in a high-throughput manner. This analysis demonstrated that the HTP of 104 plants (comprises of 57 genotypes) with lower biomass and tillering range (e.g., earlier-heading types) were correctly determined. However, fine-tuning or extension of the developed method is required for high biomass plants where spike emerges within green bushes. In conclusion, our proposed method allows significantly more reliable results for HTP detection and spike growth analysis to be achieved in application to European cultivars with earlier-heading types.

5.1 Introduction

Wheat is one of the major crop species in the world, with 762 million tons of grain produced annually (FAOSTAT 2018) and providing $\geq 20\%$ of the world's calorie and protein demand [1]. However, the increasing world population and climate change are major threats to sustainable crop production [2]. Therefore, concentrated efforts are required to increase crop yield and production to meet future needs. Information-based plant breeding and precision agriculture are fundamental for identifying suitable wheat varieties to increase wheat productivity and production. One of the important components in both crop breeding and precision agriculture is the monitoring of plant developmental growth stages to apply informed-decision-based treatments in field or greenhouse experiments. Phenology influences grain yield components both directly and indirectly [3, 4], and in this aspect, quantitative assessment of crop phenology plays an important role in precision phenotyping as a quantifier of crop performance.

According to the Feekes scale, wheat growth can be classified into four major growth stages: tillering, stem elongation, heading, and ripening. A more detailed sub-classification is made in the BBCH scale [5], with BBCH classes 49–59 representing phenology from heading to flowering. The determination of phenological stages is necessary for the interpretation of growth-related traits and stress tolerance acquired from non-invasive phenotyping. It is well-known that the major flowering time gene *PPD-H1* has a direct influence on leaf growth in barley [6], and flowering time genes have an impact on abiotic stress tolerance [7, 8]. In a study employing non-invasive phenotyping of barley growth, correlation of biomass and tipping time (BBCH49) was high [9] and resulted in a constitutive biomass QTL in the region of *PPD-H1* (Dhanagond et al., 2019). However, tipping time had to be assessed by a time-consuming visual inspection of individual plant images across time. The relationship of biomass to flowering time also holds true for wheat: both crops have delayed flowering in an environment with long growing seasons to allow longer and higher vegetative growth [10]. Similar to barley, sensitive or insensitive *Ppd-D1* alleles in wheat have been shown to correspond to differences in leaf area [11]. In winter wheat, an earlier flowering time of semidwarf cultivars was associated with reduced biomass at anthesis [12]. In dryland regions, simulations showed that higher yield derives from an increased biomass before anthesis leading to an increased grain number [4]. Non-invasive imaging experiments with a large wheat collection have been conducted to genetically dissect drought and heat-stress tolerance (unpublished data). An automated solution is urgently required for an effective

determination of flowering time-related growth stages through non-invasive imaging.

As a first step, a reliable method for spike detection is needed. Once this is established, the time point of the first detection of spikes across a time course can be determined. To date, there have been relatively few studies concerned with wheat spike detection and growth analysis from digital images. Most of them are based on single spikes and needed to cut off spikes to classify different wheat varieties using morphological image processing algorithms, Hu moments, and neural networks [13, 14, 15]. However, these methods are unsuitable for non-invasively detecting spikes from a whole plant with overlapping of leaves and young developing spikes in a high-throughput manner.

Qiongyan et al. [16] proposed a novel approach for detecting (young) spikes in digital images of wheat plants based on Law’s textural (energy) features and a neural network. This approach is based on the fact that spikes and leaves have a high color similarity but differ clearly in texture. However, when we applied this algorithm to one of our data sets, it turned out to be sensitive to the high-energy leaf edges and tillers, which led to false classifications of spike and non-spike pixels (or noisy pixels) as shown in Figure 5.1. However, their method was based on four Australian wheat varieties. In contrast, our data set is based on a diverse collection of high-yielding mainly European elite cultivars that are much more diverse in their plant architecture and produce more leaves and biomass compared to Australian genotypes. Accordingly, due to the presence of noisy pixels in the final image segmentation, the heading time point (HTP) BBCH55 was detected too early on our dataset compared to the ground truth data using their method. Thus, solely depending on Law’s textural features lead to false detection of spikes in our wheat panel. Therefore, to overcome these artifacts, an improved and extended novel approach is proposed in this paper.

The paper is structured as follows. Section 2 deals with the improved methodological framework of spike detection, including data preparation, segmentation, and post-processing algorithms. Section 3 describes the improvement of our algorithm compared to the existing method for HTP detection and the spike growth analysis. In summary (section 4), we draw conclusions regarding the performance of our algorithm and discuss its future improvements.



Figure 5.1: Limitations of wheat spike detection using Qiongyan et al. algorithm. (a) represents the detection of spike and non-spike pixels in the wheat plant. (b) Zoomed version of detected pixels in (a).

5.2 Materials and Methods

5.2.1 Dataset

We used images from one experiment with 260 diverse winter wheat cultivars of mainly Central European origin. Of these lines, 220 correspond to the collection described in [17] and represent high-yielding cultivars of the past decades. The remaining 40 lines are winter wheat elite cultivars from the Gabi-Wheat collection [18], representing a similar breeding pool. Each cultivar was represented by two biological replicates. Sowing was done in small turf trays, and 14 days after sowing (DAS) at about the 2-leaf stage, plants were placed for vernalization into a growth chamber with an 8-h light period and 4°C day/night. After 8 weeks of vernalization, turf trays were placed in a greenhouse with 15-h light and 16°/12°C during the day/night for 3 days to acclimate the plants to higher temperatures. The plants were then repotted from the trays to 2-l-volume pots and were grown in the same greenhouse for another 7 days before they were placed on the imaging system, a LemnaTec 3D Scanalyzer (LemnaTec GmbH, Aachen, Germany). They were imaged and watered daily, with watering by target weight option corresponding to 89% of the plant-available water content in the soil [19]. Temperatures in the greenhouse of the imaging system were raised over the time course of the

experiment from 16°/12°C in four steps to 30°/12°C to simulate a German spring/summer growing period, including 10 days of heat stress. In total, plants lasted 50 days on the imaging system before they were transferred to a normal greenhouse at 130 days after sowing (DAS) to grow to maturity and to evaluate the yield components. During the imaging period, the tiller number per plant was counted manually at the end of the heat period (at 125 DAS).

Images were taken from three side view angles (0°, 45°, 90°) and one top view using RGB cameras. The top view camera (a Manta G-504) had a resolution of 2,452 × 2,056 pixels with a pixel size of 3.45 × 3.45 μm, while the side view camera had a resolution of 6,576 × 4,384 pixels and a pixel size of 5.5 μm × 5.5 μm. Plant images were later visually inspected to determine the time point of heading when the ear was half out of the flag leaf (BBCH55). Here, top view images turned out not to be suitable as, from the top, an emerging ear has a very low visible area and might be easily hidden under a bending leaf. Moreover, it is hard to define how much of the ear is above the flag leaves. Therefore, this determination was done on inspecting the three side view images. In this case, only the pots were rotated; the camera is stable. Out of all 520 plants, 369 reached BBCH55 during the imaging period belonging to 202 different cultivars. These 369 plants from 202 genotypes were available for testing our spike detection algorithm. These plants exhibit strong differences in plant architecture and are challenging for this kind of analysis, for example, spikes with or without awns, short and tall plants (plant height range at harvest time from 34 to 119 cm), and especially low and high tillering genotypes ranging from 1 to 38 tillers per plant counted at 125 DAS during the imaging period. Further, the data set exhibits differences in BBCH55 timing of 29 days.

5.2.2 Methodology

The workflow for spike detection following image acquisition is shown in Figure 5.2. This algorithm was developed in the MATLAB environment (MATLAB 2019a). The methodology involved in the proposed algorithm is as follows:

In the initial step, the original image (Figure 5.3a) is converted to a grayscale image using MATLAB’s *rgb2gray* routine. To enhance the separability between the plant and background pixels, discrete wavelet transform (DWT) is applied in the preprocessing step using the Haar basis function [20]. The DWT is a single level 2-D wavelet decomposition that produces a

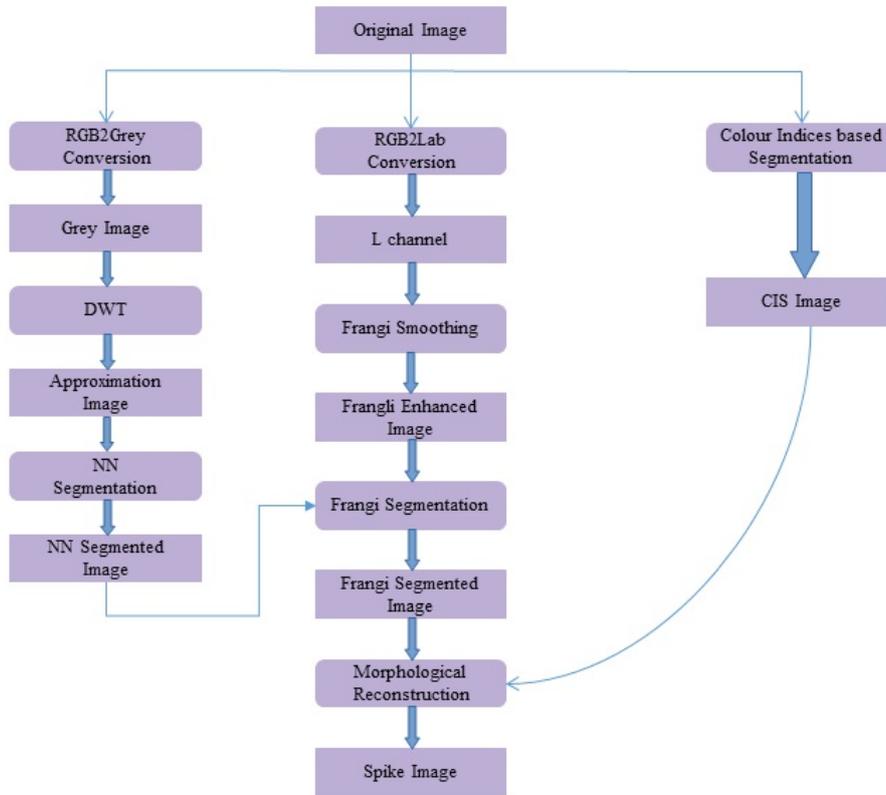


Figure 5.2: The workflow of the proposed spike detection algorithm using image processing methods. Framed rectangles represent the data modalities, and other rectangles describe the image processing operations.

featured image called an approximation coefficients image (A). This image is half the size of the original image and is useful for characterizing unique textures. Then, a neural network-based Laws texture energy method is applied to image A, as proposed in [21, 16], to segment the spike pixels from the plant pixels. Here, the segmentation of plant pixels from the background is called color index-based segmentation (CIS). Example images of the CIS and the neural network segmentation are shown in Figures 5.3b and 5.3c respectively. However, the Laws texture energy is sensitive to the high-energy noisy edges (or pixels on leaves and leaf crossings) in the plant. To eliminate those noisy edges, a combination of a multi-scale Frangi-filtered image [22] and the neural network segmented image is considered. Because the Frangi filter delivers a strength estimate of edges in the image, noisy edges can be suppressed by smoothing the image over multiple scales and orientations [22].

Therefore, this combination suppresses the tiny leaf edges and leaf crossings in the segmented image. Here, the Frangi filter is applied to an L component of the L*a*b color space image because the intensity values in the L component are closely matched with the human perception and contrast between the plant and non-plant pixels is high compared to in the a and b channels.

The Frangi-filtered image is considered one of the post-processing steps, because as a pre-processing step, it might lead to false representation of textures in the image. In other words, there might be a possibility of suppressing the spike pixels, hence modifying the unique textural characteristics of the spikes and leaves. Examples of a Frangi-filtered image and a segmented image are shown in Figures 5.3d and 5.3e, respectively. The complete spike is then recovered by applying morphological binary operations to the Frangi segmented image, as shown in Figure 5.3f.

Wavelet Decomposition

The wavelet-based texture classification is important because (1) it decorrelates the data [23] by stretching the color differences between plant and non-plant pixels in the image, (2) provides a non-redundant compressed image which reduces the computation complexity significantly compared to the original gray-scale image. Typically, wavelets are defined for 1-D signals. So, the extension to 2-D signals is usually performed by using a product of 1-D filters. The practical implementation of the wavelet transforms using different filters as follows:

$$\begin{aligned}
 A &= [L_x * [L_y * I]_{\downarrow 2,1}]_{\downarrow 1,2} \\
 H &= [L_x * [G_y * I]_{\downarrow 2,1}]_{\downarrow 1,2} \\
 V &= [G_x * [L_y * I]_{\downarrow 2,1}]_{\downarrow 1,2} \\
 D &= [G_x * [G_y * I]_{\downarrow 2,1}]_{\downarrow 1,2}
 \end{aligned} \tag{5.1}$$

where * denotes the convolution operator, ($\downarrow 2, 1$) and ($\downarrow 1, 2$) represents the downscaling along rows and columns respectively. L and G are the low and high pass filters and I is the original image.

The DWT decomposes an image into four sub-bands called approximation coefficients (A), horizontal (H), vertical (V), and diagonal (D), as shown in Figure 5.4. Sub-band A is obtained by the low-pass filtering and is accordingly called the low-resolution image, the size of which is dependent on the level of decomposition and input image size. In contrast, H, V, and D are obtained by bandpass filtering in a specific direction. Therefore, they provide detailed directional information for the image. Among these sub-bands,

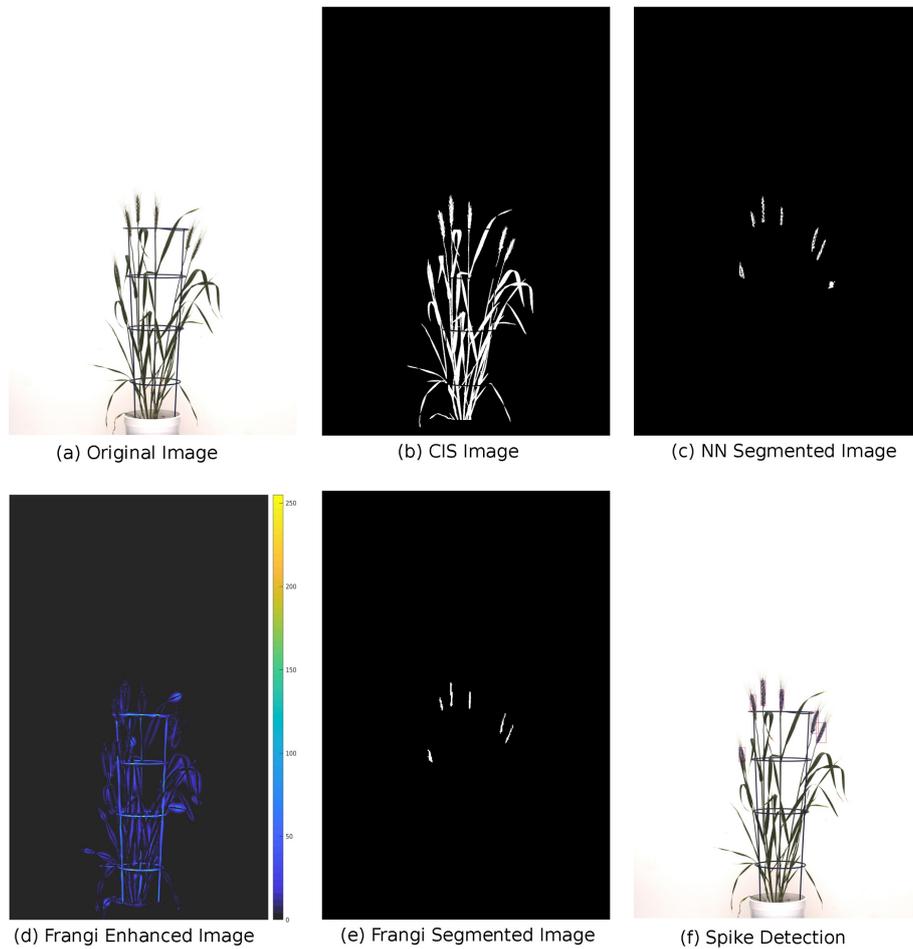


Figure 5.3: The methodology of the proposed spike detection algorithm with example images: (a) wheat plant of ID 1817KN373 with 150 days after sowing (b) Green color indices based segmented image (c) DWT + Laws textural features based NN segmented image (d) Frangi enhanced image in multiple scales and orientations (e) Final binary segmentation: one leaf crossing artifact is suppressed with Frangi enhanced image (f) Detected spikes after the morphological reconstruction.

A is an essential feature image (or coefficients image) bearing the textural information relevant to image segmentation. Consequently, the A wavelet coefficient image is used here for texture characterization.



Figure 5.4: DWT Decomposition: The coefficient image (A) is again decomposed in the multilevel DWT decomposition.

Laws textural features based image segmentation using neural networks

Laws texture energy method [24] is a classical pixel-wise texture analysis approach and it has been used in many applications [25, 26, 27, 28]. This approach uses 1-D local masks to detect various types of micro-structure textural features. The typical 1-D local masks are level detection, edge detection, and spot detection as shown in equation 5.2. However, the image is a two-dimensional and which requires the 2-D masks for the texture analysis.

$$\begin{aligned}
 L3 &= [1 \ 2 \ 1] - \text{Level detection} \\
 E3 &= [-1 \ 0 \ 1] - \text{Edge Detection} \\
 S3 &= [-1 \ 2 \ -1] - \text{Spot Detection}
 \end{aligned} \tag{5.2}$$

The 2-D masks are generated from the 1-D masks by convolving the vertical 1-D mask with the horizontal 1-D mask. For example, the mask S3L3 is calculated by a convolving vertical S3 mask with the horizontal L3 mask and it is a zero-sum mask. In contrast, the mask L3L3 is a non-zero-sum mask which is not considered for the texture analysis. The list of 2-D masks used for the texture analysis is given as the following:

$$\begin{aligned}
 L3E3 &= L3^T * E3; & E3S3 &= E3^T * E3; \\
 L3S3 &= L3^T * S3; & S3L3 &= S3^T * L3; \\
 E3L3 &= E3^T * L3; & S3E3 &= S3^T * E3; \\
 E3E3 &= E3^T * E3; & S3S3 &= S3^T * S3;
 \end{aligned} \tag{5.3}$$

The textural features are calculated in two steps [25] using 2-D masks. In the first step, the input image is convolved with all the above 2-D masks.

Then each individual resulting image is normalized with a unit standard deviation and average mean over the window size of 25. Consequently, 8 textural feature images are generated for every input image. However, these feature images have both plant and background pixels which increase the computational complexity of the neural network for spike detection.

To overcome the computational complexity of the image segmentation, the plant pixels (PP) are segmented from the background pixels using the CIS method [21] as follows

$$PP = 2g - r - b; \quad (5.4)$$

This method decorrelates the dominating green plant pixels from the background. Then a plant binary image is generated using the binarization technique (pixel value > 0), see Figure 5.3b. As a result, the number of pixels for the neural network-based segmentation is reduced significantly.

The neural network is used to perform the classification of spike and non-spike pixels in the study. In practice, the neural network is trained with a large quantity of spike and non-spike pixels from the different wheat plants. Then the trained neural network parameters are adapted to perform the spike detection in an automated manner. Here, totally 218282 spike- and 731054 non-spike pixels were extracted from 150 manually segmented images and subsequently used for training, testing and validation a network model in the sample proportion 70:15:15. The performance of the network model with the eight input nodes, one hidden layer with 10 hidden nodes and 2 output nodes was assessed using the conventional confusion matrix [TP FP; FN TN], components of which indicate the total number of correctly and incorrectly classified spike and non-spike pixels, respectively. The true positive (TP), true negative (TN) rates as well as the overall accuracy $(TP+TN)/(TP+FP+FN+TN)$ are summarized in table 5.1.

Frangi Filter Enhancement

The Frangi filter is a multi-scale second-order vessel enhancement method developed by Frangi et al. [22] that is frequently used in biomedical applications [29, 30, 31]. The Frangi filter is used for enhancement of high-contrast vessel structures or edges along with the suppression of the non-vessel structures and thin vessel edges. Since wheat shoots have multiple leaf crossings, they exhibit vessel-like thin structures producing high-energy signals similar to spikes. In turn, this can lead to false spike detection at leaf crossings by the network model, as shown in Figure 5.1. The Frangi filter is applied to

	Training	Testing	Validation	Total
Spike pixels	152793	32773	32716	218282
Non-spike pixels	511743	109627	109684	731054
TP rate (%)	96.2	96.4	96.0	96.2
TN rate (%)	99.3	99.3	99.3	99.3
Accuracy (%)	98.5	98.6	98.5	98.6

Table 5.1: The statistical performance of the neural network in the training stage.

suppress edges resulting from such leaf crossings in the neural network segmented images.

