Studies in Big Data 121

Sanjay Chaudhary
Chandrashekhar M. Biradar
Srikrishnan Divakaran
Mehul S. Raval *Editors*

Digital Ecosystem for Innovation in Agriculture

Springer

Studies in Big Data

Volume 121

Series Editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland

The series "Studies in Big Data" (SBD) publishes new developments and advances in the various areas of Big Data- quickly and with a high quality. The intent is to cover the theory, research, development, and applications of Big Data, as embedded in the fields of engineering, computer science, physics, economics and life sciences. The books of the series refer to the analysis and understanding of large, complex, and/or distributed data sets generated from recent digital sources coming from sensors or other physical instruments as well as simulations, crowd sourcing, social networks or other internet transactions, such as emails or video click streams and other. The series contains monographs, lecture notes and edited volumes in Big Data spanning the areas of computational intelligence including neural networks, evolutionary computation, soft computing, fuzzy systems, as well as artificial intelligence, data mining, modern statistics and Operations research, as well as self-organizing systems. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution, which enable both wide and rapid dissemination of research output.

The books of this series are reviewed in a single blind peer review process.

Indexed by SCOPUS, EI Compendex, SCIMAGO and zbMATH.

All books published in the series are submitted for consideration in Web of Science.

Sanjay Chaudhary · Chandrashekhar M. Biradar ·
Srikrishnan Divakaran · Mehul S. Raval
Editors

Digital Ecosystem for Innovation in Agriculture

Springer

Editors
Sanjay Chaudhary
School of Engineering and Applied Science
Ahmedabad University
Ahmedabad, Gujarat, India

Chandrashekhar M. Biradar
CIFOR-ICRAF India
Asia Continental Program
New Delhi, Delhi, India

Srikrishnan Divakaran
School of Engineering and Applied Science
Ahmedabad University
Ahmedabad, Gujarat, India

Mehul S. Raval
School of Engineering and Applied Science
Ahmedabad University
Ahmedabad, Gujarat, India

ISSN 2197-6503 ISSN 2197-6511 (electronic)
Studies in Big Data
ISBN 978-981-99-0576-8 ISBN 978-981-99-0577-5 (eBook)
https://doi.org/10.1007/978-981-99-0577-5

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

To my wife (Sunita), son (Mandar), and daughter (Anuradha)

by Sanjay Chaudhary

Dr. Ravi Prabhu, Director General, The World Agroforestry, Dr. Javed Rizvi, Dr. Maarten van Ginkel, and Dr. S. K. Dhyani

by Chandrashekhar Biradar

To my wife Vidya, son Adhithya, mother Vidya, father Divakaran, and my colleagues for their support and encouragement

by Srikrishnan Divakaran

In dedication to my wife—Hemal, son—Hetav, in memory of my mother—Charu, and father—Shirish, for their constant encouragement and support through thick and thin

by Mehul Raval

Foreword

Meeting the food demands in the 21st century seems to be a continued story of catch-up. It has become increasingly clear that the breathing time Norman Borloug and colleagues afforded us with the Green Revolution has not been effectively used to harness food demand. Although over that past half century significant progress has been achieved in reducing hunger and poverty as well as improving food security and nutrition, in the process we have seriously degraded the natural resource base (soils, water, and biodiversity), on which our food production is based. Climate change is adding an extra threat to food security. Calls are getting louder to shift to sustainable and regenerative food production systems that will feed an ever-increasing population. The agricultural science and development communities are looking to expand the technologies and toolbox that will help create efficient agricultural systems that are economically, socially, and environmentally sustainable.

One of the modern technologies that might help address the complex problems facing agriculture may be found in the innovative application of data science, big data analytics, remote sensing, Internet of Things, computer vision, machine learning, cloud computing, artificial intelligence, etc. Technologies are in place to capture big data in real-time manner, and modern farmers in advanced economies are increasingly providing and using such data for farm and resource management. In the upcoming decade, an increasing number of farms will partake in this farm modernization process. Policymakers as well as farmers will benefit from it to make better decisions. The challenge will lie in involving the resource-poor farmers in this transformative process.

