Fall 12-2-2022

Semantic Segmentation based deep learning approaches for weed detection

Puranjit Singh
University of Nebraska-Lincoln, psingh24@huskers.unl.edu

SEMANTIC SEGMENTATION BASED DEEP LEARNING APPROACHES FOR WEED DETECTION

by

Puranjit Singh

A THESIS

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Master of Science

Major: Agriculture and Biological Systems Engineering

Under the Supervision of Professor Yeyin Shi

Lincoln, Nebraska

December 2022

SEMANTIC SEGMENTATION BASED DEEP LEARNING APPROACHES FOR WEED DETECTION

Puranjit Singh, M.S.

University of Nebraska, 2022

Advisor: Yeyin Shi

Global increase in herbicide use to control weeds has led to issues such as evolution of herbicide-resistant weeds, off-target herbicide movement, etc. Precision agriculture advocates Site Specific Weed Management (SSWM) application to achieve precise and right amount of herbicide spray and reduce off-target herbicide movement. Recent advancements in Deep Learning (DL) have opened possibilities for adaptive and accurate weed recognitions for field based SSWM applications with traditional and emerging spraying equipment; however, challenges exist in identifying the DL model structure and train the model appropriately for accurate and rapid model applications over varying crop/weed growth stages and environment. In our study, an encoder-decoder based DL architecture was proposed that performs pixel-wise Semantic Segmentation (SS) classifications of crop, soil, and weed patches in the fields. The objective of this study was to develop a robust weed detection algorithm using DL techniques that can accurately and reliably locate weed infestations in low altitude Unmanned Aerial Vehicle (UAV) imagery with acceptable application speed. Two different encoder-decoder based SS models of LinkNet and UNet were developed using transfer learning techniques. We performed various measures such as backpropagation optimization and refining of the

dataset used for training to address the class-imbalance problem which is a common issue in developing weed detection models. It was found that LinkNet model with ResNet18 as the encoder section and use of ‘Focal loss’ loss function was able to achieve the highest mean and class-wise Intersection over Union scores for different class categories while performing predictions on unseen dataset. The developed state-of-art model did not require a large amount of data during training and the techniques used to develop the model in our study provides a propitious opportunity that performs better than the existing SS based weed detections models. The proposed model integrates a futuristic approach to develop a model that could be used for weed detection on aerial imagery from UAV and perform real-time SSWM applications.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor Dr. Yeyin Shi for her encouragement and support throughout my master's program. She motivated me, challenged me, and gave me freedom while working on my research project. It was undoubtedly a fruitful experience working with her and I look forward to even more during my Ph.D. in the years to follow. I also would like to thank my committee members Dr. Joe Luck, and Dr. Benjamin Riggan for their time serving on my committee and reviewing the thesis document and providing invaluable feedback for the successful completion of this project.

Finally, I would like to acknowledge my family who enabled me to be in this distant land and trusted in me despite all my shortcomings, have always been an inseparable part of my life. I am indebted to my mother – Dr. Paramjit Kaur and my brother – Dr. Jasreman Singh for their relentless support of my decisions and motivating me throughout my master's program. This journey would have been a lot more challenging without the company of kind colleagues, friends, and family. Here are some I would like to give a heartfelt shoutout: Dr. Nilovna Chatterjee (my sister-in-law), Rashmeet Kaur, Gautam Takoo, Dr. Benjamin Hage for great times and meaningful conversations.

GRANT INFROMATION

My research grant was provided by a grant from United States Department of Agriculture (USDA) National Institute of Food and Agriculture (NIFA) – An Intelligent Unmanned Aerial Application System for Site-Specific Weed management (Award Number #2021-67021-34412) and the Department of Biological Systems Engineering at University of Nebraska-Lincoln.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

The current world population of 8 billion people is expected to increase by around 1.8 billion and reach 9.7 billion by 2050 (https://www.fao.org). Increasing world population trends demand an increase in overall crop production and measures to decrease crop yield losses. There exist various biotic constraints such as weeds, pests, bacteria, fungus, etc., and abiotic factors such as change in soil moisture content, excessive rain, drought etc., that could dwindle crop yield (Oerke et al., 2006). According to a survey conducted by (Adeux et al., 2019), farmers could suffer an estimated 19-56 % decline in crop yield if weeds were left untrammeled in the fields. Weeds are one of the most important biotic constraints that compete with crops for water, sunlight, nutrients, and space. Also, weeds destroy the natural habitat of plants and animals in the fields as they provide a haven to insects and pathogens in the fields (Chauhan et al., 2020). Several chemical and mechanical weed control measures have been used to manage weeds in agriculture. Mechanical weed control techniques like tilling, hoeing, hand removal are a few practical practices which are used to handle weed infestations in small farmlands in developing countries. In developed countries like USA and Canada where farmers have large farmlands, these manual weeding techniques are not practically feasible to handle weed infestations. The most popular operandi to cover larger field areas is the use of uniform spraying of chemicals like herbicides in the fields to perform weed killings. Glyphosate-based Herbicides (GBHs) are the most used herbicides in the fields as they are capable of handling multiple weed species that grow in fields. There exist a few drawbacks associated with practicing of uniform spraying herbicide applications in the fields such as increase in economic burden due to costs associated with herbicides and human labor, environment pollution due to use of chemicals in fields, and evolution of herbicide resistant weeds in the fields which are more harmful (Veeranampalayam et al., 2020). Recently, it has been proven that excessive use of GBHs in fields lead to decrease in crop nutrition quality and disruption to rhizosphere microbial ecology (Martinez et al., 2018).