Frangi-based vessel enhancement is achieved based on Hessian and eigenvalues. The Hessian matrix of image I is computed as follows:

$$H = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} = \sigma \begin{bmatrix} \frac{\partial^2 I}{\partial x^2} & \frac{\partial^2 I}{\partial x \partial y} \\ \frac{\partial^2 I}{\partial y \partial x} & \frac{\partial^2 I}{\partial y^2} \end{bmatrix} \quad (5.5)$$

where $h_{11}, h_{12}, h_{21}, h_{22}$ are the second-order partial derivatives of the image and σ denotes a variable scaling factor.

To extract the information on structural-patterns from the Hessian matrix, the eigenvalues λ_1 and λ_2 are calculated, while σ is used for the enhancement of structures at different scales, see Table 5.2. Since we are interested in detecting and suppressing the bright vessel-like structures in the plant leaves, the image enhancement is performed under the assumption that a pixel belonging to a vessel region should have a very low value of λ_1 and a very high magnitude of λ_2 , see equation 5.6. Furthermore, the bright vessel-like structures emerge with negative λ_2 and the filter response of the corresponding pixel with $\lambda_2 > 0$ is considered as zero in the enhanced image.

$$|\lambda_1| \leq |\lambda_2| \quad (5.6)$$

The enhanced image is defined as follows:

$$I_E = \begin{cases} \text{if } \lambda_2 > 0 : & 0, \\ \text{otherwise} : & \exp\left(\frac{-R_B^2}{2\beta^2}\right) \left(1 - \exp\left(\frac{-S^2}{2c^2}\right)\right) \end{cases} \quad (5.7)$$

λ_1	λ_2	Structure pattern
N	N	noisy, no preferred direction
L	H-	vessel structure (bright)
L	H+	vessel structure (dark)
H-	H-	blob like structure (bright)
H+	H+	blob like structure (dark)

Table 5.2: Possible structural patterns in 2D images depending on eigen values λ_1 and λ_2 . H = high, L = low, +/- indicates the sign of the eigen value [22].

where $R_B = \frac{\lambda_1}{\lambda_2}$, $S = \sqrt{\lambda_1^2 + \lambda_2^2}$ and c , β are constants that control the sensitivity of the filter. The enhanced image I_E is obtained at various scales, i.e., $\sigma = 1, 3, 5, 7, 9$. Since the maximum scale approximately matches the size of the vessel to detect, the final enhanced image I_{FE} can be obtained according to [22] by taking a maximum among all scales as defined in equation 5.8.

$$I_{FE} = \max_{\sigma} I_E \quad (5.8)$$

An example of edge suppression (leaf crossings) using the Frangi filter is shown in Figure 5.5.

Consequently, the result of the neural network segmentation is subsequently filtered under consideration of leaf crossing regions detected by the Frangi filter (Figure 5.3d). This is done by eliminating the regions corresponding to leaf edges in the binary segmentation mask, see Figure 5.3e.

Spike reconstruction using morphological filters

As shown in Figure 5.3e, only some parts of the spikes were detected using the proposed algorithm compared to the CIS image in Figure 5.3b. To recover the complete spikes, the logical 'and' operation of the CIS image and the Frangi segmented image were performed. Then the morphological binary operations (erosion and dilation) were sequentially applied to recover the final spike area in the CIS image, see Figure 5.3f.

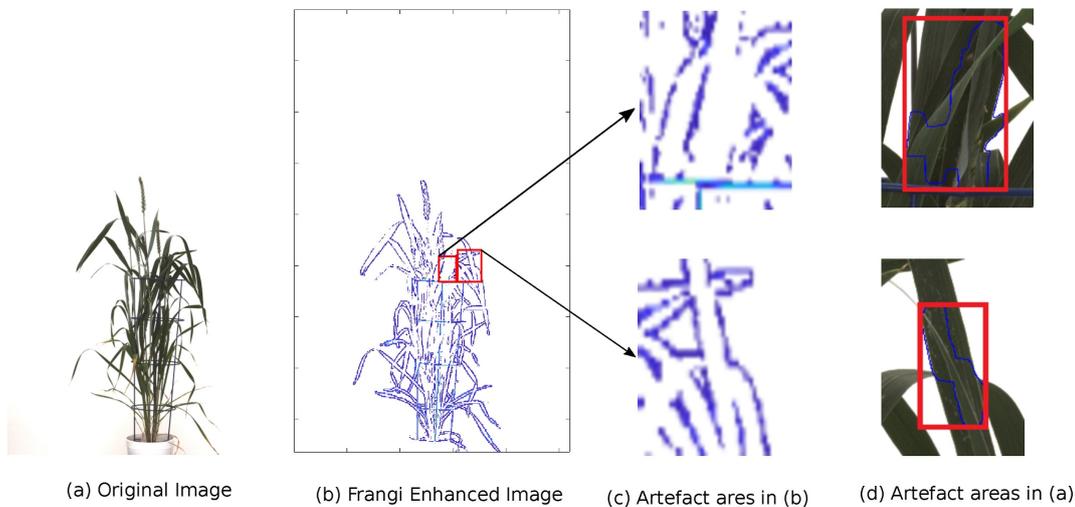


Figure 5.5: Example of suppression of leaf crossings using the Frangi Filter. From left to right: (a) original image of a wheat shoot, (b) Frangi filter enhanced image, (c) examples of Frangi enhanced regions, (d) examples of leaf crossings detected in the original image.

5.3 Results and Discussion

The above-described algorithm was applied to calculate the yield-related features at the transition from the tillering to flowering growth stages of wheat plants with an age of more than 90 DAS. Accordingly, the results of this study are presented in two sections dedicated to (i) detection of the time point of spike emergence and (ii) spike growth analysis from RGB images acquired using visible light cameras during an experiment with diverse winter wheat varieties. In the first section, the spike emergence was tested on 369 wheat plants from 202 different genotypes. Here, the HTP was defined as the first time in the imaging time course when the detected spike satisfied the minimum area constraint of 500 pixels. The spike area was then measured in a time series from the HTP to perform real-time growth analysis for a few selected plants.

Image analysis was performed on an Intel Xeon CPU E5-2640-based workstation running under the Linux OS. The algorithms were implemented under MATLAB 2019a (MathWorks Inc.). Training of a neural network on 949336 manually segmented spike and non-spike pixels using ten 2.40GHz CPUs with a total of 20 cores in parallel mode took approximately 80 s. The spike detection algorithm takes approximately 2.5 s to process an 8-megapixel test

image. However, the processing time might vary depending on the test image size.

The root mean square error (RMSE) is used for quantification of the deviations of predictions from our model and Qiongyan et al.’s model from ground truth data,:

$$\text{RMSE} = \sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (5.9)$$

where y is the ground truth value and \hat{y} is the model predicted value. For consistent comparison of performance, the Qiongyan et al. model was retrained with the European cultivars.

5.3.1 Spike Emergence

The time-series images of a single plant described in Section 2 have three orientations. Accordingly, two factors are considered to estimate the HTP from multiple orientations: the spike should (1) appear in at least two orientations and (2) remain present in all days of the experiment. This means the spike or spikes should be continuously detected until the last day to avoid false emerging time points.

Figure 5.6 shows HTP detection in the wheat plant side-view images. These nine different wheat plants from the early-flowering genotypes possessing a single spike (1817KN397, 1817KN422) and multiple spikes (remaining seven plants) were considered for the training a model because we were aware that the later-flowering genotypes, which produce more biomass, will present much greater difficulties with spike visibility due to a higher probability that the spike will be covered by leaves. Figure 5.6 indicates that HTP values obtained by the proposed method have a significant correlation with the ground-truth HTPs, with an RMSE of 1.94, whereas the Qiongyan et al. method underperforms, with an RMSE of 7.8. This indicates that the Qiongyan et al. method is highly sensitive to the leaf artifacts whose energy is similar to that of the spike pixels but that those leaf artifacts were suppressed by the proposed method, as shown in Figure 5.5.

On the other hand, the proposed method resulted in high HTP error rates of 4 days more and 3 days less for plant ID 1817KN373 and 1817KN412, respectively. For plant ID 1817KN373, this was because the spikes were narrow and the pixel-wise textural energy was similar to that of the leaves, as shown

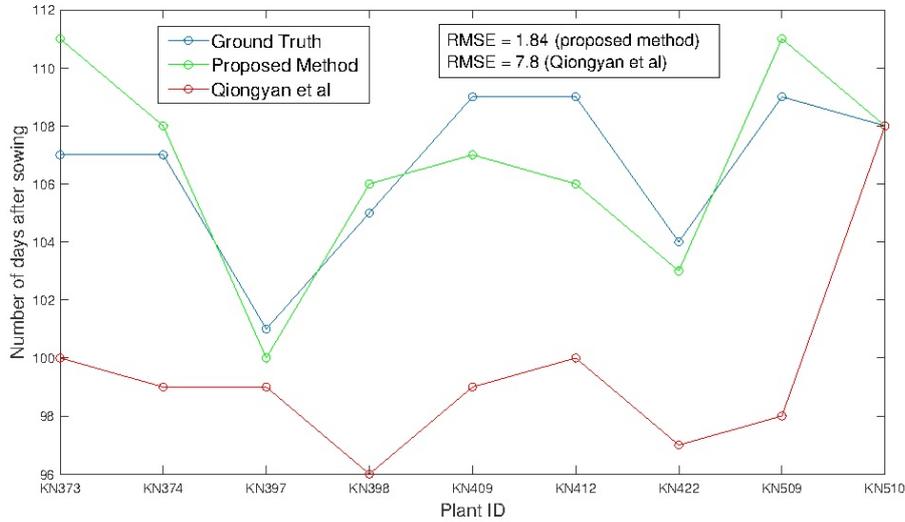


Figure 5.6: The HTP detection using Qiongyan et al. and our proposed method in comparison to the ground truth.

in Figure 5.7a compared to the other spikes in Figure 5.7b. Therefore, the HTP was detected 4 days later. In the case of plant ID 1817KN412, it turned out that the visually scored time point was determined too late, most likely by not carefully inspecting all side view angles (in the first, at 00° , the later time point looks correct, but at the 450° and 900° angles, it is visible that the earlier one is correct). Example spike images for the early HTP detection are shown in Figures 5.7c and 5.7d.

The advantages and significance of the results with the proposed method showed that it is feasible for high-throughput analysis of HTP detection. Consequently, we applied the method to all 369 diverse plants in our data set that reached heading within the imaging period. As expected, 104 plants corresponding to the supposedly earlier-heading genotypes obtained a good and reliable estimation of the true heading time point. Figure 5.8 shows the results for the high-throughput analysis of 104 plants. It is observed that the proposed method outperforms the Qiongyan et al. method, with an R^2 value of 0.776 compared to the R^2 value (0.193) of the Qiongyan et al. method. Since the European elite cultivars possess more leaves, overlay artifacts result in too early HTP detection using the Qiongyan et al. method on 90% of our data. In the remaining 265 plants, the spike emerged in the final days according to the ground truth data, and they have early-stage spike textural

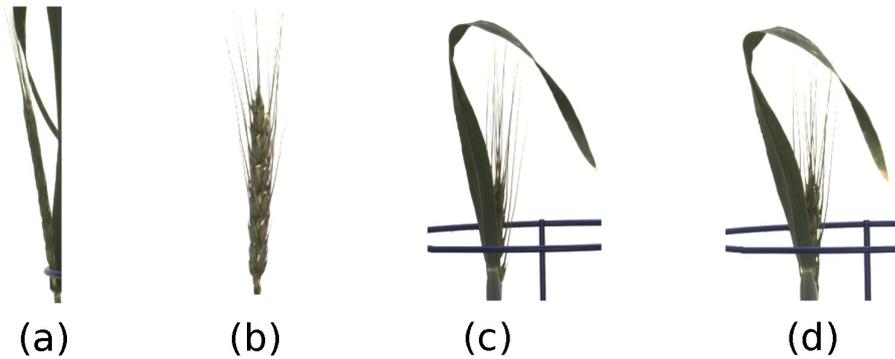


Figure 5.7: Limitations of the proposed method: (a) the early stage spike texture failed to detect in plant ID 1817KN373 (b) the detected spike texture in plant ID 1817KN373 (c) example spike geometry less than BBCH55 scale in plant ID 1817KN412 (d) spike geometry according to BBCH scale in plant ID 1817KN412.

features that are similar to the leaves. This resulted in the proposed algorithm failing to detect the spikes in the final days with a day number 0 in the output. This leads to a low value of the correlation coefficient R^2 0.0586 for the remaining 265 plants.

We compared the general plant architecture features of all 369 plants tested and classified them into three categories: (i) both plants of the genotype were classified correctly by our algorithm (94 plants from 47 genotypes), (ii) only one out the two plants of a genotype were classified correctly by our algorithm (20 plants from 10 genotypes), and (iii) none of the two plants of a genotype were classified correctly by our algorithm (Table 5.3; Supplementary Material). It turned out that the method performed better for earlier-flowering plants with an accordingly lower number of tillers and less biomass. Moreover, in 26 out of all 39 plants with awned spikes, heading time could be reliably estimated by our algorithm. This might arise from two factors: first, awned genotypes are more abundant in the earlier-flowering group and possess less biomass, and therefore spikes are less often hidden by leaves, and second, the model was trained based on nine early-flowering plant IDs with a bias toward awned types. Further, it might very well be that the fine awn structures, in general, help in the differentiation of the spike from the leaf background.

Table 5.3 shows mean phenotypic trait values, with minimum and max-

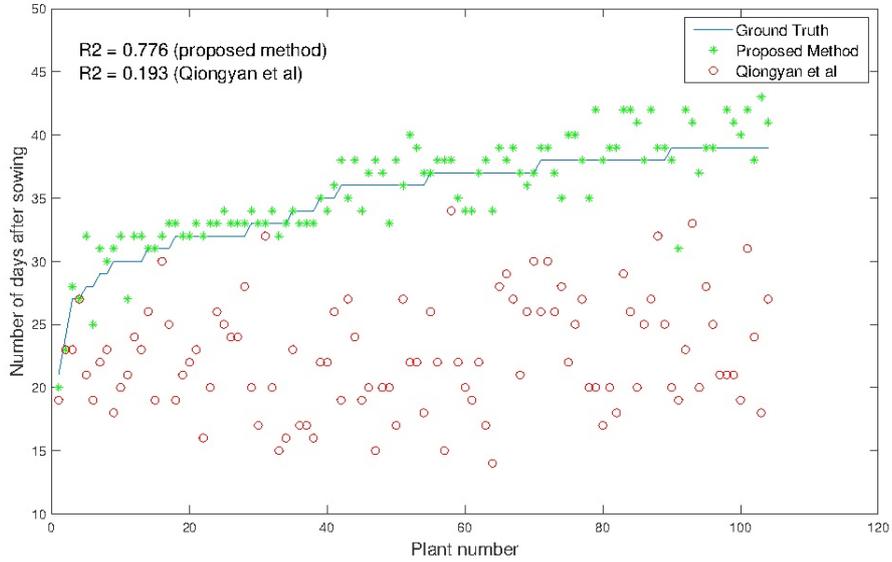


Figure 5.8: Comparison of the HTP detection using our and Qiongyan et al method vs. ground truth in 104 wheat plants.

imum in brackets, of plants successfully and non-successfully classified regarding their time point of heading.

5.3.2 Spike Area

Spike area is one of the key yield measures in wheat plants, so we have examined the total spike growth of a single wheat plant in three orientations from the spike emergence day for all images. In section 5.3.1, nine wheat plants were considered for HTP detection. Among those, only three plants (1817KN374, 1817KN409, and 1817KN422) with a single spike and two with multiple spikes are considered for the spike growth analysis. Here, the spike area of a plant per day is calculated by taking the maximum area among the three orientations. The measured area of both algorithms is validated by the RMSE and R^2 . The RMSE quantifies the difference between the ground truth area and the predicted area for all days from the ear-emergence day. The R^2 value compares the goodness of our proposed models and of the Qiongyan et al. model compared to the ground truth data.

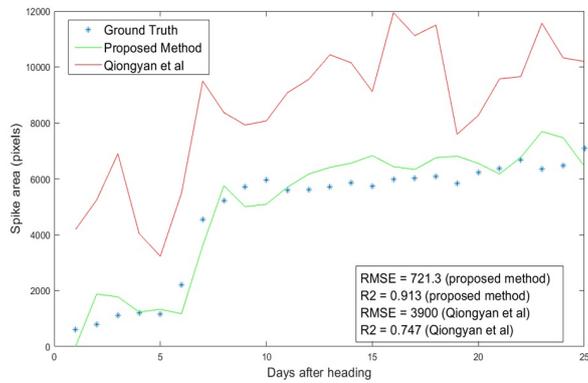
Figure 5.9 shows the results of spike growth analysis with the Qiongyan et al. method and with our proposed method compared to the ground truth

Phenotypic traits	Phenotypic mean values		
	2 out of 2 plants successful	1 plant of 2 successful	0 out of 2 plants successful
ground truth BBCH55 (DAS)	115.5 (107-120)	118.1 (101-127)	125.5 (120-130)
days to maturity (DAS)	175.4 (159-203)	185.2 (160-222)	193.8 (166-283)
presence of awns (1=no, 2=yes)	1.3	1.2	1.0
final plant height(cm)	57.1 (34-101)	64.0 (37-96)	60.9 (38-119)
Tiller number at DAS 125	7.5 (3-19)	8.4 (1-17)	11.4 (4-38)
Spike number at har- vest	7.5 (3-16)	7.8 (1-14)	9.8 (4-22)
total plant biomass at harvest (grains + straw) (g)	15.2 (5.7-26.8)	17.5 (4.5-28.1)	21.4 (8.1-48.0)
total plant straw weight at harvest (g)	9.9 (3.5-15.6)	12.8 (5.7-20.0)	15.7 (5.8-38.2)

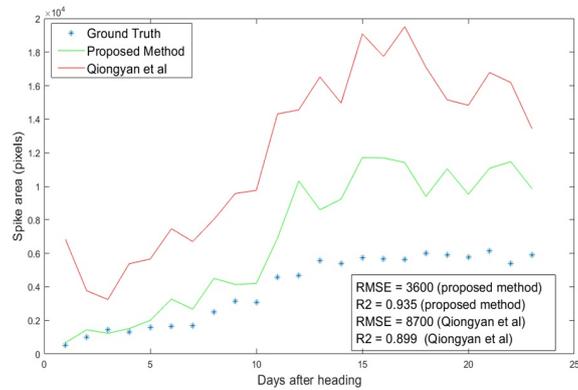
Table 5.3: Gene classification and comparison of architectural features of 369 plants.

data. Here, the ground truth data are prepared manually by segmenting the spikes using GIMP image processing software (<https://www.gimp.org>). The number of non-zero pixels in the segmented image represents the actual spike area or the ground-truth spike area of the image. This figure shows that the proposed method outperforms the Qiongyan et al. method overall, with a low RMSE and a high value of R^2 . Moreover, the RMSE is profoundly improved by more than 50% and the R^2 value is significantly improved for plant ID 1817KN373 (Figure 5.9a) and 1817KN422 (Figure 5.9c). Nevertheless, plant ID 1817KN409 (Figure 5.9b) exhibits a high RMSE compared to the other plants in the spike growth analysis.

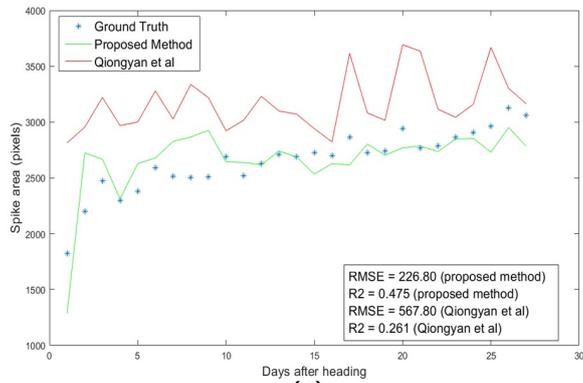
The high RMSE value for the Qiongyan et al. method is caused by the classification of leaf artifacts as spikes, which leads to an increase in the total spike area. In our method, these artifacts were eliminated using DWT and



(a)



(b)



(c)

Figure 5.9: Spike growth analysis: Day number 1 represents the first day of the ear emergence in the wheat plants. (a) plant ID 1817KN374 with multiple spikes (b) plant ID 1817KN409 with multiple spikes (c) plant ID 1817KN422 with a single spike.