This book intends to provide the reader with an overview of the frameworks and technologies involved in the digitalization of agriculture, as well as the data processing methods, decision-making processes, and innovative services/applications for enabling digital transformations in agriculture. This book has two broad sections: (1) *Frameworks, Tools, and Technologies for Transforming Agriculture* and (2) *Problems and Applications of Digital Agricultural Transformations.*

The first part offers an overview of the challenges and opportunities in transforming agriculture through efficient and cost-effective digital services and applications. The chapters in this part discuss the key aspects of building a framework for allocating digital resources necessary for developing digital services and applications in agriculture. It also discusses some of the key principles and concepts as well as technologies and tools in AI and machine learning useful for the creation of resource-efficient services and applications on these platforms.

The second part presents key principles and concepts in computer vision, machine learning, remote sensing, and artificial intelligence (AI). They demonstrate their use in developing intelligent services and applications to solve agricultural problems that arise in the context of plant phonemics, sustainable agriculture, yield prediction, and farm data integration. It also presents some analytical tools for analyzing policies for allocating farm resources and measuring farm productivity.

The value and relevance of any book can be measured by the range and depth of topics and the quality of its chapters. This book meets those criteria. The editors of this book were successful in assembling a valuable collection of chapters on technological advances to help address a key and essential challenges in agriculture. The chapters are well written by competent authors associated with globally reputed organizations. This book provides qualitative and relevant reference content on digital infrastructure for innovation in sustainable agriculture. This book offers an excellent source of knowledge and information that will help a range of professionals, from policymakers to scientists to technology developers and end-users. I congratulate the editors and authors of this book for their commendable contribution.

Austin, TX, USA
December 2022

Paul Vlek
Professor Emeritus
Former Director
Center for Development Research (ZEF)
Division of Ecology and Natural Resources
University of Bonn
Bonn, Germany

Former Executive Director of WASCAL
West African Science Service Center on Climate
Change and Adapted Land Use
Accra, Ghana

Preface

This book on *Digital Ecosystem for Innovation in Agriculture* is an attempt to provide the reader with an accessible big picture of technologies and innovations for the digital transition of agriculture, as well as the key challenges and research trends in the development of data analytical frameworks, tools, and their applications in the context of Digital Augmentation in Agriculture. "Digital Agriculture," here we refer to leveraging the digital technological advances in agriculture and agroecosystems to improve the functional productivity of processes involved in agri-food systems. Furthermore, the transformation should be economically viable and ecologically sustainable.

The increased availability of sophisticated remote sensing satellite services, widespread use of unmanned aerial vehicles (UAVs), better access to quality of data, power of cyber infrastructure, and the easy deployment of inexpensive Internet of Things (IoT) sensors, standardized interfaces, operations, and programmable frameworks have propelled applications in agricultural ecosystems. It is coupled with significant advances in data science (tools and techniques for filtering, compressing, processing, and analyzing) and big data analytics (capturing, storing, retrieving, and visualizing big data). In addition, machine learning (tools and algorithms for building models, making predictions, and performing statistical analysis on data) has accelerated the development of various niche services in Digital Agriculture.

The main focus of this book is on (i) Frameworks and Systems: Handling Big Data, employing remote sensing technology, making provisions for providing and accessing computing, storage, and services over the cloud, and the Internet of Things for collecting, filtering, storing, retrieving, integrating, and visualizing farm data; (ii) Tools and Techniques in Data Science and Machine Learning for developing models, algorithms, and services for providing niche agricultural services that involve data analytics, making predictions and providing statistical guarantees on their predictions.

The book chapters are divided into two parts. The book's first part, *Frameworks, Tools, and Technologies for Transforming Agriculture*, focuses on the critical issues in developing platforms/frameworks for effectively allocating digital infrastructure for building digital services/applications in agriculture. One of the vital issues is climate change which is now touching our daily life. It is also threatening agriculture and needs mitigation. The first chapter "A Brief Review of Tools to Promote Transdisciplinary Collaboration for Addressing Climate Change Challenges in Agriculture by Model Coupling" suggests increasing agriculture efficiencies and making room for renewable bioenergy crops. The chapter summarizes the tools that promote collaboration for developing sustainable and climate-resilient agriculture and discusses model coupling about plant and agri-sciences. The second chapter "Machine Learning and Deep Learning in Crop Management—A Review" provides a survey of the tools and techniques of machine and deep learning employed in agriculture. It discusses algorithms for crop management activities like crop yield prediction, diseases, and pest and weed detection. Satellites, drones, and on-ground sensors are essential in providing data for the digital ecosystem for agriculture innovation. However, all three modes of data collection are executed in isolation. Therefore, the need for an orchestration platform to exploit the potential of remote sensing data is presented in the third chapter "Need for an Orchestration Platform to Unlock the Potential of Remote Sensing Data for Agriculture".