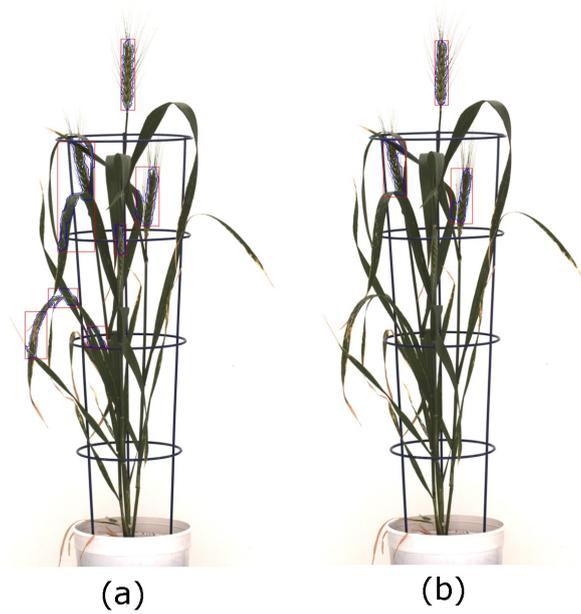


Figure 5.10: The detected leaf artifacts in (a) result in a high spike area compared to (b) for the spike growth analysis. The segmented objects are represented with blue color curves and red color rectangle boxes. (a) Qiongyan et al method, (b) proposed method.

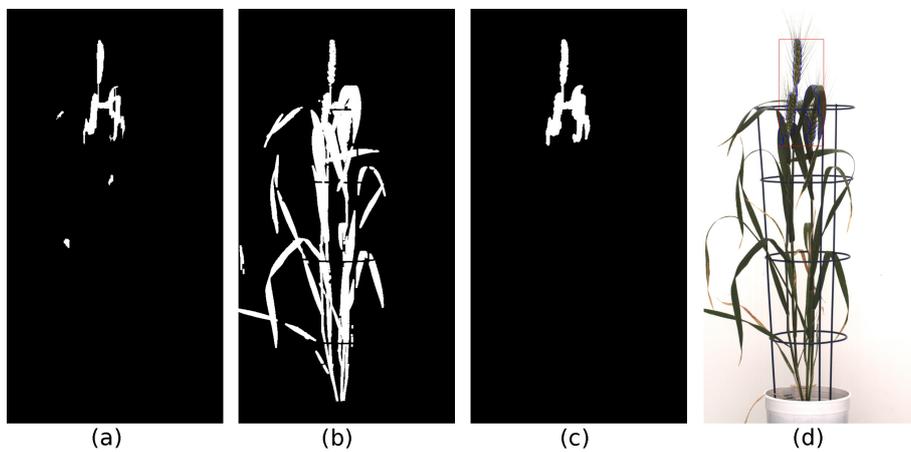


Figure 5.11: Morphological reconstruction of the spikes: (a) Frangi based spike segmentation (b) CIS image (c) Morphologically reconstructed image using a and b (d) Detected spikes in the original image represented with blue color lines and red color rectangle box.

the Frangi filter. Example images of the improved spike detection are shown in Figure 5.10. On the other hand, the high error rate observed for plant ID 1817KN409 is due to the morphological reconstruction at the final step. This leads to the fusion of neighboring spikes with the connected stems and leaves, as shown in Figure 5.11.

5.4 Conclusion

Here, we present an improved method for wheat spike detection in a test data set with 369 plants from 202 diverse winter wheat varieties corresponding to mainly high-yielding Central European varieties [17]. Our work relies on the algorithm proposed by Qiongyan et al, which was originally tailored to four Australian wheat varieties. By application to European elite cultivars, that earlier algorithm turned out to be too sensitive to the leaf crossing or overlay artifacts and aged leaves. This resulted in a high rate of false detection of spikes and, consequently, incorrect (too early) detection of heading time points. To overcome these limitations, we developed and evaluated an algorithmic pipeline extended by DWT and the Frangi filter that enable detection and suppression of high-energy regions caused by a high density of leaves. The proposed method has significantly improved the accuracy of the detection of spikes and the time point of heading, resulting in a reduction of the error rate (RMSE) by 75% compared to the Qiongyan et al. model. Similar improvement was also achieved in the analysis of spike growth, where the error rate of model predictions vs. ground truth data was reduced by 50% compared to Qiongyan et al. With these algorithmic improvements, detection of the heading time and analysis of spike growth can be performed in a high-throughput manner with sufficiently high accuracy.

In contrast to the majority of previous method studies, our approach was tested on a diverse set of genotypes with strong morphological differences regarding spike architecture (with or without awns), height, tiller number, biomass, and heading time. Such a data set is very challenging as it is easier to find an algorithm for identifying the plant organs in a small genotype set with much more uniform morphology. However, the biological truth is that many studies employ non-invasive phenotyping to screen genotype collections that exhibit a high phenotypic diversity [32, 19]. This requires algorithms with high performance across a highly heterogeneous background. Our proposed method represents a good starting point, as it correctly determined the heading date in 47 genotypes for both biological replicates and for at least one of the two biological replicates in a further 10 genotypes. These were mainly

plants from lower biomass and tillering range and, therefore, on-average earlier heading. The method is thus expected to perform well in germplasm with relatively low biomass and tillering, as would be the case for collections from hot or dry environments. However, it also clearly showed limitations in genotypes with high biomass and high tillering (mostly later-heading types), where the spike emerges within a green “bush.” The fine-tuning or extension of the developed method for reliable spike detection in such high-biomass, high-tillering genotypes will be conducted in the near future. Further, we aim at application to other existing data sets of spring barley and spring wheat collections, where ground truth data still have to be generated. It is likely that in collections with many or exclusively awned genotypes, the method would already be applicable and yield meaningful results. It is also conceivable that the presented method will work well in bi-parental mapping populations if both parents come from the lower-biomass and tiller-number spectrum.

In conclusion, the proposed approach has the potential to predict the spike yield in other cereal plants such as barley, rice, and rye over time.

In the future, we shall explore the possibility of advancing spike detection methods in an automated manner using deep learning technologies. We also plan to perform a time series analysis of spike growth over a large experimental population (\approx 500 plants) to further improve the algorithm and to deliver more sophisticated solutions for plant breeders and cereal crop researchers.

Data Availability Statement

The datasets generated for this study are available on request to the corresponding author.

Author Contributions

NN and EG conceived, designed and performed the computational experiments, analyzed the data, wrote the paper, prepared figures and tables, and reviewed drafts of the paper. KN and MR executed the laboratory experiments, acquired image data, and reviewed drafts of the paper.

Funding

This work was performed within the German Plant-Phenotyping Network (DPPN) which is funded by the German Federal Ministry of Education and Research (BMBF) (project identification number: 031A053).

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplementary Material

Cultivar	Year_of_relea	class	Quality_rec	Release_st:Hybrid.stat	Group	group in pa
Apertus	2013	2010-13	A	Deutschlan non-hybrid	SK	one
Granada	1980	1980ies	B	Deutschlan non-hybrid	SK	one
Hylux	2012	2010-13	B	Frankreich hybrid	FK	one
Mex. 17 bb	NA			Mexico non-hybrid	FK	one
Sperber	1982	1980ies	A	Deutschlan non-hybrid	SK	one
Alixan	2005	2000nds	A	Frankreich non-hybrid	FK	two
Apache	1997	1990ies	A	CZ non-hybrid	FK	two
Arlequin	2007	2000nds		Frankreich non-hybrid	FK	two
BCD 1302/83	NA			Moldavien non-hybrid	FK	two
Benni multifloret	1980	1980ies	SRW/Pr	USA/Indian non-hybrid	FL	two
Cajeme 71	1971	1970ies	HRS	Mexico non-hybrid	FK	two
Camp Remy	1980	1980ies	B	Deutschlan non-hybrid	FK	two
Centurk	1971	1970ies	HRW/ Ex	USA non-hybrid	FL	two
Helios	2013	2010-13		USA non-hybrid	SK	two
Highbury	1968	1960ies	A	GBR non-hybrid	FK	two
Hope	NA			USA/(S.Dak non-hybrid	FL	two
Hystar	2007	2000nds	B	Frankreich hybrid	FK	two
INTRO 615	NA			USA non-hybrid	FK	two
Isengrain	1996	1990ies	B	FR/SI/ES non-hybrid	FK	two
Ivanka	1998	1990ies	B1	Serbien non-hybrid	FK	two
KWS Ferrum	2012	2010-13	B	Deutschlan non-hybrid	FK	two
Maris Huntsman	1975	1970ies	A	Deutschlan non-hybrid	SK	two
Muskat	2010	2010-13	C	Deutschlan non-hybrid	FK	two
NS 22/92	NA			Serbien non-hybrid	FL	two
NS 46/90	NA			Serbien non-hybrid	FK	two
NS 66/92	NA			Serbien non-hybrid	FK	two
Phoenix	1981	1980ies	HWW/Gd	AUS:New-S non-hybrid	FL	two
Pobeda	1990	1990ies	A1-A2	Serbien non-hybrid	FK	two
Renesansa	1995	1990ies	A2-A1	Serbien non-hybrid	FK	two
Soissons	1987	1980ies	B/A	BE, Es, FR, I non-hybrid	FK	two
Tremie	1991	1990ies		ES, FR, IT, non-hybrid	FK	two
Vel	NA			USA non-hybrid	FL	two
Xanthippe	2011	2010-13	C	Deutschlan non-hybrid	SK	two
Admiral	1968	1960ies	A	Deutschlan non-hybrid	SL	#N/A
Akratos	2004	2000nds	A	Deutschlan non-hybrid	SL	#N/A
Akteur	2003	2000nds	E	Deutschlan non-hybrid	SL	#N/A
Alidos	1987	1980ies	E	Deutschlan non-hybrid	SL	#N/A
Altos	2000	2000nds	E	Deutschlan non-hybrid	SK	#N/A
Alves	2010	2010-13	A	Deutschlan non-hybrid	SL	#N/A
Anapolis	2013	2010-13	C	Deutschlan non-hybrid	SK	#N/A
Anthus	2005	2000nds	B	Deutschlan non-hybrid	SK	#N/A
Apollo	1984	1980ies	C	Deutschlan non-hybrid	SL	#N/A
Aquila	1979	1970ies	C	GRB/IT non-hybrid	SL	#N/A
Ares	1983	1980ies	B	Deutschlan non-hybrid	SL	#N/A
Aristos	1997	1990ies	A	Deutschlan non-hybrid	SL	#N/A
Arktis	2010	2010-13	E	Deutschlan non-hybrid	SK	#N/A

Aron	1992 1990ies	E	Deutschlan non-hybrid SL	#N/A
Asketis	1998 1990ies	A	Deutschlan non-hybrid SL	#N/A
Astron	1989 1980ies	A	Deutschlan non-hybrid SL	#N/A
Aszita	2005 2000nds	B	Deutschlan non-hybrid SL	#N/A
Atomic	2012 2010-13	A	Deutschlan non-hybrid SK	#N/A
Avalon	1980 1980ies	B/Gd	GBR; non-hybrid FK	#N/A
Avenir	2013 2010-13	A	Deutschlan non-hybrid SK	#N/A
Basalt	1980 1980ies	B	Deutschlan non-hybrid SL	#N/A
Batis	1994 1990ies	A	Deutschlan non-hybrid SL	#N/A
Benno	1973 1970ies	E	Deutschlan non-hybrid SL	#N/A
Biscay	2000 2000nds	C	Deutschlan non-hybrid SK	#N/A
Bombus	2012 2010-13	C	Deutschlan non-hybrid SK	#N/A
Boregar	2007 2000nds	A	Frankreich non-hybrid FK	#N/A
Boxer	2013 2010-13	C	Deutschlan non-hybrid SL	#N/A
Brigand	1979 1970ies	BISC; D	GBR non-hybrid SK	#N/A
Brillant	2005 2000nds	A	Deutschlan non-hybrid SK	#N/A
Bussard	1990 1990ies	E	Deutschlan non-hybrid SL	#N/A
Butaro	2009 2000nds	E	Deutschlan non-hybrid SL	#N/A
Capelle Desprez	1946 1940ies	C	FR, CHL, GÈ non-hybrid SL	#N/A
Caphorn	2000 2000nds		Frankreich non-hybrid FK	#N/A
Capone	2012 2010-13	A	Deutschlan non-hybrid SK	#N/A
Cardos	1998 1990ies	A	Deutschlan non-hybrid FK	#N/A
Carenius	2006 2000nds	B	Deutschlan non-hybrid SK	#N/A
Caribo	1968 1960ies	B	Deutschlan non-hybrid SL	#N/A
Carimulti	1975 1970ies	C	Deutschlan non-hybrid SL	#N/A
Carisuper	1975 1970ies	A	Deutschlan non-hybrid SL	#N/A
Chevalier	2005 2000nds	A	AT, CZ, LT, non-hybrid SK	#N/A
Claire	1999 1990ies	C	IE/UK non-hybrid SK	#N/A
Colonia	2011 2010-13	B	DE/BE/HU non-hybrid SK	#N/A
Contra	1990 1990ies	C	Deutschlan non-hybrid SK	#N/A
Cordiale	2003 2000nds		England non-hybrid FK	#N/A
Cubus	2002 2000nds	A	Deutschlan non-hybrid SK	#N/A
Dekan	1999 1990ies	B	Deutschlan non-hybrid SK	#N/A
Desamo	2013 2010-13	B	Deutschlan non-hybrid SK	#N/A
Diplomat	1966 1960ies	A	Deutschlan non-hybrid SL	#N/A
Discus	2007 2000nds	A	Deutschlan non-hybrid SK	#N/A
Disponent	1975 1970ies	A	Deutschlan non-hybrid SL	#N/A
Drifter	1999 1990ies	B	Deutschlan non-hybrid SK	#N/A
Durin	NA		Frankreich non-hybrid SK	#N/A
Edgar	2010 2010-13	B	Deutschlan non-hybrid SK	#N/A
Edward	2013 2010-13	B	Deutschlan non-hybrid SK	#N/A
Einstein	2004 2000nds	B	GB non-hybrid SK	#N/A
Elixer	2012 2010-13	C	Deutschlan non-hybrid SK	#N/A
Ellvis	2002 2000nds	A	Deutschlan non-hybrid SK	#N/A
Enorm	2002 2000nds	E	Deutschlan non-hybrid SK	#N/A
Esket	2007 2000nds	A	Deutschlan non-hybrid SK	#N/A
Estivus	2012 2010-13	A	Deutschlan non-hybrid SK	#N/A

Event	2009 2000nds	E	Deutschlan non-hybrid SK	#N/A
Famulus	2010 2010-13	E	Deutschlan non-hybrid SK	#N/A
Fedor	2007 2000nds	A	Deutschlan non-hybrid SK	#N/A
Flair	1996 1990ies	B	Deutschlan non-hybrid SK	#N/A
Florian	2010 2010-13	E	Deutschlan non-hybrid SK	#N/A
Florida	1984 1980ies		USA non-hybrid FK	#N/A
Forum	2012 2010-13	A	DE/EE/PO/ non-hybrid SK	#N/A
GaUCHo	1993 1990ies		USA non-hybrid SK	#N/A
Genius	2010 2010-13	E	Deutschlan non-hybrid SK	#N/A
Glaucus	2011 2010-13	A	Deutschlan non-hybrid SK	#N/A
Global	2009 2000nds	B	DE/AT non-hybrid SK	#N/A
Gordian	2013 2010-13	B	Deutschlan non-hybrid SK	#N/A
Götz	1978 1970ies	B	Deutschlan non-hybrid SK	#N/A
Gourmet	2013 2010-13	E	Deutschlan non-hybrid SK	#N/A
Greif	1989 1980ies	B	Deutschlan non-hybrid SK	#N/A
Hermann	2004 2000nds	C	Deutschlan non-hybrid SK	#N/A
Herzog	1986 1980ies	A	Deutschlan non-hybrid SK	#N/A
Hybery	2010 2010-13	B	Frankreich hybrid FK	#N/A
Hybred	2003 2000nds	B	DE/FR hybrid FK	#N/A
Hylland	2009 2000nds	B	DE/HU hybrid FK	#N/A
Ibis	1991 1990ies	A	Deutschlan non-hybrid SL	#N/A
Impression	2005 2000nds	A	Deutschlan non-hybrid SK	#N/A
Inspiration	2007 2000nds	B	Deutschlan non-hybrid SK	#N/A
Intro	2011 2010-13	B	Deutschlan non-hybrid SK	#N/A
Jafet	2008 2000nds	E	Deutschlan non-hybrid SK	#N/A
JB Asano	2008 2000nds	A	Deutschlan non-hybrid SK	#N/A
Jenga	2007 2000nds	A	Deutschlan non-hybrid SK	#N/A
Joker	2012 2010-13	A	Deutschlan non-hybrid SK	#N/A
Joss	1972 1970ies	C	Deutschlan non-hybrid SK	#N/A
Kalahari	2010 2010-13	B	DE/BE non-hybrid SK	#N/A
Kanzler	1980 1980ies	B	Deutschlan non-hybrid SL	#N/A
Kerubino	2004 2000nds	E	Deutschlan non-hybrid SK	#N/A
Knirps	1985 1980ies	B	Deutschlan non-hybrid SL	#N/A
Kobold	1978 1970ies	B	Deutschlan non-hybrid SL	#N/A
Kometus	2011 2010-13	A	Deutschlan non-hybrid SK	#N/A
Konsul	1990 1990ies	B	Deutschlan non-hybrid SL	#N/A
Kontrast	1990 1990ies	A	Deutschlan non-hybrid FK	#N/A
Kormoran	1973 1970ies	A	Deutschlan non-hybrid SL	#N/A
Kraka	1982 1980ies	A	Deutschlan non-hybrid SL	#N/A
Kranich	1969 1960ies	A	Deutschlan non-hybrid SK	#N/A
Kredo	2009 2000nds	B	Deutschlan non-hybrid SK	#N/A
Kronjuwel	1980 1980ies	B	Deutschlan non-hybrid SK	#N/A
Kurt	2013 2010-13	B	Deutschlan non-hybrid SK	#N/A
KWS Cobalt	2013 2010-13	A	Deutschlan non-hybrid SK	#N/A
KWS Milaneco	2013 2010-13	E	Deutschlan non-hybrid SL	#N/A
KWS Pius	2010 2010-13	A	Deutschlan non-hybrid SK	#N/A
KWS Santiago	2011 2010-13	C	England non-hybrid SK	#N/A

Lambriego Inia	1980	1980ies		Chile	non-hybrid FK	#N/A
Landsknecht	2013	2010-13	C	Deutschlan	non-hybrid SK	#N/A
Limes	2003	2000nds	B	Deutschlan	non-hybrid SK	#N/A
Linus	2010	2010-13	A	Deutschlan	non-hybrid SK	#N/A
Lucius	2006	2000nds	A	Deutschlan	non-hybrid SK	#N/A
Ludwig	1998	1990ies	A	Deutschlan	non-hybrid SL	#N/A
Magister	2005	2000nds	E	Deutschlan	non-hybrid SK	#N/A
Magnus	2000	2000nds	A	Deutschlan	non-hybrid SL	#N/A
Manager	2006	2000nds	B	Deutschlan	non-hybrid SK	#N/A
Matrix	2010	2010-13	B	Deutschlan	non-hybrid SK	#N/A
Meister	2010	2010-13	A	Deutschlan	non-hybrid SK	#N/A
Memory	2013	2010-13	B	Deutschlan	non-hybrid SK	#N/A
Mentor	2012	2010-13	B	Deutschlan	non-hybrid SK	#N/A
Mex. 3			NA	Mexico	non-hybrid FK	#N/A
Mironovska 808	1963	1960ies	STR/Mediu	Ukraine	non-hybrid FL	#N/A
Monopol	1975	1970ies	E	Deutschlan	non-hybrid SL	#N/A
Mulan	2006	2000nds	B	Deutschlan	non-hybrid SK	#N/A
NaturaStar	2002	2000nds	A	Deutschlan	non-hybrid SL	#N/A
Nelson	2011	2010-13	E	Deutschlan	non-hybrid SK	#N/A
Nimbus	1975	1970ies	B	Deutschlan	non-hybrid FK	#N/A
Oakley	2008	2000nds	C	UK/BE	non-hybrid SK	#N/A
Obelisk	1987	1980ies	B	NE; DE	non-hybrid SK	#N/A
Oberst	1980	1980ies	A	Deutschlan	non-hybrid SL	#N/A
Orcas	2010	2010-13	B	Deutschlan	non-hybrid FK	#N/A
Orestis	1988	1980ies	B	Deutschlan	non-hybrid SK	#N/A
Oxal	2010	2010-13	B	Deutschlan	non-hybrid SK	#N/A
Pantus	1966	1960ies	A	Deutschlan	non-hybrid SK	#N/A
Paroli	2004	2000nds	A	Deutschlan	non-hybrid SK	#N/A
Patras	2012	2010-13	A	Deutschlan	non-hybrid SK	#N/A
Pegassos	1994	1990ies	A	AT, CH, DE,	non-hybrid FK	#N/A
Piko	1994	1990ies	B	Deutschlan	non-hybrid FK	#N/A
Pionier	2013	2010-13	A	Deutschlan	non-hybrid SK	#N/A
Potenzial	2006	2000nds	A	Deutschlan	non-hybrid SK	#N/A
Premio	2006	2000nds	B	Frankreich	non-hybrid FK	#N/A
Primus	2009	2000nds	B	Deutschlan	non-hybrid SK	#N/A
Profilus	2008	2000nds	A	Deutschlan	non-hybrid SK	#N/A
Progress	1969	1960ies	A	Deutschlan	non-hybrid SK	#N/A
Rebell	2013	2010-13	A	Deutschlan	non-hybrid SK	#N/A
Rektor	1980	1980ies	E	Deutschlan	non-hybrid SL	#N/A
Ritmo	1993	1990ies	B	Deutschlan	non-hybrid SK	#N/A
Robigious	2004	2000nds	B	England	non-hybrid SK	#N/A
Rumor	2013	2010-13	B	Deutschlan	non-hybrid FK	#N/A
Saturn	1973	1970ies	C	Deutschlan	non-hybrid SK	#N/A
Schamane	2005	2000nds	A	Deutschlan	non-hybrid SK	#N/A
Severin	1980	1980ies	E	Deutschlan	non-hybrid SK	#N/A
Siete Cerros	1966	1960ies	HWS/Pr Mi	Mexiko	non-hybrid FK	#N/A
Skagen	2006	2000nds	E	Deutschlan	non-hybrid SK	#N/A

Skalmeje		2006 2000nds	C	Deutschlan non-hybrid SK	#N/A
Skater		2000 2000nds	B	Deutschlan non-hybrid SK	#N/A
Sokrates		2001 2000nds	A	Deutschlan non-hybrid SL	#N/A
Solstice		2001 2000nds		England non-hybrid SK	#N/A
Sonalika		1967 1960ies	HWS/Hg	Indien non-hybrid FK	#N/A
Sorbas		1985 1980ies	B	Deutschlan non-hybrid SK	#N/A
Sponsor		1994 1990ies		FR, IE non-hybrid SK	#N/A
SUR99820	NA			Frankreich non-hybrid	#N/A
SW Topper		2002 2000nds	E	Deutschlan non-hybrid SK	#N/A
SY Ferry		2012 2010-13	B	Deutschlan non-hybrid SK	#N/A
Tabasco		2008 2000nds	C	Deutschlan non-hybrid SK	#N/A
Tabor		1979 1970ies	A	Deutschlan non-hybrid SK	#N/A
Tambor		1993 1990ies	A	Deutschlan non-hybrid SL	#N/A
Tarso		1992 1990ies	A	Deutschlan non-hybrid SK	#N/A
Terrier		2001 2000nds	B	Deutschlan non-hybrid SK	#N/A
Tiger		2001 2000nds	A	Deutschlan non-hybrid SL	#N/A
TJB 990-15		1980 1980ies	BISC; D	GBR non-hybrid SK	#N/A
Tobak		2011 2010-13	B	Deutschlan non-hybrid SK	#N/A
Tommi		2002 2000nds	A	Deutschlan non-hybrid SK	#N/A
Topfit		1972 1970ies	B	Deutschlan non-hybrid SL	#N/A
Toronto		1990 1990ies	A	Deutschlan non-hybrid SK	#N/A
Torrild		2005 2000nds	A	Deutschlan non-hybrid SK	#N/A
Transit		1994 1990ies	A	Deutschlan non-hybrid SK	#N/A
Triple dirk \S\''''	NA			Australien non-hybrid FK	#N/A
Tuareg		2005 2000nds	A	Deutschlan non-hybrid SK	#N/A
Türkis		2004 2000nds	A	Deutschlan non-hybrid SK	#N/A
Urban		1980 1980ies	E	Deutschlan non-hybrid SK	#N/A
Vuka		1975 1970ies	A	Deutschlan non-hybrid SL	#N/A
Winnetou		2002 2000nds	C	Deutschlan non-hybrid SK	#N/A
WW 4180 (Kongo)		2012 2010-13		Deutschlan non-hybrid FK	#N/A
Zappa		2009 2000nds	C	Deutschlan non-hybrid SK	#N/A
Zentos		1989 1980ies	E	Deutschlan non-hybrid SL	#N/A
Zobel		2006 2000nds	A	Deutschlan non-hybrid SK	#N/A

Bibliography

- [1] Hans-Joachim Braun, Gary Atlin, Thomas Payne, et al. Multi-location testing as a tool to identify plant response to global climate change. *Climate change and crop production*, 1:115–138, 2010.
- [2] Mark Tester and Peter Langridge. Breeding technologies to increase crop production in a changing world. *Science*, 327(5967):818–822, 2010.
- [3] J. Snape, K. Butterworth, E. Whitechurch, and A. J. Worland. *Waiting for Fine Times: Genetics of Flowering Time in Wheat*, pages 67–74. Springer Netherlands, Dordrecht, 2001.
- [4] Kunpu Zhang, Jichun Tian, Liang Zhao, Bin Liu, and Guangfeng Chen. Detection of quantitative trait loci for heading date based on the doubled haploid progeny of two elite chinese wheat cultivars. *Genetica*, 135(3):257–265, Apr 2009.
- [5] A Witzemberger, H Hack, et al. Explanations of the bbch decimal code for the growth stages of cereals-with illustrations. *Gesunde Pflanzen*, 1989.
- [6] Benedikt Digel, Elahe Tavakol, Gabriele Verderio, Alessandro Tondelli, Xin Xu, Luigi Cattivelli, Laura Rossini, and Maria von Korff. Photoperiod-h1 (ppd-h1) controls leaf size. *Plant Physiology*, 172(1):405–415, 2016.
- [7] Ermias Habte, Lukas M Müller, Munqez Shtaya, Seth J Davis, and Maria von Korff. Osmotic stress at the barley root affects expression of circadian clock genes in the shoot. *Plant, cell & environment*, 37(6):1321–1337, 2014.
- [8] Adel H Abdel-Ghani, Rajiv Sharma, Celestine Wabila, Sidram Dhanagond, Saed J Owais, Mahmud A Duwayri, Saddam A Al-Dalain, Christian Klukas, Dijun Chen, Thomas Lübberstedt, et al. Genome-wide association mapping in a diverse spring barley collection reveals

- the presence of qtl hotspots and candidate genes for root and shoot architecture traits at seedling stage. *BMC plant biology*, 19(1):216, 2019.
- [9] Kerstin Neumann, Yusheng Zhao, Jianting Chu, Jens Keilwagen, Jochen Reif, Benjamin Kilian, and Andreas Graner. Genetic architecture and temporal patterns of biomass accumulation in spring barley revealed by image analysis. *BMC Plant Biology*, 17, 08 2017.
- [10] James Cockram, Huw Jones, Fiona J. Leigh, Donal O’Sullivan, Wayne Powell, David A. Laurie, and Andrew J. Greenland. Control of flowering time in temperate cereals: genes, domestication, and sustainable productivity. *Journal of Experimental Botany*, 58(6):1231–1244, 04 2007.
- [11] Zifeng Guo, Guozheng Liu, Marion S. Röder, Jochen C. Reif, Martin W. Ganal, and Thorsten Schnurbusch. Genome-wide association analyses of plant growth traits during the stem elongation phase in wheat. *Plant Biotechnology Journal*, 16(12):2042–2052, 2018.
- [12] Rafael E. Maeoka, Victor O. Sadras, Ignacio A. Ciampitti, Dorivar R. Diaz, Allan K. Fritz, and Romulo P. Lollato. Changes in the phenotype of winter wheat varieties released between 1920 and 2016 in response to in-furrow fertilizer: Biomass allocation, yield, and grain protein concentration. *Frontiers in Plant Science*, 10:1786, 2020.
- [13] Bi Kun, Jiang Pan, Li Lei, Shi Benyi, and Wang Cheng. Non-destructive measurement of wheat spike characteristics based on morphological image processing. *Transactions of the Chinese Society of Agricultural Engineering*, 2010(12), 2010.
- [14] Bi Kun, Jiang Pan, Tang Chongwei, Huang Feifei, and Wang Cheng. The design of wheat variety bp classifier based on wheat ear feature. *Chinese Agricultural Science Bulletin*, 6:465–467, 2011.
- [15] Kun Bi, Fei-fei Huang, and Cheng Wang. Quick acquisition of wheat ear morphology parameter based on imaging processing. In Yuanxu Yu, Zhengtao Yu, and Jingying Zhao, editors, *Computer Science for Environmental Engineering and EcoInformatics*, pages 300–307, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.
- [16] Li Qiongyan, Jinhai Cai, Bettina Berger, Mamoru Okamoto, and Stanley J. Miklavcic. Detecting spikes of wheat plants using neural networks with laws texture energy. *Plant Methods*, 13(1):83, Oct 2017.

- [17] Kai Voss-Fels, Andreas Stahl, Benjamin Wittkop, Carolin Lichthardt, Sabrina Nagler, Till Rose, Tsu-Wei Chen, Holger Zetzsche, Sylvia Seddig, Mirza Baig, Agim Ballvora, Matthias Frisch, Elizabeth Ross, Ben Hayes, Matthew Hayden, Frank Ordon, Jens Léon, H. Kage, Wolfgang Friedt, and Rod Snowdon. Breeding improves wheat productivity under contrasting agrochemical input levels. *Nature Plants*, 5, 07 2019.
- [18] Christine Zanke, Jie Ling, Joerg Plieske, Sonja Kollers, Erhard Ebmeyer, Viktor Korzun, Odile Argillier, Gunther Stiewe, Maike Hinze, Kerstin Neumann, Martin Ganal, and Marion Röder. Whole genome association mapping of plant height in winter wheat (*triticum aestivum* l.). *PLoS ONE*, 9, 12 2014.
- [19] Sidram Dhanagond, Guozheng Liu, Yusheng Zhao, Dijun Chen, Michele Grieco, Jochen Reif, Benjamin Kilian, Andreas Graner, and Kerstin Neumann. Non-invasive phenotyping reveals genomic regions involved in pre-anthesis drought tolerance and recovery in spring barley. *Frontiers in Plant Science*, 10:1307, 2019.
- [20] Radomir S. Stanković and Bogdan J. Falkowski. The haar wavelet transform: Its status and achievements. *Computers and Electrical Engineering - CEE*, 29:25–44, 01 2003.
- [21] K Bi, P Jiang, L Li, B Shi, and C Wang. Non-destructive measurement of wheat spike characteristics based on morphological image processing. *Transactions of the Chinese Society of Agricultural Engineering(TCSAE)*, 26:212–216, 12 2010.
- [22] Alejandro F. Frangi, Wiro J. Niessen, Koen L. Vincken, and Max A. Viergever. Multiscale vessel enhancement filtering. In William M. Wells, Alan Colchester, and Scott Delp, editors, *Medical Image Computing and Computer-Assisted Intervention — MICCAI'98*, pages 130–137, Berlin, Heidelberg, 1998. Springer Berlin Heidelberg.
- [23] Yanqin Fan. On the approximate decorrelation property of the discrete wavelet transform for fractionally differenced processes. *Information Theory, IEEE Transactions on*, 49:516 – 521, 03 2003.
- [24] Kenneth I Laws. Textured image segmentation. Technical report, University of Southern California Los Angeles Image Processing INST, 1980.
- [25] Tianhorng Chang and C-C Jay Kuo. Texture analysis and classification with tree-structured wavelet transform. *IEEE Transactions on image processing*, 2(4):429–441, 1993.

- [26] Ching-Fen Jiang and Mu-Long Chen. Segmentation of ultrasonic ovarian images by texture features. In *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol. 20 Biomedical Engineering Towards the Year 2000 and Beyond (Cat. No. 98CH36286)*, volume 2, pages 850–853. IEEE, 1998.
- [27] Christina I Christodoulou, Constantinos S Pattichis, Marios Pantziaris, and Andrew Nicolaides. Texture-based classification of atherosclerotic carotid plaques. *IEEE transactions on medical imaging*, 22(7):902–912, 2003.
- [28] Stavroula Gr Mougiakakou, Spyretta Golemati, Ioannis Gousias, Andrew N Nicolaides, and Konstantina S Nikita. Computer-aided diagnosis of carotid atherosclerosis based on ultrasound image statistics, laws’ texture and neural networks. *Ultrasound in medicine & biology*, 33(1):26–36, 2007.
- [29] M. Shahid and I. A. Taj. Robust retinal vessel segmentation using vessel’s location map and frangi enhancement filter. *IET Image Processing*, 12(4):494–501, 2018.
- [30] Attila Budai, Rüdiger Bock, Andreas Maier, Joachim Hornegger, and Georg Michelson. Robust vessel segmentation in fundus images. *International journal of biomedical imaging*, 2013, 2013.
- [31] M Vazquez, N Huyhn, and JM Chang. *Multi-Scale vessel extraction using curvilinear filter-matching applied to digital photographs of human placentas*. PhD thesis, Ph. D. thesis, California State University, Long Beach, 2001.
- [32] Nora Honsdorf, Timothy John March, Bettina Berger, Mark Tester, and Klaus Pillen. High-throughput phenotyping to detect drought tolerance qtl in wild barley introgression lines. *PLOS ONE*, 9(5):1–13, 05 2014.

6. Awn Image Analysis and Phenotyping Using BarbNet

Paper 5: Narisetti N, Awais M, Khan M, Stolzenburg F, Stein N, Gladilin E. Awn Image Analysis and Phenotyping Using BarbNet. *Plant Phenomics* 2023;0081. DOI: 10.34133/plantphenomics.0081

Impact Factor: 6.96 (according to Clarivate releases the Journal Citation Report (JCR 2023))

Abstract

Consideration of the properties of awns is important for the phenotypic description of grain crops. Awns have a number of important functions in grasses, including assimilation, mechanical protection and seed dispersal and burial. An important feature of the awn is the presence or absence of barbs, - tiny hook-like single-celled trichomes on the outer awn surface that can be visualized using microscopic imaging. There are, however, no suitable software tools for the automated analysis of these small, semi-transparent structures in a high-throughput manner. Further, automated analysis of barbs using conventional methods of pattern detection and segmentation is hampered by high variability of their optical appearance including size, shape and surface density. In this work, we present a software tool for automated detection and phenotyping of barbs in microscopic images of awns which is based on a dedicated deep learning model (BarbNet). Our experimental results show that BarbNet is capable of detecting barb structures in different awn phenotypes with an average accuracy of 90%. Furthermore, we demonstrate that phenotypic traits derived from BarbNet-segmented images enable a quite robust categorization of four contrasting awn phenotypes with an accuracy of > 85%. Based on the promising results of this work, we see that the proposed model has potential applications in the automation of barley awns sorting for plant developmental analysis.

6.1 Introduction

Awns are bristle-like extensions of the glumes and/or husks of many grass species including the major crop plants such as wheat, rice, barley, and rye. The awns provide protection against pests and foraging animals. They are considered important due to their physiological role in photosynthesis [1]. Barley plants, for example, typically have long, thin awns extending from the outer husk called “lemma”.

Most awns bear barb-like structures on the surface which primarily aid in the dispersal of the plant’s seeds by facilitating, in the case of wild grass species, the adherence to the animal furs [2]. The barbs are upward-oriented single-celled trichomes with highly silicified cell walls [3] having different sizes and structures. Their presence gives awns a rough texture which can be harmful during manual harvest or lowers barley quality when producing barley as a feed crop. The barbed barley awn is called a “rough” awn. These cultivars bear large and more dense barbs covering the awns from the apical to the basal part. However, the density and size of the barbs vary across the species and cultivars, defining the intensity of awn roughness. Individual barley cultivars lack completely any barbs, except a few at the apical part of the awns, giving the awns a completely smooth surface texture and cultivars are referred to as smooth awn barley.

To look into the genetic control of the intensity of barb formation in barley, a genome-wide association study (GWAS) was conducted on 1,000 barley accessions [4]. Two highly significant loci were associated with the awn roughness trait and one major gene *HvRaw1* could be isolated (Milner et al. 2019).

For the genetic mapping of the second genetic locus and to gain insight into the independent and interactive effect of both loci, we conducted a genetic mapping study using a biparental F2 population derived from a cross between a rough variety “Barke” and a smooth mutant “MHOR597”, both of which were confirmed to carry different allelic states at the two loci controlling awn roughness. By genotyping we confirmed the segregation of all expected nine possible genetic classes (own unpublished data), however, we observed only four different phenotypic classes of awn roughness, bearing barbs of varying density and structure, which is explained by the dominant inheritance of the wild-type alleles of both genes and the additive epistasis between both loci (own unpublished data).

To enable high-throughput screening with the unbiased classification of all different phenotypic classes, we evaluated different techniques of awn imaging to measure barb structure and density. One approach, scanning electron microscopy (SEM) allows to visualize barbs at the highest resolution and a great level of detail [4]. These techniques allow researchers to analyze the size, shape, and arrangement of the barbs on the surface of plant awns, as well as other features such as the presence of glandular structures or the distribution of pigments.

Barbs can also be monitored using conventional light microscopy, however, here they are only distinguishable on the edges of the awns, see Figure 6.2a. For analysis of visible light images, conventional analytic tools such as ImageJ [5], and GIMP [6] can be used. However, accurate segmentation and quantification of small, semi-transparent and often also occluding barb structures are challenging and time expensive. In the absence of appropriate software tools, analysis of barbs in awn images (e.g., counting, measurement) is widely done manually. This approach requires a lot of time and human effort and can be hardly scaled to large datasets. Therefore, automated image processing algorithms are required to detect, quantify and classify the barb structures in different crop plants.

Since we are interested not only in the detection and counting of barbs but also in a more comprehensive assessment of barb morphology, a consistent image segmentation approach was used in this study. The accurate segmentation of barbs like tiny objects is a critical and challenging task because of its limited spatial resolution, low contrast and large variability. Deep learning-based methods have emerged as powerful tools for segmenting small objects like cells [7], nuclei [8] and subcellular structures [9] in biomedical images. Because they can learn complex features and patterns that are difficult to capture with traditional image analysis methods. In the case of plant image analysis, several approaches were proposed for the detection of plant organs like roots [10, 11, 12], shoots [13, 14] and flowers [15] and their high-throughput phenotyping. However, very few studies for automated analysis of small and optically variable plant organs, such as grain spikes [16, 17] were presented in the literature. To the best of our knowledge, no appropriate tools suitable for accurate barb segmentation in awn images are known.

This paper presents a deep learning, convolutional neural network (CNNs) based approach to the segmentation of barbs in microscopic images of barley cultivars. For the task of image segmentation, a well-known encoder-decoder CNN architecture with fully convolutional layers is used. Our approach to

barb segmentation is based on the extension of the U-net segmentation model from Ronneberger et. al [18]. Here, we present our methodological framework including the proposed U-net-based framework for barb detection, ground truth data generation, and training and evaluation procedures. Then, the results of experimental investigations are presented including a model performance by application to the segmentation of test images and genotype-phenotypic analysis.

6.2 Materials and Methods

6.2.1 Plant Material

In this study, we used an F2 population that was segregating for two independent rough awn loci: Raw1 (A) on chromosome 5H and Raw7HS (B) on chromosome 7H, which were previously detected in a GWAS study (Milner et al. 2019). The population was derived from the cross between rough awn variety “Barke”, wild-type at both loci (AABB), and the smooth awned, x-ray mutagenized, mutant named “MHOR597” (GBIS/I, IPK Genebank) carrying the recessive mutant alleles at both loci (aabb). For image analysis, we selected 70 plants based on their either homozygous genotypic constitution at either locus. The plants were grown under controlled greenhouse conditions: 21°C day and 17°C night temperature, with 16 hours duration of artificial light.