However, it is crucial to develop strategies to connect multimodal data. In that context, the fourth chapter "An Algorithmic Framework for Fusing Images from Satellites, Unmanned Aerial Vehicles (UAV), and Farm Internet of Things (IoT) Sensors" shows an algorithmic framework for constructing higher-dimensional maps through data fusion of satellite images and unmanned aerial vehicles with multisensor farm data. Remote sensing plays a critical role in mapping and monitoring crops on a large scale. The availability of open-source data and cloud resources play a significant role in developing remote sensing-based solution. Therefore, the fifth chapter "Globally Scalable and Locally Adaptable Solutions for Agriculture" focuses on using open-source high-resolution spectral, spatial, and temporal resolution satellite data, open-source cloud-based platforms, and big data algorithms for agriculture. Continuous knowledge management (KM) can trigger innovation in agriculture. Therefore, developing a theoretical framework to guide the KM process is crucial. The sixth chapter "A Theoretical Framework of Agricultural Knowledge Management Process in the Indian Agriculture Context", the last chapter of the part, uses the famous Indian milk cooperative sector as an example, derives various systemic factors, and guides agri-organization through KM processes.

The book's second part, *Problems and Applications of Digital Agricultural Transformations*, presents specific challenges for Digital Agriculture that employs computer vision, machine learning, and remote sensing tools and techniques. This part spans from chapter "Simple and Innovative Methods to Estimate Gross Primary Production and Transpiration of Crops: A Review" to "Computer Vision Approaches for Plant Phenotypic Parameter Determination". The most significant carbon and water fluxes in agroecosystems are gross primary production (GPP) and transpiration (TR) of crops. Crop yield estimate using GPP and transpiration measurement can improve irrigation in cropland. The seventh chapter "Simple and Innovative Methods to Estimate Gross Primary Production and Transpiration of Crops: A Review" reviews simple and innovative methods to estimate gross primary production and transpiration. It reviews the state of the science, including in situ and remote sensing methods, while focusing on the biophysical foundation. The growth of computation power has facilitated the creation of 3D models of plants or trees. Virtual plants can simulate crop growth (in silico) compared to the natural environment. The eighth chapter "Role of Virtual Plants in Digital Agriculture" overviews the role of virtual plants in Digital Agriculture, showcasing that in silico implementation is an alternative to time-consuming, labor-intensive actual field trials. Finally, this chapter covers the concept of VP modeling, its applications, and some challenges in its application. Anthropogenic activities can impact soil carbon pools and phytomass on a vast scale. Orchards and plantations can affect the carbon pool, which needs to be carefully studied. The ninth chapter "Remote Sensing for Mango and Rubber Mapping and Characterization for Carbon Stock Estimation—Case Study of Malihabad Tehsil (UP) and West Tripura District, India" provides a case study involving remote sensing for mango and rubber mapping and characterization for carbon stock estimation. It uses Sentinel-2 data and machine learning to classify tree species. It demonstrates that simultaneous high-resolution phytomass and soil mapping with geospatial techniques significantly enhances India's capability to monitor and model terrestrial carbon pools.

Deep learning (DL) and computer vision (CV) advances are penetrating agriculture and natural resource management. The next set of chapters showcases such use cases of DL. The tenth chapter "Impact of Vegetation Indices on Wheat Yield Prediction Using Spatio-Temporal Modeling" presents the use of spatial-temporal modeling to study the impact of vegetation indices on wheat yield prediction. It showcases the use of convolutional neural networks (CNNs) and long short-term memory (LSTM) for yield prediction. Irrigation is a crucial phase of crop cultivation, and its scheduling and water management play essential roles in arid regions. Therefore, estimating the crop-specific water requirement at the farm and a more significant level is necessary. The eleventh chapter "Farm-Wise Estimation of Crop Water Requirement of Major Crops Using Deep Learning Architecture" illustrates the use of deep learning in evaluating farm-wise crop water requirements of major crops. It showcases platform development for adequately managing water resources across states in India.