6.2.2 Image Data

The roughness of the awn surface is determined by the density of the barbs, which is controlled independently by each of the two mentioned loci. Therefore, we needed to perform robust phenotyping by optically assessing the awns at specific “central” and “basal” locations. To achieve this, we collected three awns from the centre of the main tiller spikes of each plant at the harvest stage and taped them on an imaging slide. We used a digital microscope (Keyence VHX) to capture images of the adaxial side of the awns at 100x magnification (Figure 6.1).

6.2.3 Genotyping

To study the association between barb density and allelic state of the plants at both loci, the genotyping was performed using Kompetitive Allele-Specific

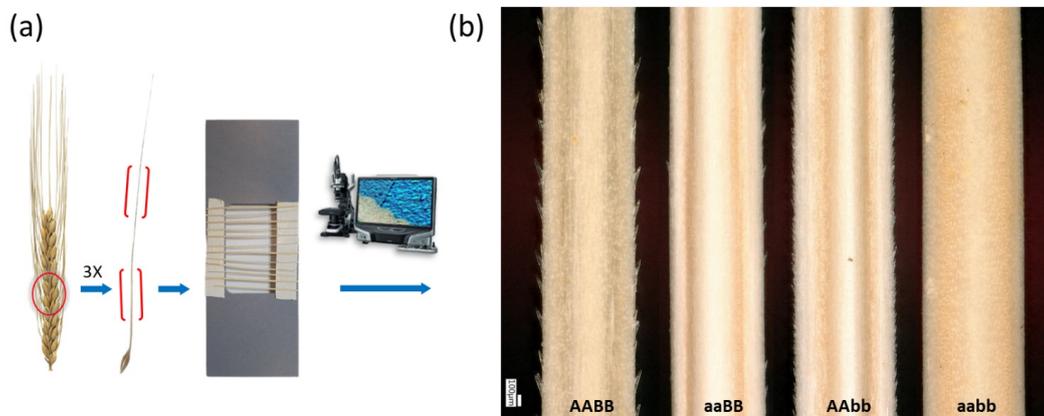


Figure 6.1: Awn imaging (a) The central three awns were collected from the main spike of each plant. The central and basal parts of the awns were taped on an imaging slide and the micrographs were generated under a digital microscope. (b) The representative micrographs of the adaxial side of the basal part of the awns. All possible homozygous genetic classes for the two alternative alleles at the two awn roughness controlling loci (“A” and “B”) show varying density and size of barbs.

PCR (KASP) markers (from PACE® at 3crbio.com), which enable bi-allelic scoring at both loci.

6.2.4 Ground Truth Generation

In order to develop CNN segmentation models, a representative set of ground truth images with an accurate annotation of fore- and background image regions is required. In this study, ground truth images of different microscopic awn images were generated using ImageJ [5]. This tool contains an image-filling option in the toolbar which allows for the efficient annotation or filling of image regions by manually drawing a polygon around each barb in the image. Figure 6.2(b) shows an example of generated ground truth image. Manual annotation of barbs using ImageJ takes between 10-30 minutes per image depending on the number and structural complexity of the barbs of a given awn image.

6.2.5 Barb Detection using CNN

The BarbNet model is a modified version of the original encoder-decoder CNN architecture of U-net [18], which is designed for semantic segmentation

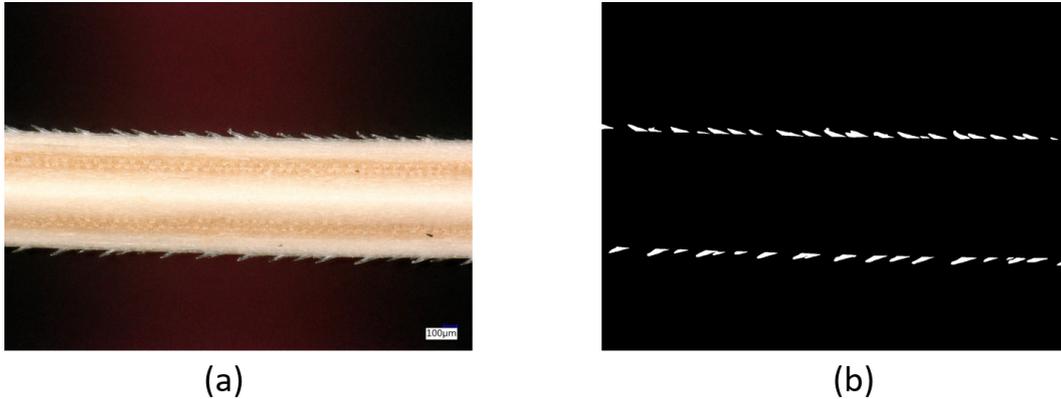


Figure 6.2: Example of barley awn image captured under the microscope. (a) Original image with the resolution of 1200 x 1600 (b) Ground truth mask generated by ImageJ.

of barley awn barbs. Unlike the U-net model, BarbNet includes batch normalization [19] after each convolution layer to enhance network performance and stability by normalizing the feature maps at respective levels [19, 20]. Dropout layers are not used in BarbNet because combining batch normalization and dropout layers can lead to poor results [21]. Additionally, the kernel size is increased to improve the segmentation quality of varying and elongated target patterns [22]. Finally, the depth of BarbNet is increased to 5 compared to the original U-net depth of 4 due to the larger input image size. Table 6.1 provides a detailed comparison of convolutional parameters with respect to the original U-net.

The U-net framework was utilized for the task of barb detection on barley awns by incorporating the suggested modifications. Training and testing phases of the network were conducted on input images at their original resolution of 1024 x 1600.

The encoder network is responsible for extracting features from input image patches. It consists of five encoder blocks, each with two convolutional layers with 5 x 5 filters, followed by batch normalization [19] and ReLU activation function [23]. Max-pooling operations are used to down-sample the feature maps by half of their original size [24, 25]. These steps enable a more efficient aggregation of image features. The bridge encoder block without a max-pooling layer is applied to generate 512 feature maps of size 32 x 50. The decoder network is responsible for upsampling the feature maps and generating the final output. It is composed of four decoder blocks, where the output from the bridge encoder is up-sampled using 3 x 3 transpose

Table 6.1: Convolutional parameters of the original U-net and BarbNet.

Convolutional parameters	Original U-net	BarbNet
Kernel size	3x3	5x5
Transposed kernel size	2x2	3x3
Stride	1x1	2x2
Padding	unpadded	padding with zeros
Depth	4	5
Number of filters	(64, 128, 256, 512, 1028)	(16, 32, 64, 128, 256, 512)

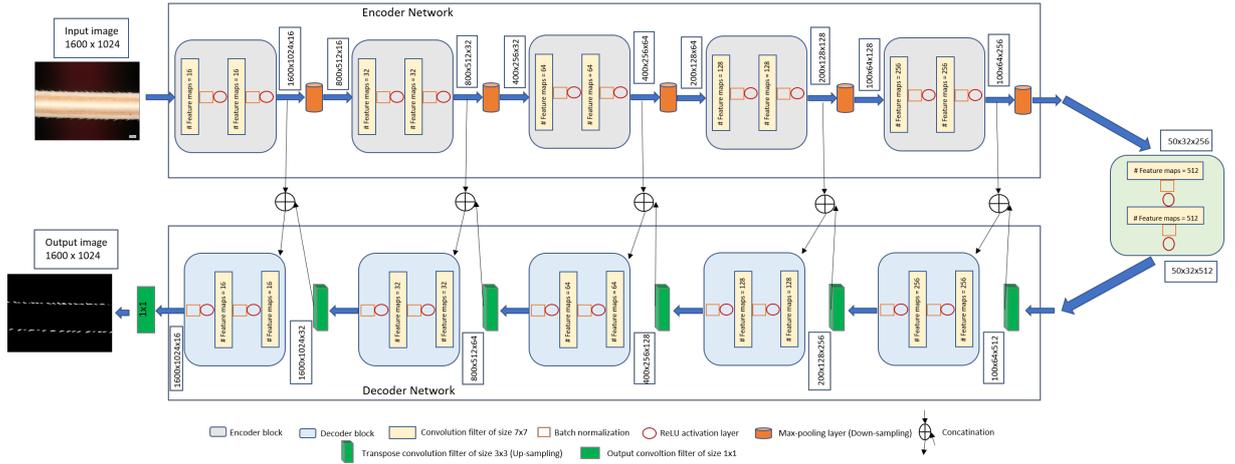


Figure 6.3: The proposed U-net architecture for barb detection on barley awn images.

convolution and stride 2. The resulting feature map is concatenated with the corresponding encoder feature maps and subsequently passed through a convolutional layer with decreasing channel depth (128, 64, 32, and 16). The output of the final decoder block is fed into a convolutional layer with the logistic function [26] to classify each pixel as barb or non-barb in the image. The output of the proposed architecture is a probability image with values ranging between 0 and 1 of size 1024 x 1600, similar to the input image shown in Figure 6.3. Overall, the U-net-based model demonstrates strong

performance in detecting barb pixels in images.

6.2.6 Performance Metrics

During the training and testing phase, the performance of the proposed model (BarbNet) is evaluated using binary cross entropy loss function [25] and the Dice coefficient (DC) [27]. The binary cross-entropy loss function compares each pixel prediction (0: non-barb, 1: barb) with the ground truth pixel and averaged all pixels losses to calculate the total image loss in the training stage.

$$\text{Binary cross entropy loss} = 1/N \sum_{i=1}^N -(Y_i \cdot \log \hat{Y}_i + (1 - Y_i) \cdot \log(1 - \hat{Y}_i)) \quad (6.1)$$

The DC quantifies the similarity between the model’s predicted segmentation and the ground truth segmentation, with values ranging from 0 to 1. A DC value of 0 indicates completely false segmentation (0% similarity), while a DC value of 1 indicates perfect segmentation (100% similarity). It is calculated as 2 times the area of overlap divided by the total number of pixels in both the model predicted and ground truth binary images.

$$DC = \frac{2 * (\hat{Y} \cap Y)}{\hat{Y} \cup Y} = \frac{2 * \sum_{i=1}^N \hat{y}_i y_i}{\sum_{i=1}^N \hat{y}_i + \sum_i y_i} \quad (6.2)$$

The symbols \hat{Y} and Y represent the predicted and ground truth binary images, respectively. The values of \hat{y}_i and y_i correspond to the output of pixel i in the predicted and ground truth binary images, with possible values of either 0 or 1.

6.2.7 Computational implementation

The BarbNet model was developed under Python 3.8 using TensorFlow [28] with Keras API. Furthermore, image processing operations such as reading, cropping, and training data preparation were performed using PIL, Numpy [29], and Scikit-Image [30] packages. The model was trained on a GPU machine with a Linux operating system (Ubuntu 20.04 LTS, Intel(R) Core (TM) i7-10700K CPU @ 3.80GHz) and NVIDIA RTX 3090 with 24GB VRAM graphic card.

As stated above, 348 images were annotated using ImageJ to train a proposed BarbNet model on barely awn images. Afterwards, the dataset was

partitioned into a training set and a validation set in the ratio of 85:15, respectively, based on our experience and literature [31, 32]. Each image is cropped from the original resolution (1200 x 1600) to 1024 x 1600 and normalized in the range of [0, 1] using the min-max method to ensure feature consistency in the CNN network without losing the distribution of original data. The BarbNet’s initial weights were set randomly with zero mean and SD of 0.05, as suggested by Krizhevsky et. al [33]. An Adam optimizer [34] was applied in the model optimization process to improve the segmentation performance on training data. Since the output of the model is a binary segmentation, the binary cross-entropy loss function [25] was used to measure the error rate of the model during the training stage. The model was trained for 100 epochs with a batch size of 12 as per system constraints. The learning rate of the Adam optimizer is initialized with 0.001 and it updates the model weights during each training iteration of the model. To avoid a too-quick convergence of the model to a suboptimal solution, a dynamic learning rate scheduler was introduced to reduce the learning rate by a factor of 0.2 until 0.0001 if the validation loss is not improved in the next 5 iterations. This results in overfitting in the case of a large learning rate and getting stuck on the suboptimal solution in the case of a too-small learning rate can be eliminated [35]. Finally, an early stopping criterion is introduced if training cross-entropy loss is not improved in the next 10 iterations.

In the next step, the optimized BarbNet model was used for the segmentation of barbs structures and subsequent assessment of their phenotypic traits. The MATLAB 2021a routines performing this analysis were compiled to a single license-free executable tool which can be downloaded from electronic Data Archive Library (e!DAL).

The output layer of the model consisted of a logistic activation function, which produced a probability map ranging from 0 to 1 for the segmentation. To convert this probability map into a binary image, a threshold value, denoted as tsh , was chosen. Since the probability of barb pixels is higher than that of background pixels, a threshold value of $tsh \geq 0.5$ was selected to classify all high-probability pixels as barb pixels in the final segmentation. In the post-processing step, segmented objects with an area less than 15 pixels were removed to eliminate false (awn barbs) positives by the phenotypic traits calculation.

6.3 Results

6.3.1 Training and Validation of BarbNet

The training and validation of BarbNet were performed on a set of total 348 images which were subdivided into training and validation subsets with a ratio of 85:15, respectively. The dataset contains different awn phenotypes including barbs of different sizes and densities (smooth, sparse, moderate and dense). During the training stage of the network, the performance of the model was analyzed using binary cross-entropy loss and Dice coefficient (DC) at each epoch. Figure 6.4 shows the training and validation of BarbNet over 75 epochs. It shows that training loss (Figure 6.4a) was minimized and platen the curve 50 epochs. Simultaneously, training DC (Figure 6.4b) achieved more than 0.91 from epoch number 50. However, the generalized performance of the optimized model is analyzed using validation metrics. The BarbNet model achieved a maximum validation Dice coefficient of 0.91 and a minimum validation loss of 0.0076 at epoch number 61. Thenceforth, model training was not improved and terminated at epoch number 72.

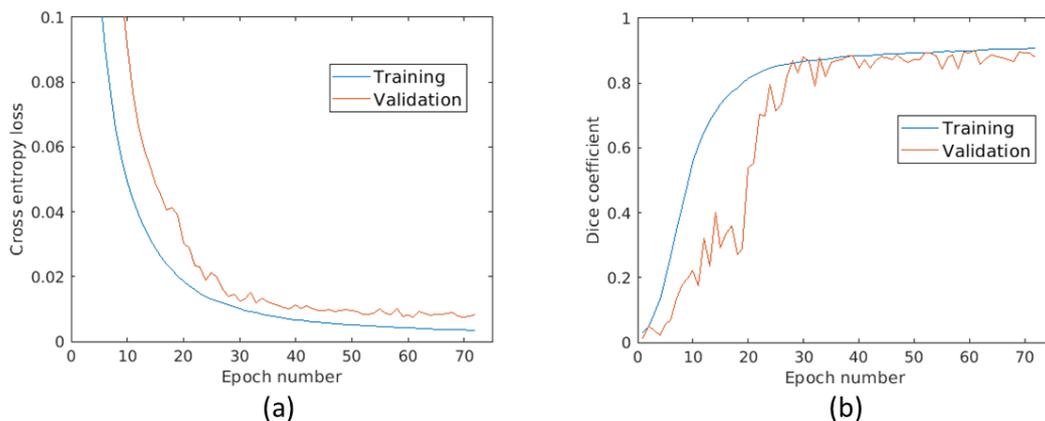


Figure 6.4: Training and validation performance of the BarbNet model over 75 epochs with respect to (a) the cross entropy loss and (b) the Dice coefficient. X- and Y-axes represent the epoch number and performance measure, respectively.

6.3.2 Evaluation and Comparison of U-net Models

The BarbNet model represents an extension of the U-net framework from [18]. In order to adapt the U-net to the task of barb (small object) segmentation several modifications were introduced including (i) inclusion of batch

normalization (original U-net + BN); (ii) exclusion of dropout layers (original U-net + BN + No Dropout); (iii) increased kernel size of 5 (original U-net + BN + No Dropout + K5) and (iv) increased depth of the model to 5 and reduced number of filters (original U-net + BN + No Dropout + K5 + D5). The last U-net modification (iv) turned out to show the best performance of barb segmentation is further termed as BarbNet. For the evaluation and comparison of these models, the same image sets for training and validation were used. Figure 6.5 shows a comparison of performance of all four modified U-net models vs the original U-net for the task of barb segmentation. It shows that the validation loss of both "original U-net + BN + No Dropout + K5" and BarbNet decreases after epoch number 40 and converges earlier than the remaining three U-net models with a higher validation loss. However, BarbNet outperforms all the other U-net models with respect to the Dice accuracy measure of image segmentation beginning from the epoch number 30. The details on validation metrics for all four modified and original U-net models are summarized in Table 6.2.

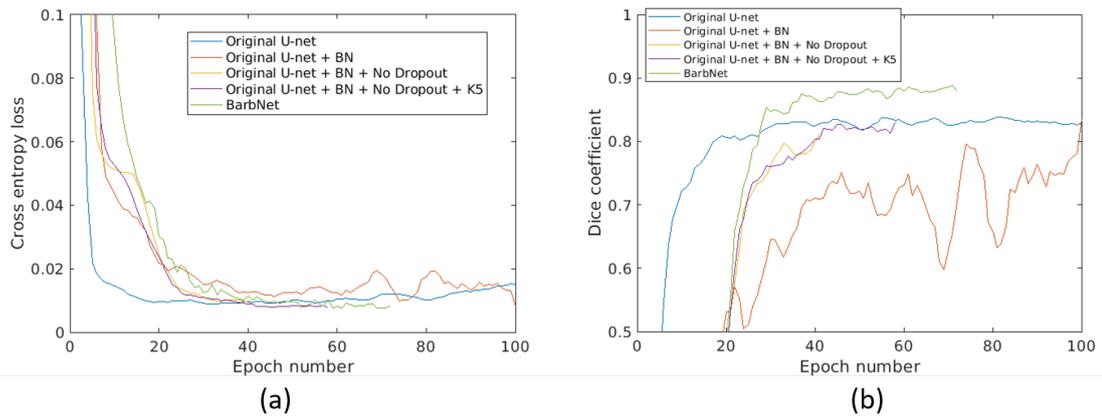


Figure 6.5: Comparison of performance metric of BarbNet vs U-net with respect to (a) the cross entropy loss and (b) the Dice coefficient. X- and Y-axes represent the epoch number and performance measure, respectively.

The performance of all CNN segmentation models is also evaluated on a test set of completely unseen 19 images that were not used for model training, see Table 6.3. It shows that BarbNet outperforms the original as well as all other modified U-net models with a mean accuracy of more than 90% for all test images. Examples of BarbNet segmented images including smooth (image number 2), sparse (image number 9), moderate (image number 11) and dense (image number 14) barb distributions can be found in Supplementary information, Figure S1, S2, S3 and S4 respectively. Thereby, the largest

Table 6.2: Comparison of performance of four modified vs. original U-net model over 100 training epochs with respect to the cross entropy loss and the Dice coefficient.

Model	Cross entropy loss	Dice coefficient
Original U-net	0.0081	0.85
Original U-net+BN	0.0082	0.84
Original U-net+BN+No Dropout	0.0086	0.82
Original U-net+BN+No Dropout+K5	0.0070	0.85
BarbNet	0.0076	0.91

differences in performance between BarbNet and other U-net models were observed for the smooth awn phenotype. The average computational time required for the BarbNet to segment a 1600 x 1024 test image on a PC with Intel(R) Xeon(R) CPU E5-2640 2.40GHz CPU is 1.10 seconds.

Table 6.3: Comparison of mean Dice coefficient of U-net models on a test dataset of 19 images.

Model	Mean Dice coefficient
Original U-net	0.82
Original U-net+BN	0.86
Original U-net+BN+No Dropout	0.80
Original U-net+BN+No Dropout+K5	0.83
BarbNet	0.93

6.3.3 Evaluation of Phenotypic Traits

In addition to segmentation performance, phenotypic properties of awns calculated using our automated segmentation models were compared to manually segmented (ground truth) data. For this purpose, 10 phenotypic fea-

tures are proposed in this study. Further information on the definition of traits is included in the Supplementary Information, see Table S1. Out of 10 traits, only 3 important features for awns and barbs characterization named total barb count, mean barb area and mean barb length in pixels are presented. Figure 6.6 shows the correlation between the ground truth (x-axis) and predicted barbs count (y-axis) over 156 awn images. Differently from the above-mentioned set of ground truth segmented 348 images, this dataset is independently prepared by biologists for evaluation of barb count predictions. Here, the p-value represents that the BarbNet is a highly significant (< 0.05) model which exhibits a higher correlation coefficient of determination (R-squared) value of 0.86. It indicates that BarbNet exhibits 86% conformity between the ground truth and predicted barbs count.

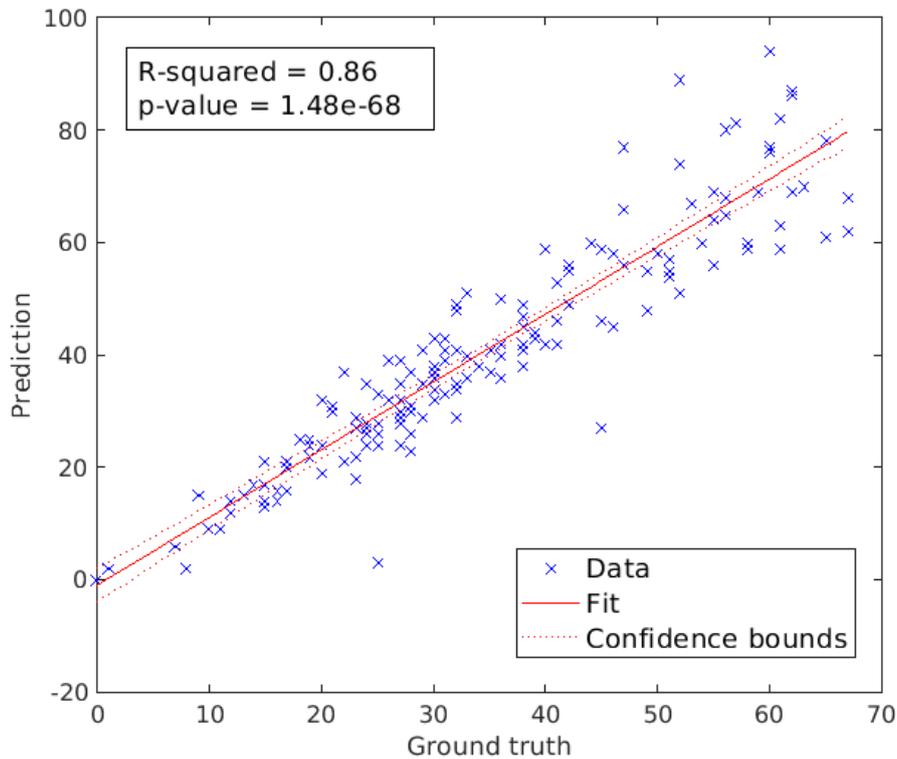


Figure 6.6: Correlation between the ground truth (x-axis) and predicted (y-axis) barbs count over 156 awn images. Each point represents the total number of barbs per awn image. The red colour solid line and dotted lines represent a fitted curve and 95% confidence bounds, respectively. The R-squared value indicates good conformity between ground truth and BarbNet count of image segmentation and trait calculation.

6.3.4 Genotypic-Phenotypic Classification

Four major awn phenotypes were observed in this study including smooth (aabb) sparse (aaBB), moderate (AABB) and dense (AAbb) on the basis of the density and size of the barbs. Here, A and B are two independently segregating (unlinked) genes each represented by a dominant wild-type (capital letter) and a recessive mutant (small letter) allele. In an F2 progeny, the genes may segregate among others into four distinct, fully homozygous genotypic classes i.e., AABB, aaBB, AAbb or aabb. In total, 326 awn images of basal (176) and central (150) regions from these respective four genotypes were used. Further details on data distribution among the four genotypic classes are listed in Table 6.4 and 6.5.

Next, we tested how accurately awn regions (basal or central) can be assigned to one of these four genetic classes using features derived from BarbNet-segmented images. For this purpose, 10 phenotypic features and their pair-wise combinations were compared. The classification was performed using unsupervised k-means clustering after normalizing the trait values in between 0 and 1 using the min-max method, and subsequently starting with estimates of cluster centroids obtained from original density clusters as shown in Figure 6.7. Our proposed prediction model accurately clustered the phenotype, confirming the precision of our predictions in classifying the phenotype into the expected number of clusters defined by the genotype. The best-performing pair of features turn out to be total barb count and barb area (average accuracy of 86%) and total barb count and barb length (average accuracy of 88%) on awns central regions, see Table 6.4 and 6.5.

Table 6.4: F1-score of four genotypes with count and area features.

Count vs Area	Basal (# of images)	Central (# of images)
aabb	0.89 (36)	0.91 (35)
aaBB	0.68 (62)	0.85 (60)
AABB	0.90 (40)	0.81 (32)
AAbb	0.75 (38)	0.85 (23)
Mean F1-score	0.81 (176)	0.86 (150)

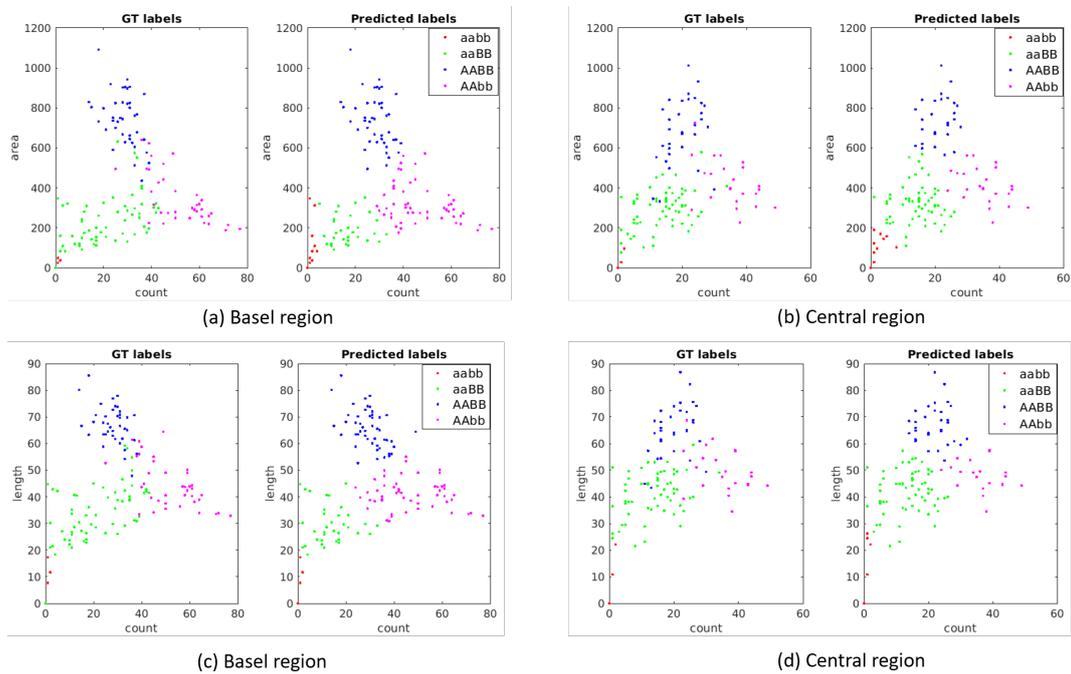


Figure 6.7: Clustering of phenotypic traits on four different barb distributions (smooth, sparse, moderate and dense) using k-means. (a) basal region: Count vs Area, mean F1 score: 0.81 (b) central region: Count vs Area, mean F1 score: 0.86 (c) basal region: Count vs Length, mean F1 score: 0.81 (d) central region: Count vs Length, mean F1 score: 0.88. These four genotypes were well clustered with an F1-score of $> 85\%$ on central regions of awns in both phenotypic traits. Area and length features are measured in pixels.

Table 6.5: F1-score of 4 genotypes with count and length features.

Count vs Length	Basal (# of images)	Central (# of images)
aabb	0.96 (36)	0.97 (35)
aaBB	0.72 (62)	0.90 (60)
AABB	0.87 (40)	0.83 (32)
AAbb	0.69 (38)	0.82 (23)
Mean F1-score	0.81 (176)	0.88 (150)

6.4 Discussion & Conclusion

Awns are important for the phenotypic description of grain crops. But accurate detection of barbs on awn surfaces is challenging because of their limited spatial resolution, low contrast and large variability. Taking advantage of the latest advancements in deep learning, the U-net model was considered and applied to perform automatic analysis of awns in a high-throughput manner. However, such direct transfer led to unsatisfactory performance because the original U-net model was not designed for the awn image analysis. Hence, we modified the U-net architecture to fully automated the segmentation of barbs in microscopic images. Our experimental results have shown a remarkable accuracy of the modified U-net model (BarbNet) for fully automated segmentation of barbs of different sizes, shapes and densities.

The training and validation performance of the BarbNet model has improved over the number of epochs. Besides, the model before epoch number 30 under-performed and showed the worst performance in the validation stage. However, due to the dynamic reduction in the learning rate by a factor of 0.2 a stable performance of approximately 90% Dice coefficient is achieved from epoch number 50. Finally, the optimized model with maximum validation DC 0.91 and minimum cross-entropy loss 0.0076 at epoch number 61 is saved for the automated awn phenotyping.

Comparison of the original vs. four modified U-net models lead to a selection of the most accurate and robust one (BarbNet) crossover all awn phenotypes. Thereby, it was observed that the original U-net as well as the "original U-net + BN" models failed to converge within the first 100th epochs. We draw the reduced accuracy of these two models to the presence of dropout layers removing small or fine-grained structures with low probability values which effectively reduces the capacity of the model. This results in low validation Dice coefficient values, especially for the "original U-net + BN" model. On the other side, three other models with only batch normalization but without dropout layers turn out to exhibit a rapid convergence. Because batch normalization normalises the activations, gradients of the training model are stabilized which reduces the likelihood of vanishing or exploding gradients. As a result, a more stable gradient flow enables smoother and more efficient optimization, leading to faster convergence of these models.

In addition, by reducing the number of filters with increased depth of the model, BarbNet focuses on extracting more localized and detailed features

from the overall large awn images. It results in the DC of BarbNet being improved by 6% compared to the other U-net models (Table 6.2). This is also reflected in the higher DC of the BarbNet model (≥ 90) on all unseen awn images with different barb distributions compared to other U-net models. Consequently, all four modified U-net models but, in particular, BarbNet exhibit a superior accuracy and robustness by segmentation of barbs crossover different awn phenotyping.

Furthermore, phenotypic characterization and classification of awns show a highly significant correlation for the barb's count indicating that barb detection and phenotyping using BarbNet is practically very near to the human-supervised one. However, the difference in count estimation between ground truth and BarbNet might result from occluding neighbour barbs. This may lead to a reduction in barb count per image compared to the ground truth.

Since accurate pixel-wise segmentation and not just a region detection model was used in our approach, the genotype of segmented barbs can be accurately and comprehensively characterized in terms of different phenotypic features. Our results on genotypic-phenotypic classification indicate distinct clustering of all four observed phenotypic classes and show its correspondence with the associated genotypic class. Moreover, it indicates that both genes contribute independently to the awn roughness trait: gene "A" controls the barb density, while gene "B" regulates the barb size. Furthermore, the location of the pink and blue clusters on each graph suggests that gene "A" plays a more significant role in defining the awn roughness trait.

Similarly, our experiments with different genotypes explain that a combination of barb density and size provides the best results on central regions of awns (i.e., $> 85\%$) than basal regions (81%). In particular, the classification accuracy of genotype aaBB is improved by more than 15%. Because the basal regions of this genotype are smoother with almost no barbs than the central region and this results in less classification accuracy by unsupervised k-means algorithm. However, the accuracy reduction of 5% in the central regions of genotype AABB is due to the false positives that occurred between the moderate and dense phenotypes. Since awns have tightly connected neighbourhood barbs in the dense phenotype they tend to have higher lengths similar to the moderate barb density. Therefore, one can achieve the best and most balanced results when features of barb (area and length) and features of awns (count) are combined to perform phenotype-genotype analysis.

In conclusion, our approach provides an efficient solution with 90% accuracy for the detection of barbs in barley awn images. Our model turns out to be robust enough for the detection of barbs in different awn phenotypes. Even though our model exhibits quite accurate detection rates, the topic of barb and, more generally, small organ detection is still an emergent topic in the broader field of plant phenotyping. Presently, our model occasionally ignores tiny barbs, which can be attributed to the large-scale downsampling and network depth. In addition, neighbour barbs in the dense phenotypes are occluded together, leading to larger error rates in phenotypic traits. These limitations may be overcome by extending the training set of ground truth images, especially in dense and sparse phenotypes. Furthermore, alternative segmentation and region detection CNNs such as DeepLab and YOLO can be taken into consideration for capturing small objects such as barbs and other plant organs.

Acknowledgments

General

Thanks to our technicians Mary Ziems and Archana Chelamalashetty for helping in data generation.

Author Contributions

N.N. and E.G. conceived the computational study, performed data analysis, prepared figures and tables, wrote the manuscript. M.A. performed biological experiments, wrote and corrected the manuscript. M.K. performed biological experiments. N.S. conceived the biological study, read and reviewed the manuscript. F.S. co-supervised the study, read and reviewed the manuscript. All authors reviewed the manuscript.

Funding

The research was financially supported by a grant of the German Research Association DFG (STE 1102/17-1) to NS.

Competing interests

The authors declare that there is no conflict of interest regarding the publication of this article.

Data availability

The datasets analyzed and source code during the current study are available from the corresponding author upon reasonable request.

Supplementary Material

Supplementary Material

Awn image analysis and phenotyping using BarbNet

Narendra Nariseti, Muhammad Awais, Muhammad Khan , Frieder Stolzenburg, Nils Stein, and Evgeny Gladilin

Table S1. Description barb traits.

Trait Name	Description
Total Count	Total number detected barbs per awn image.
Lower Count	Number of detected barbs in the lower edge of the awn.
Upper Count	Number of detected barbs in the upper edge of the awn.
Area	Number of pixels of segmented barbs. It measures the size of barbs.
Length	Major axis length of barbs approximated by fitting ellipse to the barb object. It measures the elongated length of each barb.
Circularity	Measures the roundness of segmented barbs. It indicates the shape of barbs like circle or elliptical.
Perimeter	Measures the length of barbs boundary. It indicates how large is the size of barbs.
Compactness	Ratio of area of barb to the area of ellipse. It measures how irregular are the boundaries of barbs.
Lower_dist	Euclidian distance between the neighborhood barbs in the lower edge of the awn. It measures how close barbs are places on the awn.
Upper_dist	Euclidian distance between the neighborhood barbs in the upper edge of the awn. It measures how close barbs are places on the awn.

Figure S1. Examples of barb segmentation on smooth awn phenotype using original U-net and BarbNet: (a) original image, (b) segmentation using the original U-net vs. ground truth (DC=0), (c) segmentation using BarbNet vs ground truth (DC = 0.83). For the visualization purpose, original images were cropped to the region of interest.



(a)



(b)



(c)

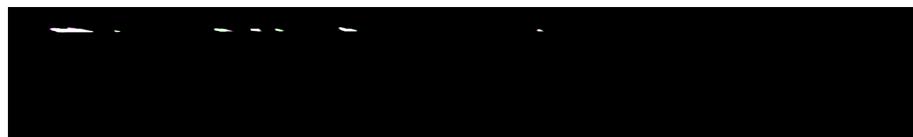
Figure S2. Examples of barb segmentation on sparse awn phenotype using original U-net and BarbNet: (a) original image, (b) segmentation using the original U-net vs. ground truth (DC=0.83), (c) segmentation using BarbNet vs ground truth (DC = 0.94). For the visualization purpose, original images were cropped to the region of interest.



(a)



(b)

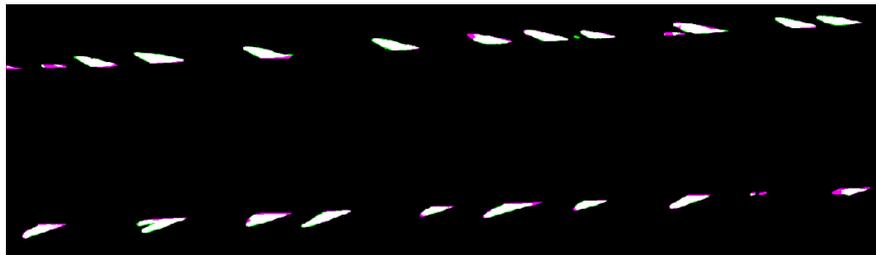


(c)

Figure S3. Examples of barb segmentation on moderate awn phenotype using original U-net and BarbNet: (a) original image, (b) segmentation using the original U-net vs. ground truth (DC=0.83), (c) segmentation using BarbNet vs ground truth (DC = 0.94). For the visualization purpose, original images were cropped to the region of interest.



(a)



(b)



(c)

Figure S4. Examples of barb segmentation on dense awn phenotype using original U-net and BarbNet: (a) original image, (b) segmentation using the original U-net vs. ground truth (DC=0.88), (c) segmentation using BarbNet vs ground truth (DC = 0.95). For the visualization purpose, original images were cropped to the region of interest.



(a)



(b)



(c)

Bibliography

- [1] Tilahun Abebe, Roger P Wise, and Ronald W Skadsen. Comparative transcriptional profiling established the awn as the major photosynthetic organ of the barley spike while the lemma and the palea primarily protect the seed. *The Plant Genome*, 2(3), 2009.
- [2] Rivka Elbaum, Liron Zaltzman, Ingo Burgert, and Peter Fratzl. The role of wheat awns in the seed dispersal unit. *Science*, 316(5826):884–886, 2007.
- [3] Eva Kondorosi, François Roudier, and Emmanuel Gendreau. Plant cell-size control: growing by ploidy? *Current opinion in plant biology*, 3(6):488–492, 2000.
- [4] Sara G Milner, Matthias Jost, Shin Taketa, Elena Rey Mazón, Axel Himmelbach, Markus Oppermann, Stephan Weise, Helmut Knüpffer, Martín Basterrechea, Patrick König, et al. Genebank genomics highlights the diversity of a global barley collection. *Nature genetics*, 51(2):319–326, 2019.
- [5] Chong Huang, Michael F Becker, John W Keto, and Desiderio Kovar. Annealing of nanostructured silver films produced by supersonic deposition of nanoparticles. *Journal of Applied Physics*, 102(5):054308, 2007.
- [6] The GIMP Development Team. Gimp.
- [7] Yousef Al-Kofahi, Alla Zaltsman, Robert Graves, Will Marshall, and Mirabela Rusu. A deep learning-based algorithm for 2-d cell segmentation in microscopy images. *BMC bioinformatics*, 19(1):1–11, 2018.
- [8] Hemaxi Narotamo, J Miguel Sanches, and Margarida Silveira. Segmentation of cell nuclei in fluorescence microscopy images using deep learning. In *Pattern Recognition and Image Analysis: 9th Iberian Conference, IbPRIA 2019, Madrid, Spain, July 1–4, 2019, Proceedings, Part I*, pages 53–64. Springer, 2019.