Usually, remote sensing satellite data is available as multispectral images. More spectral information can increase the accuracy of the machine learning algorithms used in Digital Agriculture. Hyperspectral sensing (HyS) provides very high spectral resolution and is useful in land use and land cover classification (LULC). The twelfth chapter "Hyperspectral Remote Sensing for Agriculture Land Use and Land Cover Classification" presents LULC using hyperspectral sensing. A review of current algorithms for processing HyS datasets is carried out in this article. It includes validating various atmospheric correction (AC) models, dimensionality reduction techniques, and classification methods. In plant breeding, phenotypic trait measurement, i.e., morphological and physiological characteristics, is necessary to develop improved crop varieties. Computer vision-based techniques have emerged as an efficient method for non-invasive and non-destructive plant phenotyping. The thirteenth chapter "Computer Vision Approaches for Plant Phenotypic Parameter Determination" presents Computer Vision Approaches for Plant Phenotypic Parameter Determination. A deep learning-based encoder-decoder network is developed to recognize and count the number of spikes from visual images of wheat plants.

Ahmedabad, India — Sanjay Chaudhary
New Delhi, India — Chandrashekhar M. Biradar
Ahmedabad, India — Srikrishnan Divakaran
Ahmedabad, India — Mehul S. Raval
December 2022

Acknowledgements We are thankful to

- Contributing authors
- Springer
- Janusz Kacprzyk, Aninda Bose, and Nareshkumar Mani
- Ahmedabad University
- Center for International Forestry Research (CIFOR) and World Agroforestry (ICRAF)
- Family members

Contents

Computer Vision Approaches for Plant Phenotypic Parameter Determination

Alka Arora, Tanuj Misra, Mohit Kumar, Sudeep Marwaha, Sudhir Kumar, and Viswanathan Chinnusamy

Abstract Climate change and the growing population are major challenges in the global agriculture scenario. High-quality crop genotypes are essential to counter the challenges. In plant breeding, phenotypic trait measurement is necessary to develop improved crop varieties. Plant phenotyping refers to studying the plant's morphological and physiological characteristics. Plant phenotypic traits like the number of spikes/panicle in cereal crops and senescence quantification play an important role in assessing functional plant biology, growth analysis, and net primary production. However, conventional plant phenotyping is time-consuming, labor-intensive, and error-prone. Computer vision-based techniques have emerged as an efficient method for non-invasive and non-destructive plant phenotyping over the last two decades. Therefore to measure these traits in high-throughput and non-destructive way, computer vision-based methodologies are proposed. For recognition and counting of number of spikes from visual images of wheat plant, a deep learning-based encoder-decoder network is developed. The precision, accuracy, and robustness (F_1-score) of the approach for spike recognition are found as 98.97%, 98.07%, and 98.97%, respectively. For spike counting, the average precision, accuracy, and robustness are 98%, 93%, and 97%, respectively. The performance of the approach demonstrates that the encoder-decoder network-based approach is effective and robust for

A. Arora (✉) · S. Marwaha
ICAR-Indian Agricultural Statistics Research Institute (IASRI), New Delhi, India
e-mail: Alka.Arora@icar.gov.in

S. Marwaha
e-mail: Sudeep.Marwaha@icar.gov.in

T. Misra
Teaching Cum Research Associate, Rani Lakshmi Bai Central Agricultural University, Jhansi, UP, India

M. Kumar
Computer Application, ICAR-Indian Agricultural Statistics Research Institute (IASRI), New Delhi, India

S. Kumar · V. Chinnusamy
ICAR-Indian Agricultural Research Institute (IARI), New Delhi, India
e-mail: viswa.chinnusamy@gmail.com

S. Chaudhary et al. (eds.), *Digital Ecosystem for Innovation in Agriculture*,
Studies in Big Data 121, https://doi.org/10.1007/978-981-99-0577-5_13

spike detection and counting. For senescence quantification, machine learning-based approach has been proposed which segments the wheat plant into different senescence and greenness classes. Six machine learning-based classifiers: decision tree, random forest, KNN, gradient boosting, naïve Bayes, and artificial neural network (ANN) are trained to segment the senescence portion from wheat plants. All the classifiers performed well, but ANN outperformed with 97.28% accuracy. After senescence segmentation, percentage of senescence area is also calculated. A GUI-based desktop application, m—Senescencica has been developed, which processes the input images and generates output for senescence percentage, plant height, and plant area.