- [9] Arif Ahmed Sekh, Ida S Opstad, Gustav Godtlielsen, Åsa Birna Birgisdottir, Balpreet Singh Ahluwalia, Krishna Agarwal, and Dilip K Prasad. Physics-based machine learning for subcellular segmentation in living cells. *Nature Machine Intelligence*, 3(12):1071–1080, 2021.
- [10] Tao Wang, Mina Rostamza, Zhihang Song, Liangju Wang, G McNickle, Anjali S Iyer-Pascuzzi, Zhengjun Qiu, and Jian Jin. Segroot: a high throughput segmentation method for root image analysis. *Computers and Electronics in Agriculture*, 162:845–854, 2019.
- [11] Narendra Narisetti, Michael Henke, Christiane Seiler, Rongli Shi, Astrid Junker, Thomas Altmann, and Evgeny Gladilin. Semi-automated root image analysis (saria). *Scientific reports*, 9(1):1–10, 2019.
- [12] Narendra Narisetti, Michael Henke, Christiane Seiler, Astrid Junker, Jörn Ostermann, Thomas Altmann, and Evgeny Gladilin. Fully-automated root image analysis (faria). *Scientific Reports*, 11(1):1–15, 2021.
- [13] Ildar Rakhmatulin, Andreas Kamilaris, and Christian Andreasen. Deep neural networks to detect weeds from crops in agricultural environments in real-time: a review. *Remote Sensing*, 13(21):4486, 2021.
- [14] Narendra Narisetti, Michael Henke, Kerstin Neumann, Frieder Stolzenburg, Thomas Altmann, and Evgeny Gladilin. Deep learning based greenhouse image segmentation and shoot phenotyping (deepshoot). *Frontiers in Plant Science*, 13, 2022.
- [15] Kun Hu, Zhiyong Wang, Guy Coleman, Asher Bender, Tingting Yao, Shan Zeng, Dezhen Song, Arnold Schumann, and Michael Walsh. Deep learning techniques for in-crop weed identification: A review. *arXiv preprint arXiv:2103.14872*, 2021.
- [16] Sajid Ullah, Michael Henke, Narendra Narisetti, Klára Panzarová, Martin Trtílek, Jan Hejato, and Evgeny Gladilin. Towards automated analysis of grain spikes in greenhouse images using neural network approaches: A comparative investigation of six methods. *Sensors*, 21(22):7441, 2021.
- [17] Tanuj Misra, Alka Arora, Sudeep Marwaha, Viswanathan Chinnusamy, Atmakuri Ramakrishna Rao, Rajni Jain, Rabi Narayan Sahoo, Mrinmoy Ray, Sudhir Kumar, Dhandapani Raju, et al. Spikesegnet-a deep learning approach utilizing encoder-decoder network with hourglass for

- spike segmentation and counting in wheat plant from visual imaging. *Plant methods*, 16(1):1–20, 2020.
- [18] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.
- [19] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*, 2015.
- [20] Shibani Santurkar, Dimitris Tsipras, Andrew Ilyas, and Aleksander Madry. How does batch normalization help optimization?, 2019.
- [21] Xiang Li, Shuo Chen, Xiaolin Hu, and Jian Yang. Understanding the disharmony between dropout and batch normalization by variance shift, 2018.
- [22] Chao Peng, Xiangyu Zhang, Gang Yu, Guiming Luo, and Jian Sun. Large kernel matters – improve semantic segmentation by global convolutional network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.
- [23] Forest Agostinelli, Matthew Hoffman, Peter Sadowski, and Pierre Baldi. Learning activation functions to improve deep neural networks. *arXiv preprint arXiv:1412.6830*, 2014.
- [24] Limin Wang, Sheng Guo, Weilin Huang, and Yu Qiao. Places205-vggnet models for scene recognition. *arXiv preprint arXiv:1508.01667*, 2015.
- [25] Ranjeet Ranjan Jha, Gaurav Jaswal, Divij Gupta, Shreshth Saini, and Aditya Nigam. Pixisegnet: Pixel-level iris segmentation network using convolutional encoder-decoder with stacked hourglass bottleneck. *IET Biometrics*, 9(1):11–24, 2019.
- [26] Rob A Dunne and Norm A Campbell. On the pairing of the softmax activation and cross-entropy penalty functions and the derivation of the softmax activation function. In *Proc. 8th Aust. Conf. on the Neural Networks, Melbourne*, volume 181, page 185. Citeseer, 1997.
- [27] Kelly H Zou, Simon K Warfield, Aditya Bharatha, Clare MC Tempany, Michael R Kaus, Steven J Haker, William M Wells III, Ferenc A Jolesz, and Ron Kikinis. Statistical validation of image segmentation quality

- based on a spatial overlap index1: scientific reports. *Academic radiology*, 11(2):178–189, 2004.
- [28] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Craig Chen Zhifeng Citro, Greg S Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467*, 2016.
- [29] Stéfan van der Walt, S Chris Colbert, and Gael Varoquaux. The numpy array: a structure for efficient numerical computation. *Computing in science & engineering*, 13(2):22–30, 2011.
- [30] Stefan Van der Walt, Johannes L Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D Warner, Neil Yager, Emmanuelle Gouillart, and Tony Yu. scikit-image: image processing in python. *PeerJ*, 2:e453, 2014.
- [31] Alessandro Crimi, Spyridon Bakas, Hugo Kuijf, Bjoern Menze, and Mauricio Reyes. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: Third International Workshop, BrainLes 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 14, 2017, Revised Selected Papers*, volume 10670. Springer, 2018.
- [32] V. Roshan Joseph. Optimal ratio for data splitting. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, n/a(n/a), 2022.
- [33] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.
- [34] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [35] Yoshua Bengio. Practical recommendations for gradient-based training of deep architectures. In *Neural networks: Tricks of the trade*, pages 437–478. Springer, 2012.

7. Synthesis of Results

This cumulative dissertation is based on five peer-reviewed publications. All these articles were published in international scientific journals with a total impact factor of about 30 (as of January 2024). All articles followed the primary goal of the work, which is to develop automatic plant image segmentation and phenotyping solutions for high-throughput plant phenotyping. This chapter aims to summarize the state-of-the-art methods, method improvements, and main results of five publications, synthesize the individual results into broader conclusions, discuss the contributions, implications and limitations of the performed research, and provide an outlook of future directions.

Image-based plant phenotyping is crucial in modern agriculture and plant science to improve crop yield by studying phenotypic traits. However, the effectiveness and precision of trait computation depend critically on the segmentation of target plant organs (root, shoot, spike and barbs in awns) in the images. Furthermore, the appearance of plant organs varies dynamically during plant growth. Therefore, more advanced and automatic image processing solutions are required to process and quantify large amounts of varying image data. The presented work approached the problem by developing advanced computer vision solutions for root and shoot image segmentation, Wheat spike detection and barb segmentation in Barley awns for high-throughput plant phenotyping. The summary of each application of plant phenotyping addressed in this thesis is given below.

7.1 Addressing Research Objectives

7.1.1 Objective 1

Development of a semi-automated root image segmentation algorithm and phenotyping tool using traditional image processing algorithms and evaluating the tool's performance with state-of-the-art tools.

Currently, most root image analysis tools (WinRHIZO [22], EZ-RHIZO [23], SmartRoot[24] and, IJ_Rhizo [25])) are tailored to artificial mediums, such as roots grown in agar gels and scanned roots. With a single threshold value, we can segment images captured under artificial mediums with contrasting backgrounds different from soil. However, this solution cannot segment inhomogeneous and contrast varying roots growing under heterogeneous and noisy soil backgrounds (soil-root images). Moreover, advanced deep-learning solutions are unavailable because there are no publicly available datasets and ground truth annotations on soil-root images.

To address the above limitations and consequently to develop deep learning-based solutions, first, a semi-automated root image analysis (saRIA) method based on adaptive thresholding [53] and morphological filtering was developed. The proposed method in saRIA tolerates global inhomogeneity of image intensity, such as vertical image gradient in our soil-root images, by calculating pixel-wise threshold based on its local neighbourhood compared to the global thresholding methods in IJ_Rhizo (bi-level threshold) [25] and GiAroots (single-level threshold) [54]. The roots in the segmented image are elongated line-/curve-like structures that differ from non-root blobs concerning their area, length, and shape (i.e., eccentricity). They are removed by applying morphological filtering, lacking in the above state-of-the-art methods. This method performs rapid segmentation and quantification of roots under natural (soil) and artificial environments (agar and scanned mediums) by adjusting a small set of algorithmic parameters.

The performance of saRIA is measured in both image segmentation and phenotypic traits using dice similarity coefficient (DC) and correlation coefficient of determination (R^2), respectively, concerning state-of-the-art methods and/or tools. Results showed that automatic root image segmentation of saRIA over 100 images achieved more than 80% accuracy compared to the global thresholding methods in IJ_Rhizo and GiAroots. It concludes that root structures in soil background are highly inhomogeneous with varying contrast, and segmentation of these structures requires different threshold values for different parts of the image. However, selecting morphological parameters makes it difficult to automate the image segmentation of large and small root structures. Because of small root architectures from the early growth stage, plants are sensitive to noisy objects similar to the roots in the image. They also, reflected this result in the phenotypic traits, including root length, surface area and volume calculation. It indicates that root traits obtained using saRIA achieved more than 80% correlation with p-value <0.05 compared to the manually calculated traits using SmartRoot [24]. Neverthe-

less, these limitations can be overcome by tuning parameters and removing noise manually for each image in the saRIA tool.

The saRIA segmentation, root tracing and trait calculation algorithms are significantly faster than conventional manual segmentation [24] and are freely available as a precompiled executable program ¹. Therefore, saRIA can be used to generate ground-truth segmentation data to develop highly rapidly demanded advanced machine learning/deep learning techniques.

7.1.2 Objective 2

Research on CNN models for the segmentation of roots from soil back-ground and compare the performance of the proposed model with the state-of-the-art CNN models for root image analysis.

Semi-automatic image analysis solutions effectively analyze single or fewer images to detect and extract multiple root traits. In real-time scenarios, biologists often conduct experiments over several plants for multiple days, resulting in the need to analyze a large number (several thousand) of images. However, semi-automatic approaches face significant challenges when dealing with large image datasets. They require manual algorithmic parameter selection, which hampers their ability to automatically segment large image datasets with inhomogeneous and noisy root system architectures. These limitations underscore the necessity for advanced computer vision solutions, specifically those based on convolutional neural networks (CNN), for fully automatic soil-root image segmentation.

The RootNav 2.0 [55] and SegRoot [56] models, while advanced in their own right, face specific limitations in certain experimental setups. The RootNav 2.0 tool, for instance, was primarily developed for roots grown on germination paper with high contrast between root and (blue) background pixels. It is expected to perform poorly when applied to other imaging modalities, such as noisy roots grown in soil images. Similarly, the SegRoot model, while suitable for segmenting roots from soil background in minirhizotron systems, struggles to identify fine, fuzzy, and low-contrast roots. This limitation could affect the accuracy of the following phenotypic traits, including the estimated root biomass and other geometric features. These challenges highlight the need for more versatile and accurate solutions.

¹<https://ag-ba.ipk-gatersleben.de/saria.html>

A CNN approach based on the U-net architecture [44] designed to enhance the quality of automatic segmentation of root structures of different sizes, shapes, and optical contrasts from the soil environment. This approach transfers the location and pixel information of the feature maps from the encoder to the decoder. In contrast, existing solutions like SegRoot only transfer the location of feature maps to the decoder for image segmentation. The method is called fully-automated root image analysis (faRIA), which fully automates the segmentation of roots and the computation of phenotypic traits. During the experimental setup, greyscale root images of Maize plants acquired with a Near-infrared (NIR) based IPK plant phenotyping system were selected and annotated by agronomists using the saRIA tool for the ground truth segmentation. The performance of the trained model is analyzed using binary cross-entropy loss, Dice coefficient (DC), precision and recall.

The results show that the proposed U-net framework (faRIA) for the segmentation of roots in soil significantly outperforms the SegRoot method on our data set by improving the binary cross-entropy loss by the factor 10 and DC by 20%, respectively. It can be concluded from this study that extensive feature information (i.e., both location and pixel values), including low and high-level abstract features in the faRIA model, enables the detection of both primary and secondary low-contrast roots compared to the only max-pooling indices in the SegRoot model. Furthermore, the segmentation of full images resized to 1024x1024 using faRIA exhibits false positive segmentation with low recall and high precision at the soil-root region boundary compared to the original full-size images. It results in false phenotypic values for the root system architecture analysis. Therefore, it is vital to calculate the accurate phenotypic traits by preserving root structures for high-throughput plant phenotyping.

In addition, the presence of bias in a model can hinder its generalization capabilities. Our results show that the performance of the proposed U-net architecture is more robust, with an accuracy of around 80% than SegRoot on segmentation of unseen Barley and Arabidopsis soil-root images captured under rhizotron and UV imaging systems, respectively. The study concludes that the faRIA model adequately adapted to new, previously unseen data drawn from the same distribution (soil-root images) as the one used to create the model. In conclusion, end users without technical and IT knowledge can automatically perform high-throughput root system phenotyping using the

faRIA tool ².

7.1.3 Objective 3

Investigation of shallow learning-based neural networks and different deep learning models for the segmentation of different plant shoot images (Arabidopsis, Barley and Maize) at different developmental stages (e.g., juvenile, adult) in different views (e.g., top and side views) and evaluate their performance with the proposed CNN model.

To fully characterize the relationship between a plant’s genotype and phenotype, the shoot system needs to be studied in addition to the root system. However, segmenting different parts of different plant types at different developmental stages in different views is complicated by factors such as inhomogeneous illumination, shadows, occlusions, reflections, and the dynamic optical appearance of growing plants. These limitations make it challenging to establish invariant criteria for detecting parts such as leaves, flowers, fruits, and spikes, especially for plant types such as Arabidopsis, Maize, and Barley, and at developmental stages ranging from juvenile to adult. Moreover, phenotyping of thousands and millions of plant images in a high-throughput manner requires segmentation algorithms that are fully automated, accurate, and efficient.

The state-of-the-art conventional solutions based on reference frames [29, 30] and pixel-based supervised learning methods [31, 32] are generally not sufficiently robust for unsupervised analysis of sizeable variable image data. Henke et al. [30] proposed a registration-classification approach based on the same plant’s fluorescent and visible light RGB images. However, the requirement of multimodal data limits the application of this method to uni-modal plant phenotyping systems. Furthermore, pixel-based supervised learning methods [31, 32] are sensitive to noise due to the missing neighbourhood information or patterns in the image. Also, they require expertise in manual parameter tuning for each image or feature engineering techniques for supervised methods such as shallow neural network models. As of the published works of the thesis, no appropriate deep learning solutions for accurate segmentation of different plant shoot (Arabidopsis, Barley and Maize) images are known.

²<https://ag-ba.ipk-gatersleben.de/faria.html>

In this work, a deep learning-based U-net model was developed to automate and efficiently segment shoot images (DeepShoot)³. This model can be applied to multiple plants in both top and side views, including Arabidopsis, Barley, and Maize, and does not require reference frames. In contrast to many current solutions that are limited to a single plant type or view, the proposed U-net model can generate multi-level features, such as neighbourhood information, colour, spatial patterns, and textural features. This is a significant advancement over the state-of-the-art solution proposed by Adam et al., which only considered neighbourhood information. The performance of the proposed U-net model was evaluated with different encoder backbones, including vgg19, resnet50, and xception. This approach has the potential to revolutionize automated and efficient plant phenotyping.

The shoot image analysis results summarize that the proposed U-net model significantly outperforms $DC > 0.9$ for all listed plants compared to the pixel-based supervised learning methods [32] have DC between 0.5 and 0.8. Moreover, the U-net model accurately segments the optically variable structures like brown colour stressed leaves and low contrast leaves in Arabidopsis top view and Barley side view images, respectively. It confirms that segmentation of plant shoot structures requires multi-level features, which are missing in most state-of-the-art solutions. On the other hand, deep learning models with complex architectures (resnet50 and xception encoder backbones) and many redundant features result in the model's underperformance compared to the lower-depth architectures of the proposed U-net and vgg19. Therefore, it is essential to maintain the complexity of features for the robust segmentation of optically varying shoot images. In addition, the generalizability of the proposed U-net model was also tested on an unseen publicly available Arabidopsis dataset, grown under soil background from Scharr et al. [57]. This study concludes that our model can segment quite different optical backgrounds or field-like images until target plant structures are similar to images used in the training sets. However, the proposed model's limitation was primarily seen in side-view shoot images, where shadowed areas with thin and twisted leaves are more likely to be misclassified.

7.1.4 Objective 4

Enhancement of wheat spike detection algorithm in 2D images using shallow learning-based neural networks and studying the robustness of the model on European cultivars.

³<https://ag-ba.ipk-gatersleben.de/ds.html>

In addition to the phenotyping of plant primary (root and shoot) systems, the need and demand for image-based Wheat spike phenotyping are increasing to contribute to studies on improving crop yield for food security and sustainability. Nevertheless, there have been only a small number of studies focused (until this publication) on Wheat spike detection and growth analysis [34, 35, 36, 37] from greenhouse digital images because of the need for imaging data and labelled ground truth. Qiongyan et al. [37] proposed a novel approach for detecting (young) spikes in digital images of Wheat plants based on Law’s textural (energy) features and a neural network. However, by application to European elite cultivars, this algorithm was sensitive to the high-energy leaf crossing or overlay artefacts and aged leaves, which led to false classifications of spike and non-spike pixels.

To overcome the limitations of the Qiongyan et al. method, a reliable three-step method for spike detection was proposed based on wavelet textural features [58], Frangi filter [59] and artificial neural networks. In our method, discrete wavelet transform enhances the separability between the plant and background pixels by resulting in a coefficient image bearing the textural information relevant to spike segmentation. Later, the Frangi filter was applied to suppress high-energy edges resulting from such leaf crossings in the neural network segmented images.

The study’s findings demonstrate that the proposed approach performs better than the current solution Qiongyan et al. [37], explicitly created for Wheat cultivars grown in Australia. Notably, there was a 75% improvement in the precision of spike heading time point detection for European Wheat cultivars with bushy leaves surrounding the spike. It is also evident in the time series growth analysis of the predicted spike area. These results conclude that (i) textural features extracted from the discrete wavelet transform image are much more robust than the original image, and (ii) introducing the Frangi filter suppresses the bushy leaves surrounded by spikes in European cultivars. Moreover, the genotype-phenotype study also indicates that the proposed model is well-suited for plants where spikes emerge quickly with high energy patterns (low biomass plants). However, the model failed to detect very smooth and green colour spikes, which take a long heading time to emerge in plants. In summary, automatic spike phenotyping of European cultivars is a good fit for the suggested spike detection method in 2D images. Furthermore, ground truth data creation for deep learning applications can benefit from its use.

7.1.5 Objective 5

Developing a CNN model for the segmentation of barbs in microscopic images of Barley cultivars and studying the robustness of the proposed model by phenotypic characterization and classification of awns.

In addition to the spikes, awns play a significant role in the phenotypic description of grain crops like Barley and Wheat. *Awn's roughness* is a property determined by the density and presence of structures known as barbs. Until now, many studies have been conducted on the genetic control of the intensity of barb formation in Barley using gene analysis [60]. However, the potential of linking awn genotype with phenotypic traits still needs to be included. To our knowledge, no appropriate image processing solutions are suitable for accurate barb segmentation and phenotyping in awn images.

The identification of barbs on awn surfaces poses a challenge due to their inconsistent appearance, low contrast, and limited spatial resolution. To address this issue, we utilized cutting-edge deep learning techniques by initially applying the original U-net model [44] for automated analysis of awns in a high-throughput manner. However, we encountered unsatisfactory results because the original U-net model needed to be specifically tailored for awn image analysis. As a solution, we modified the U-net architecture called BarbNet⁴. to enable full automation of barb segmentation in microscopic images. The major modifications include, unlike the original U-net model, BarbNet includes batch normalization after each convolution layer to enhance network performance and stability by normalizing the feature maps at respective levels. Dropout layers are not used in BarbNet because combining batch normalization and dropout layers can lead to poor results [61].

The experimental results show that the proposed BarbNet model without dropout layer and redundant features improved DC by 6% compared to the originally developed U-net model for biomedical image analysis. This results in higher DC = 93% for BarbNet on the test dataset compared to the original and three modified U-net models (DC <85%). It briefs that reducing the number of filters with increased depth of the model and combining batch normalization and dropout layers penalizes detecting highly variable tiny barbs. Furthermore, U-net models with more features make diagnosing and addressing the barb segmentation problem harder. Also, this limita-

Chapter 6

⁴<https://doi.ipk-gatersleben.de/DOI/ce489d84-ed66-4171-ad97-cf4696099389/09144b48-8336-48d7-86a2-f00dfd988cba/2>

tion was observed in BarbNet for segmenting very tiny barbs in the basal region of awns. In addition to the segmentation performance, three important phenotypic traits for awn and barb characterization: total barb count, mean barb area, and mean barb length in pixels are presented. Our experiments with different genotypes explain that a combination of barb density and size provides the best results on central regions of awns (i.e., >85%) than basal regions (81%). However, one can achieve the best and most balanced gene classification results by combining multiple phenotypic traits of barbs. Therefore, the proposed model may be used to automate the sorting of Barley awns to study plant development.

7.1.6 Objective 6

Developing a GUI-based front end for efficiently handling the above algorithmic frameworks.

Developed plant phenotyping algorithms are designed to solve specific problems and require programming skills to implement. However, many plant biologists need to gain these skills, which can make these solutions difficult to use. To address this, GUI-based solutions with an intuitive interface, visualization capabilities, and experimental design and analysis tools have been developed to enable biologists to tackle complex biological problems more effectively.

User-friendly GUI-based Windows and Linux precompiled Matlab executable programs named saRIA, faRIA, and DeepShoot were developed for root-and-shoot image analysis. These tools allow end-users with limited technical knowledge to perform automated analysis of plant images intuitively and transparently. The tools can be used to analyze single images or large image datasets to detect and extract multiple plant traits, with the ability to define a custom region of interest and eliminate artefacts caused by external factors. The saRIA/faRIA and DeepShoot tools calculate 70 root and 35 shoot traits, respectively. A command line-based tool for both Windows and Linux is available for awn phenotyping. More information on the definition of traits can be found in the Supplementary information accompanying the corresponding articles.