Keywords Computer vision · Deep learning · Image analysis · Machine learning · m—Senescencica · Plant phenomics

1 Introduction

Food grain production must be doubled by 2050 to meet the demand of growing population (Ray et al., 2013). This is a major challenge due to climate change induced stresses and slow rate of genetic gain in conventional crop improvement programs. High-quality crop varieties are crucial to counter the challenges. Plant phenotypic trait measurement is necessary to develop improved crop varieties. In this connection, plant phenomics refers to studying the plant's morphological and physiological characteristics. Plant phenotypic traits quantification of germplasm lines and mapping population in a given environment is necessary for gene mapping and trait pyramiding. Manual recording of physiological traits is used in traditional methods, which is time-consuming and labor-intensive, and may be error-prone when recording a large number of genotypes. Computer vision-based techniques have recently emerged as an efficient framework for non-destructive plant phenotyping. Plant phenotypic traits like number of spikes/panicle in cereal crops and senescence quantification play an important role in assessing functional plant biology, growth analysis, and net primary production. Conventional techniques of measuring these traits are tedious, time-consuming, and error-prone for phenotyping large dataset. In this chapter, computer vision-based approaches for measuring the phenotypic traits (number of spikes/panicle in cereal crops and senescence quantification) are presented.

2 Recognizing and Counting of Spikes in Visual Images of a Wheat Plant

Wheat spike is the grain-bearing organ, and spike number is the key measure in yield determination of the plant. Manual or conventional technique of counting number of spikes based on naked-eye observation is tedious and time-consuming to record large number of genotypes. Recently, computer vision (integration of image analysis and machine learning techniques)-based technologies acquire strong attention in recognition and counting of spikes through image processing. A computer vision-based approach is presented in this chapter for identifying spikes in visual images (VIS) of wheat plants.

2.1 Image Acquisition

In this study, visual images of the plant were taken using 6576 × 4384 pixel RGB camera from three different side view directions (angles: 0°, 120°, and 240°) with respect to the initial position of the plant. To reduce the issue of spikes overlapping, three side view directions were taken into consideration. This LemnaTec imaging facilities (LemnaTec GmbH, Aachen, Germany) are installed at Nanaji Deshmukh Plant Phenomics Center at Indian Agriculture Research Institute (IARI), New Delhi (Misra et al., 2020). Hundred wheat plants were grown in pot under controlled environmental condition with recommended cultural practices. Images were captured during reproductive stage of the plant by maintaining a uniform background for better image processing and stored in PNG format. After image acquisition, spikes number per plant pot was recorded manually to validate the developed approach.

2.2 Architecture of the Deep Learning Approach

The developed approach consists of two deep learning networks: Patchify Network (PN) and Refinement Network (RN). PN extracts spatial and local contextual features at patch level (a small area of the image that overlaps), while RN refines the segmented output of PN (as it sometimes contain segments spikes inaccurately) (Fig. 1). The convolutional encoder-decoder deep learning network with hourglasses serves as the approach's backbone and bottleneck network for pixel-by-pixel segmentation of objects (here, spikes). For retrieving the feature map depiction clutching the spatial and factual information from the instruction digital representation, an encoder network with three encoder blocks (each encoder block consisting of two convolution layers followed by a ReLU and Max-pooling layer of window size 2 × 2) was constructed. Three encoder blocks each had 16 filters, 64 filters, and 128 filters to encode the features. Three decoder blocks make up the decoder network, and each block's two-transpose convolution layers are, used to upsample the incoming feature maps. In contrast to the encoder network used to renew the features, there were 128, 64, and 16 filters in each of the three-decoder blocks, correspondingly. For better localization of the segmented features, the upsampled feature maps are combined with the appropriate encoded feature maps. The hourglass network is made of three hourglasses, each of which is made of a series of residual blocks that have three

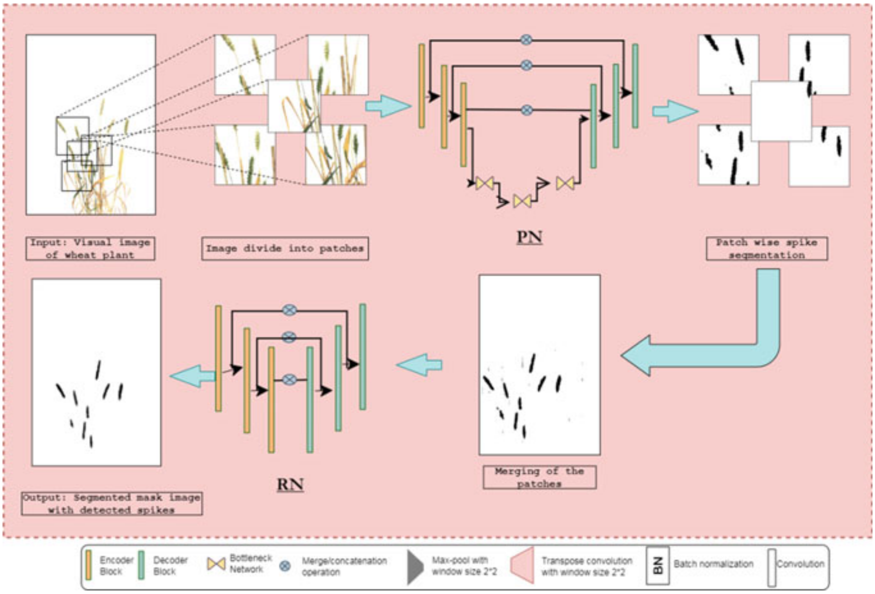

Fig. 1 Input image is split into patches before entering into PN. Patch-wise segmented mask images are the output of PN and are concatenated which contains some erroneous segmentation of spikes. The output are then refined using RN network. Refined mask image contains spike regions only

convolution layers with the filter sizes 1 × 1, 3 × 3, and 1 × 1. This allows for more confident spike segmentation by focusing on the key features that are affected by scale, viewpoints, and occlusions. Based on the optical performances, the number of encoders, decoders, and hourglasses was estimated on empirical evidence.

2.3 *Training of the Deep Learning Model*

For developing the deep learning model, image dataset comprising 300 images (3 direction images of 100 plants) was divided randomly into training dataset and test dataset at 85% and 15% ratio, respectively. The deep learning model was developed on Linux operating system with 32 GB RAM and NVIDIA GeForce GTX 1080 Ti graphics card (with memory 11 GB). The training dataset consists of 255 images (*i.e.,* 85% of the total dataset). The images were divided into patches before entering into the PN. The popular optimizer "Adam" with learning rate 0.0005 was used to update the weight of the network, and "binary cross-entropy" (Misra et al., 2021, 2022) loss function was utilized to for predicting the spikes and non-spikes pixels. As it is a binary class classification problem (spike pixels or not), "binary cross-entropy" is used to calculate the loss function at pixel level. Both the networks (PN and RN) were trained separately for 100 epochs with batch size 32 (due to the system

Table 1 List of hyperparameters

Optimizer	Learning rate	Epoch	Batch size	Loss function
Adam	0.0005	100	32	Binary cross-entropy

Table 2 Performance of SpikeSegNet in spike segmentation

E1	E2	JI	Accuracy	Precision	Recall	F-measure
0.0016	0.0487	0.9982	0.9807	0.9897	0.9889	0.9897

constraints) and then merged to confine a single network. The hyperparameters used in developing the model are given in Table 1.

2.4 Result

SpikeSegNet, which consists of both networks (PN and RN), was trained sequentially. The pixel-wise segmentation performance of the developed model was measured by the performance metrics (Type I classification error (E_1), Type II classification error (E_2), Jaccard index (JI), accuracy, precision, recall, and F-measure) presented in Table 2.

Precision = TP/(TP + FP); Recall = TP/(TP + FN); Accuracy = (TP + TN)/(TP + TN + FP + FN); TP: True positive; TN: True negative; FP: False positive; FN: False negative.