7.2 Conclusion & Future Work

As outlined, image-based high-throughput plant phenotyping systems require more advanced and fully automated computer vision solutions to deal with complex and diverse plant structures. However, the integration of sophisticated computer vision solutions in plant science has been hindered by limited data accessibility and annotation. Initially, a semi-automated image analysis approach for root image segmentation (saRIA) and spike detection in Wheat plants was developed to overcome these limitations. Eventually, deep learning-based solutions for a broad spectrum of applications, including root (faRIA), shoot (DeepShoot) image analysis (BarbNet) and phenotyping, were developed in this thesis.

For the first, results of semi-automated approaches for root and spike image segmentation show that both models achieved more than 80% accuracy on automatic image segmentation and phenotyping traits calculation. However, the accuracy of these models can be improved by tuning the algorithm's parameters according to target patterns in the image. Moreover, these tools provide solutions to agronomists for analysing small datasets, variational imaging setups like roots grown in any medium, and the rapid generation of ground truth data. However, they limit the analysis of large amounts of images for high-throughput plant phenotyping.

Later, with the help of semi-automated approaches, fully automated CNN models for image segmentation using low-budget hardware systems were developed. The results show that CNN models achieved approximately 90% accuracy on root, shoot and barb segmentation in awn images. Furthermore, the study on model complexity concludes that the proposed CNN models outperform the deep-depth CNN architectures, which complicates the interpretability of the model. This results in more generalizable CNN models for segmenting target plant structures are similar to images used in the training sets. The major limitation of this work is the generalizability of the developed methods to broad fields like biomedical image analysis. All methods were exclusively developed in this work based on real-time data from plant phenotyping experiments. In conclusion, end-users with limited technical knowledge can perform analysis of large amounts of images for high-throughput plant phenotyping.

The computer vision methods that are presented in this thesis work brought improvements in automated plant image phenotyping for plant science; there is a vast scope for improving and developing advanced deep learn-

ing algorithms if ground truth data is available. Since all developed methods in this thesis are based on binary segmentation, they can be extended to multi-class segmentation, for example, the segmentation of primary and secondary roots and healthy and stressed/diseased leaves. Also, deep learning algorithms can address the artificial filling of gaps in the root system that occurred due to inhomogeneous scene illumination. Regarding shoot image analysis, multiple plants are often screened together in a single image, and the separation of these plants can be addressed in the future using graph theory and/or a deep learning approach.

Bibliography

- [1] Ines Janusch, Walter Kropatsch, and Wolfgang Busch. Topological image analysis and (normalised) representations for plant phenotyping. *Proceedings - 16th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, SYNASC 2014*, pages 579–586, 02 2015.
- [2] Roland Pieruschka and Uli Schurr. Plant phenotyping: past, present, and future. *Plant Phenomics*, 2019, 2019.
- [3] Mark Tester and Peter Langridge. Breeding technologies to increase crop production in a changing world. *Science*, 327(5967):818–822, 2010.
- [4] JW Snape, K Butterworth, E Whitechurch, and AJ Worland. Waiting for fine times: genetics of flowering time in wheat. In *Wheat in a global environment*, pages 67–74. Springer, 2001.
- [5] Kunpu Zhang, Jichun Tian, Liang Zhao, Bin Liu, and Guangfeng Chen. Detection of quantitative trait loci for heading date based on the doubled haploid progeny of two elite chinese wheat cultivars. *Genetica*, 135(3):257–265, 2009.
- [6] Lei Li, Qin Zhang, and Danfeng Huang. A review of imaging techniques for plant phenotyping. *Sensors*, 14(11):20078–20111, 2014.
- [7] Lance R Veldboom and Michael Lee. Genetic mapping of quantitative trait loci in maize in stress and nonstress environments: I. grain yield and yield components. *Crop Science*, 36(5):1310–1319, 1996.
- [8] Marion I Menzel, Susanne Tittmann, Jonas Buehler, Stella Preis, Norbert Wolters, Siegfried Jahnke, Achim Walter, Antonia Chlubek, Ariel Leon, Normen Hermes, et al. Non-invasive determination of plant biomass with microwave resonators. *Plant, cell & environment*, 32(4):368–379, 2009.

- [9] L Waltman. Van eck, nj (2013), “a smart local moving algorithm for large-scale modularity-based community detection”. *The European Physical Journal*, 86(11):471, 2013.
- [10] Randy T Clark, Robert B MacCurdy, Janelle K Jung, Jon E Shaff, Susan R McCouch, Daniel J Aneshansley, and Leon V Kochian. Three-dimensional root phenotyping with a novel imaging and software platform. *Plant physiology*, 156(2):455–465, 2011.
- [11] Marcus Jansen, Frank Gilmer, Bernhard Biskup, Kerstin A Nagel, Uwe Rascher, Andreas Fischbach, Sabine Briem, Georg Dreissen, Susanne Tittmann, Silvia Braun, et al. Simultaneous phenotyping of leaf growth and chlorophyll fluorescence via growscreen fluoro allows detection of stress tolerance in arabidopsis thaliana and other rosette plants. *Functional Plant Biology*, 36(11):902–914, 2009.
- [12] Corrado Costa, Francesca Antonucci, Federico Pallottino, Jacopo Aguzzi, Da-Wen Sun, and Paolo Menesatti. Shape analysis of agricultural products: a review of recent research advances and potential application to computer vision. *Food and Bioprocess Technology*, 4(5):673–692, 2011.
- [13] Hendrik Poorter, Fabio Fiorani, Mark Stitt, Uli Schurr, Alex Finck, Yves Gibon, Björn Usadel, Rana Munns, Owen K Atkin, François Tardieu, et al. The art of growing plants for experimental purposes: a practical guide for the plant biologist. *Functional Plant Biology*, 39(11):821–838, 2012.
- [14] Anjali S Iyer-Pascuzzi, Olga Symonova, Yuriy Mileyko, Yueling Hao, Heather Belcher, John Harer, Joshua S Weitz, and Philip N Benfey. Imaging and analysis platform for automatic phenotyping and trait ranking of plant root systems. *Plant physiology*, 152(3):1148–1157, 2010.
- [15] Noah Fahlgren, Malia A Gehan, and Ivan Baxter. Lights, camera, action: high-throughput plant phenotyping is ready for a close-up. *Current opinion in plant biology*, 24:93–99, 2015.
- [16] David Deery, Jose Jimenez-Berni, Hamlyn Jones, Xavier Sirault, and Robert Furbank. Proximal remote sensing buggies and potential applications for field-based phenotyping. *Agronomy*, 4(3):349–379, 2014.
- [17] Neil R Baker. Chlorophyll fluorescence: a probe of photosynthesis in vivo. *Annu. Rev. Plant Biol.*, 59:89–113, 2008.

- [18] Qin Zhang. *Precision agriculture technology for crop farming*. Taylor & Francis, 2016.
- [19] Gregory P Asner, David E Knapp, Christopher B Anderson, Roberta E Martin, and Nicholas Vaughn. Large-scale climatic and geophysical controls on the leaf economics spectrum. *Proceedings of the National Academy of Sciences*, 113(28):E4043–E4051, 2016.
- [20] Halimatu Sadiyah Abdullahi, F Mahieddine, and Ray E Sheriff. Technology impact on agricultural productivity: A review of precision agriculture using unmanned aerial vehicles. In *Wireless and Satellite Systems: 7th International Conference, WiSATS 2015, Bradford, UK, July 6-7, 2015. Revised Selected Papers 7*, pages 388–400. Springer, 2015.
- [21] Massimo Minervini, Hanno Scharr, and Sotirios A Tsaftaris. Image analysis: the new bottleneck in plant phenotyping [applications corner]. *IEEE signal processing magazine*, 32(4):126–131, 2015.
- [22] J-L Arsenault, S Poulcur, C Messier, and R Guay. WinRHIZO: a root-measuring system with a unique overlap correction method. *HortScience*, 30(4):906D–906, 1995.
- [23] Patrick Armengaud, Kevin Zambaux, Adrian Hills, Ronan Sulpice, Richard J Pattison, Michael R Blatt, and Anna Amtmann. Ez-rhizo: integrated software for the fast and accurate measurement of root system architecture. *The Plant Journal*, 57(5):945–956, 2009.
- [24] Guillaume Lobet, Loïc Pagès, and Xavier Draye. A novel image-analysis toolbox enabling quantitative analysis of root system architecture. *Plant physiology*, 157(1):29–39, 2011.
- [25] Alain Pierret, Santimaitree Gonkhamdee, Christophe Jourdan, and Jean-Luc Maeght. Ij_rhizo: an open-source software to measure scanned images of root samples. *Plant and soil*, 373(1):531–539, 2013.
- [26] Kerstin A Nagel, Alexander Putz, Frank Gilmer, Kathrin Heinz, Andreas Fischbach, Johannes Pfeifer, Marc Faget, Stephan Blossfeld, Michaela Ernst, Chryssa Dimaki, et al. Growscreen-rhizo is a novel phenotyping robot enabling simultaneous measurements of root and shoot growth for plants grown in soil-filled rhizotrons. *Functional Plant Biology*, 39(11):891–904, 2012.

- [27] Boris Rewald and Jhonathan E Ephrath. Minirhizotron techniques. In *Plant Roots: The Hidden Half, Fourth Edition*, pages 735–750. CRC press, 2013.
- [28] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, 1979.
- [29] Sruti Das Choudhury, Srinidhi Bashyam, Yumou Qiu, Ashok Samal, and Tala Awada. Holistic and component plant phenotyping using temporal image sequence. *Plant methods*, 14(1):1–21, 2018.
- [30] Michael Henke, Astrid Junker, Kerstin Neumann, Thomas Altmann, and Evgeny Gladilin. A two-step registration-classification approach to automated segmentation of multimodal images for high-throughput greenhouse plant phenotyping. *Plant Methods*, 16(1):1–10, 2020.
- [31] Unseok Lee, Sungyul Chang, Gian Anantrio Putra, Hyoungeok Kim, and Dong Hwan Kim. An automated, high-throughput plant phenotyping system using machine learning-based plant segmentation and image analysis. *PloS one*, 13(4):e0196615, 2018.
- [32] Jason Adams, Yumou Qiu, Yuhang Xu, and James C Schnable. Plant segmentation by supervised machine learning methods. *The Plant Phenome Journal*, 3(1):e20001, 2020.
- [33] Mark Tester and Peter Langridge. Breeding technologies to increase crop production in a changing world. *Science*, 327(5967):818–822, 2010.
- [34] Bi Kun, Jiang Pan, Tang Chongwei, Huang Feifei, and Wang Cheng. The design of wheat variety bp classifier based on wheat ear feature. *Chinese Agricultural Science Bulletin*, 27(6):464–468, 2011.
- [35] Kun Bi, Pan Jiang, Lei Li, Benyi Shi, and Cheng Wang. Non-destructive measurement of wheat spike characteristics based on morphological image processing. *Transactions of the Chinese Society of Agricultural Engineering*, 26(12):212–216, 2010.
- [36] Kun Bi, Fei-fei Huang, and Cheng Wang. Quick acquisition of wheat ear morphology parameter based on imaging processing. In *International Workshop on Computer Science for Environmental Engineering and EcoInformatics*, pages 300–307. Springer, 2011.

- [37] Li Qiongyan, Jinhai Cai, Bettina Berger, Mamoru Okamoto, and Stanley J Miklavcic. Detecting spikes of wheat plants using neural networks with laws texture energy. *Plant Methods*, 13(1):1–13, 2017.
- [38] Rivka Elbaum, Liron Zaltzman, Ingo Burgert, and Peter Fratzl. The role of wheat awns in the seed dispersal unit. *Science*, 316(5826):884–886, 2007.
- [39] Andre Esteva, Katherine Chou, Serena Yeung, Nikhil Naik, Ali Madani, Ali Mottaghi, Yun Liu, Eric Topol, Jeff Dean, and Richard Socher. Deep learning-enabled medical computer vision. *NPJ digital medicine*, 4(1):5, 2021.
- [40] Alaa S Al-Waisy, Rami Qahwaji, Stanley Ipson, Shumoos Al-Fahdawi, and Tarek AM Nagem. A multi-biometric iris recognition system based on a deep learning approach. *Pattern Analysis and Applications*, 21:783–802, 2018.
- [41] Shaveta Dargan and Munish Kumar. A comprehensive survey on the biometric recognition systems based on physiological and behavioral modalities. *Expert Systems with Applications*, 143:113114, 2020.
- [42] Sakib Mostafa, Debajyoti Mondal, Karim Panjvani, Leon Kochian, and Ian Stavness. Explainable deep learning in plant phenotyping. *Frontiers in Artificial Intelligence*, 6, 2023.
- [43] Mark Everingham, Luc Van Gool, Christopher KI Williams, John Winn, and Andrew Zisserman. The pascal visual object classes (voc) challenge. *International journal of computer vision*, 88:303–338, 2010.
- [44] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.
- [45] Liang Zheng, Yi Yang, and Qi Tian. Sift meets cnn: A decade survey of instance retrieval. *IEEE transactions on pattern analysis and machine intelligence*, 40(5):1224–1244, 2017.
- [46] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 580–587, 2014.

- [47] Michael P Pound, Jonathan A Atkinson, Alexandra J Townsend, Michael H Wilson, Marcus Griffiths, Aaron S Jackson, Adrian Bulat, Georgios Tzimiropoulos, Darren M Wells, Erik H Murchie, et al. Deep machine learning provides state-of-the-art performance in image-based plant phenotyping. *Gigascience*, 6(10):gix083, 2017.
- [48] Clément Douarre, Richard Schielein, Carole Frindel, Stefan Gerth, and David Rousseau. Transfer learning from synthetic data applied to soil-root segmentation in x-ray tomography images. *Journal of Imaging*, 4(5):65, 2018.
- [49] Tanuj Misra, Alka Arora, Sudeep Marwaha, Viswanathan Chinnusamy, Atmakuri Ramakrishna Rao, Rajni Jain, Rabi Narayan Sahoo, Mrinmoy Ray, Sudhir Kumar, Dhandapani Raju, et al. Spikesegnet-a deep learning approach utilizing encoder-decoder network with hourglass for spike segmentation and counting in wheat plant from visual imaging. *Plant methods*, 16(1):1–20, 2020.
- [50] Yu Jiang and Changying Li. Convolutional neural networks for image-based high-throughput plant phenotyping: a review. *Plant Phenomics*, 2020, 2020.
- [51] Yezi Zhu, Marc Aoun, Marcel Krijn, Joaquin Vanschoren, and High Tech Campus. Data augmentation using conditional generative adversarial networks for leaf counting in arabidopsis plants. In *BMVC*, page 324, 2018.
- [52] Junjie Chen and Xinghua Shi. A sparse convolutional predictor with denoising autoencoders for phenotype prediction. In *Proceedings of the 10th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics*, pages 217–222, 2019.
- [53] Derek Bradley and Gerhard Roth. Adaptive thresholding using the integral image. *Journal of graphics tools*, 12(2):13–21, 2007.
- [54] Taras Galkovskyi, Yuriy Mileyko, Alexander Bucksch, Brad Moore, Olga Symonova, Charles A Price, Christopher N Topp, Anjali S Iyer-Pascuzzi, Paul R Zurek, Suqin Fang, et al. Gia roots: software for the high throughput analysis of plant root system architecture. *BMC plant biology*, 12(1):1–12, 2012.
- [55] Robail Yasrab, Jonathan A Atkinson, Darren M Wells, Andrew P French, Tony P Pridmore, and Michael P Pound. Rootnav 2.0: Deep

- learning for automatic navigation of complex plant root architectures. *GigaScience*, 8(11):giz123, 2019.
- [56] Tao Wang, Mina Rostamza, Zhihang Song, Liangju Wang, G McNickle, Anjali S Iyer-Pascuzzi, Zhengjun Qiu, and Jian Jin. Segroot: a high throughput segmentation method for root image analysis. *Computers and Electronics in Agriculture*, 162:845–854, 2019.
- [57] H Scharr, M Minervini, A Fischbach, and SA Tsaftaris. Annotated image datasets of rosette plants. Technical report, Forschungszentrum Jülich GmbH, 2014.
- [58] Yanqin Fan. On the approximate decorrelation property of the discrete wavelet transform for fractionally differenced processes. *IEEE Transactions on Information Theory*, 49(2):516–521, 2003.
- [59] Alejandro F Frangi, Wiro J Niessen, Koen L Vincken, and Max A Viergever. Multiscale vessel enhancement filtering. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI’98: First International Conference Cambridge, MA, USA, October 11–13, 1998 Proceedings 1*, pages 130–137. Springer, 1998.
- [60] Sara G Milner, Matthias Jost, Shin Taketa, Elena Rey Mazón, Axel Himmelbach, Markus Oppermann, Stephan Weise, Helmut Knüpffer, Martín Basterrechea, Patrick König, et al. Genebank genomics highlights the diversity of a global barley collection. *Nature genetics*, 51(2):319–326, 2019.
- [61] Xiang Li, Shuo Chen, Xiaolin Hu, and Jian Yang. Understanding the disharmony between dropout and batch normalization by variance shift, 2018.

Curriculum Vitae



Personal information

Name: Narendra Narisetti

Nationality: Indian

Email: narisetti@ipk-gatersleben.de, nari.narisetti@gmail.com

Education

- 2014-2017 **M.S. in Computers and Communication Technology**,
Thesis: “Development and Evaluation of a Menu Detection Algorithm Based on a Supervised Learning Method”,
Saarland University, Saarbruecken, Germany.
Grade: 2.1 (Best-1.0, Worst-5.0).
- 2009-2013 **B.S. in Electronics and Communication Engineering**,
Thesis: “Arduino Based Industrial Automization Using Real Time Clock”,
Jawaharlal Nehru Technological University, Hyderabad, India.
Grade: 8.1 (Best-10.0, Worst-1.0).

Professional Experience

- Since 07/2018 **Scientific Researcher**, Image Analysis group, IPK Gatersleben, Germany.
- Development of algorithms and software tools for high-throughput plant phenotyping.
 - Image analysis with the focus on segmentation, classification using machine/deep learning.
 - Derivation of quantitative traits of plant morphology, development and health.
 - Design and implementation of GUI based software solutions.
- 12/2017 – **Software Engineer**, AkknaTek GmbH, Saarbruecken, Germany.
- 06/2018
- Design and implementation of image analysis algorithms for the medical data.
 - Software verification like code reviews, unit test and error handling.
 - Technical documentation of the software on Microsoft Visio.
- 07/2017 – **Student Assistant**, Max Plank Institute for Informatics, Saarbruecken, Germany.
- 10/2017
- Image annotation and analysis for pose estimation of objects in real-time videos.

Internships

- 09/2012 – South Central Railways (Indian Government Company),
11/2012 Topic: Arduino Based Industrial Automization Using GSM (Embedded Systems)
- 01/06/2012 – Bharat Sanchar Nigam Limited (Indian Government Company),
30/06/2012 Topic: Functioning of WiMAX

Technical skills

Programming Languages	C, C++, MATLAB, VHDL, Python, HTML
Software Framework	TesnorFlow, PyTorch, Keras, PIL, scikit-learn, OpenCV, MS Office, Git
IDE	Visual Studio, Eclipse, Arduino, Xilinx, PyCharm
Operating Systems	Windows, Linux

Language skills

Telugu	Native	German	B1
English	C1	Hindi	B2

Hobbies

- Playing Badminton and Cricket, Jogging, Hiking and Cycling
- Reading technical articles

Publications

1. Awn image analysis and phenotyping using BarbNet. Plant Phenomics 5 (2023) 0081. <https://dx.doi.org/10.34133/plantphenomics.0081>
2. Line12D-Net: A deep learning approach to solving 2D linear elastic boundary value problems on image domains. iScience 27 (2024) 109519. <https://dx.doi.org/10.1016/j.isci.2024.109519>
3. FDM data driven U-Net as a 2D Laplace PINN solver. Sci. Rep. 13 (2023) 9116. <https://dx.doi.org/10.1038/s41598-023-35531-8>
4. Deep learning-based greenhouse image segmentation and shoot phenotyping (DeepShoot). Front. Plant Sci. 13 (2022) 906410. <https://dx.doi.org/10.3389/fpls.2022.906410>
5. Fully-automated root image analysis (faRIA). Sci. Rep. 11 (2021) 16047. <https://dx.doi.org/10.1038/s41598-021-95480-y>
6. Automated spike detection in diverse European wheat plants using textural features and the Frangi filter in 2D greenhouse images. Front. Plant Sci. 11 (2020) 666. <https://dx.doi.org/10.3389/fpls.2020.00666>
7. Semi-automated Root Image Analysis (saRIA). Sci. Rep. 9 (2019) 19674. <https://dx.doi.org/10.1038/s41598-019-55876-3>

Ort, Datum / Place, Date
Wernigerode, 03.03.2025

Narendra