E_1 indicates that a very small number of pixels were incorrectly identified. The created model's accuracy is almost 99%, and spikes can be identified with an average precision of 98.97%. For spike counting, the "analyze particles"—function of imageJ (Abràmoff et al., 2004)—was applied on the output of SpikeSegNet network, i.e., binary image containing spike region only. For spike counting, the average precision, accuracy, and robustness are 98%, 93%, and 97%, respectively. The performance of SpikeSegNet indicates that it is an important advance forward toward high-throughput phenotyping of wheat plant.

In the next section, another approach has been presented for the senescence quantification.

3 Machine Learning-Based Plant Senescence Quantification

Senescence is the last stage of the wheat crop cycle, and it is at this time that nutrients start to flow back into the developing grain from the plant. The first and most significant change in wheat senescence is the damaging of chloroplasts, which result in breakdown of photosynthetic pigments such as chlorophyll in leaf (Nikolaeva et al., 2010). Due to damage in chlorophyll, color of the leaf changes from the usual deep green to yellow and finally brown (Fig. 3).

Measuring of plant senescence is an important aspect as this helps to select the best genotypes tolerable to senescence in the stressed conditions. Conventionally, senescence is measured by manual scoring, in which an expert assign senescence score by observing the plant. But this method has many drawbacks, first of all it is subjective in nature and highly biased. This method is time-consuming as in any breeding program, there is a large population of grown plants. Manually measuring senescence for such a large population of plants is time-consuming and prone to errors. With the availability of image data, image-based measuring of plant phenotypic parameters is gaining the interest of researchers. It is high throughput and non-destructive in nature.

Here, a computer vision-based approach has been proposed for plant senescence quantification (Kumar, 2020). This is the plant pixels classification problem to classify plant pixels into each of the defined classes. Six classes were decided for the senescence dry, yellow, pale yellow, dark green, and light green and one background class by observing the senescence pattern in Fig. 2. Around 1000 pixels values were sampled from the image data. Sampled dataset was divided in two training and test sets with 75–25% ratio, respectively. Six machine learning-based classifiers (ANN, naïve Bayes, random forest, gradient boosting classifier, decision tree, K-nearest neighbors) were trained on training data by using scikit-learn library. In order to get the best models parameters, tuning with tenfold cross-validation was used to select the best-performing models.

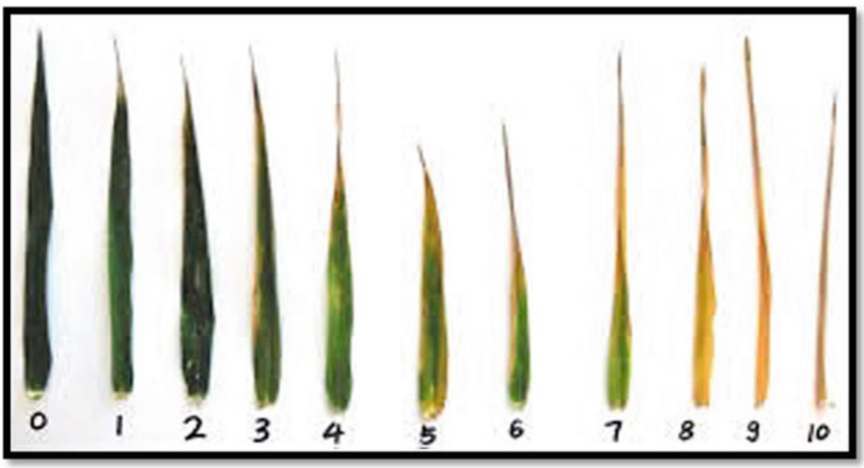

Fig. 2 Changes in leaf color due to senescence. Initially, the leaf is green, but due to senescence, it results in change in leaf color. It changes to pale yellow to yellow and at last to brown

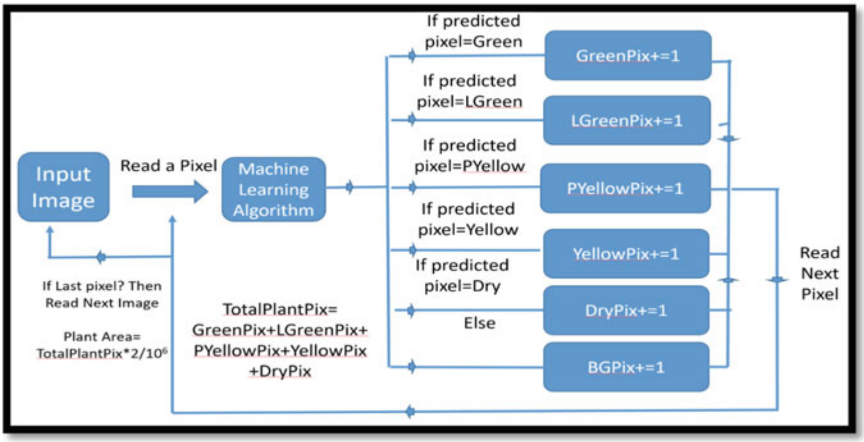

Fig. 3 Flowchart for senescence quantification

Precision, recall, and F1-scores were measured on test data. Among all the trained classifiers, ANN outperformed with 97.28% test accuracy. After pixel classification, the total number of pixels in each class was counted. Sum of all the pixels provides the total number of plant pixels (Fig. 3). Upon dividing total pixels in each class by total plant pixels gives the percentage of pixels in each class. Division by zero in Python causes zero division error, and this exception is handled by using try and catch block.

Among the six defined classes, yellow, pale yellow, and brown account for senescence classes. Hence, the percentage sum for those three classes gave the senescence percentage.

In this study, four approaches have been presented based on artificial intelligence in the area of plant phenomics. All these approaches gave promising results in the area of phenomics. Artificial intelligence techniques have tremendous potential in determination of other plant phenomics parameters with the ultimate goal for yield estimation.

4 Conclusion

In the era of modern plant phenotyping, computer vision-based technologies are much needed to counter the challenges that exist in traditional plant phenotyping of recognizing and counting spikes and senescence quantification in wheat plant. In this study, a deep learning approach has been developed for recognizing and counting spikes form visual images of wheat plant with satisfactory precision, accuracy, and robustness performances. Besides, machine learning-based approaches also performs a promising result in plant senescence quantification. As conventional phenotyping is the rate limiting step in utilization of vast genomics resources generated in different crops, hence, development of these techniques is critical for utilization of germplasm resources to develop high yielding and climate-resilient crop varieties. The methods developed in this study are cost—and time-effective and will be useful in both crop improvement and crop management. These approaches are significant step forward in the area of high-throughput wheat yield phenotyping and can be extended to other cereal crops also.

References

Abràmoff, M. D., Magalhães, P. J., & Ram, S. J. (2004). Image processing with ImageJ. *Biophotons Intention, 11*(7), 36–42.

Kumar, M. (2020). Wheat Plant Senescence Quantification using Machine Learning Algorithms (Master's thesis). Indian Agricultural Statistics Research Institute, IARI, New Delhi.

Misra, T., Arora, A., Marwaha, S., Chinnusamy, V., Rao, A. R., Jain, R., ... & Goel, S. (2020). 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.

Misra, T., Arora, A., Marwaha, S., Jha, R. R., Ray, M., Varghese, E., Kumar, S., Nigam, A., Sahoo, R. N., & Chinnusamy, V. (2021). Web-SpikeSegNet: Deep learning framework for recognition and counting of spikes from visual images of wheat plants. *IEEE Access, 9*, 76235–76247.

Misra, T., Arora, A., Marwaha, S., Ranjan Jha, R., Ray, M., Kumar, S., & Chinnusamy, V. (2022). Yield-SpikeSegNet: An extension of SpikeSegNet deep-learning approach for the yield estimation in the wheat using visual images. *Applied Artificial Intelligence, 36*(1), 2137642.

Nikolaeva, M. K., Maevskaya, S. N., Shugaev, A. G., & Bukhov, N. G. (2010). Effect of drought on chlorophyll content and antioxidant enzyme activities in leaves of three wheat cultivars varying in productivity. *Russian Journal of Plant Physiology, 57*(1), 87–95.

Ray, D. K., Mueller, N. D., West, P. C., & Foley, J. A. (2013). Yield trends are insufficient to double global crop production by 2050. *PloS one, 8*(6), e66428.