Precision agriculture suggests various Site-Specific Weed Management (SSWM) strategies in performing the weeding applications at locations where weed infestations are detected in the fields. Computer vision techniques have been extensively used to extract useful information from the data acquired using various remote sensing techniques such as Satellites, Unmanned Aerial Vehicles, Ground robots to generate weed maps that could be used to perform the tasks of SSWM applications. (Wu et al., 2021). Traditionally machine vision and image processing techniques were used to develop weed detection models, these models were trained to detect weed infestations that existed in-between crops row spacing and did not perform well in classifying intra-row weed infestations (Perez et al., 2000). The availability of high-resolution imagery collected using Unmanned Aerial Vehicles hovering at lower altitudes have helped in extraction of useful features which have been used with various Machine Learning algorithms such as Support Vector Machines, Decision Tress, Random Forests to develop weed detection models in recent past (Liakos et al., 2018). Due to their better learning capabilities on features extracted from these images, ML algorithms helped achieve higher accuracy scores to perform the classifications tasks between crops and weeds (Bakhshipour et al., 2018). However, training a weed detection has been a complex task due to the homogenous crop-weed characteristics such as color similarity, high occlusions, similar reflective index, similar shape, and texture features (Hasan et al., 2021).

DL techniques using CNN have revolutionized computer vision tasks by creating state-of-art models to perform the complex tasks of image classification, object detection, semantic segmentation, and instance segmentation (Girshick et al., 2015; Ren et al., 2015; Liu et al., 2016; Ronneberger et al., 2015; He et al., 2017). CNN based deep learning approach have gained much importance over the past few years in performing classifications and detection tasks because of their feature generation and better self-learning capabilities in comparison with Machine Learning algorithms and traditional conventional image processing or vegetation index-based approaches. The training of DL models requires systems with high configurational capabilities to perform the complex computations required during the training phase. The availability of systems having better computer hardware, software capabilities and use of GPUs have helped in training DL models to achieve higher accuracy scores. Weed detection models have been developed using Object Detection and Semantic Segmentation techniques (He et al., 2017; Ronneberger et al., 2015). Object detection performs the task of generating bounding boxes around detected objects during prediction and Semantic segmentation models generate pixel-wise mapping of different class categories detected during predictions. The Semantic Segmentation approaches have gained much importance in recent past as they better perform the tasks of scene understanding and help in generating better accuracy scores during predictions. Due to their better scene understanding capabilities these models have been used for real-time applications such as self-driving cars, and augmented reality (Hao et al., 2020; Garcia-Garcia et al., 2017). Pixel-wise classifications using Semantic Segmentation approaches are not fully exploited in resolving weed detection tasks due to unavailability of data to be used for training these models. A major challenge faced while training Semantic Segmentation models is the class-imbalance issue, which is a common concern while training Deep Learning model. This study proposes a two-step approach to handle the class imbalance problem and generate optimized Semantic Segmentation based Deep Learning models to generate weed maps with better accuracy scores. The two-step optimized strategy is explained in detail in Chapter 3 of this thesis.

This thesis has been organized as three chapters with Chapter 1 as the Introduction. Chapter 2 gives a detailed review on the Computer Vision techniques that have been used to develop real-time weed detection models in the past two decades. Chapter 3 presents research conducted to develop a robust weed detection model using Semantic Segmentation based Deep Learning approach.

1.1 References:

Oerke, E. C. (2006). Crop losses to pests. The Journal of Agricultural Science, 144(1), 31-43. https://doi.org/10.1017/S0021859605005708

Chauhan, B. S. (2020). Grand challenges in weed management. Frontiers in Agronomy, 1, 3. https://doi.org/10.3389/fagro.2019.00003

Adeux, G., Vieren, E., Carlesi, S., Bàrberi, P., Munier-Jolain, N., & Cordeau, S. (2019). Mitigating crop yield losses through weed diversity. Nature Sustainability, 2(11), 1018-1026. https://doi.org/10.1038/s41893-019-0415-y

Veeranampalayam Sivakumar, A. N., Li, J., Scott, S., Psota, E., J Jhala, A., Luck, J. D., & Shi, Y. (2020). Comparison of object detection and patch-based classification deep learning models on mid-to late-season weed detection in UAV imagery. Remote Sensing, 12(13), 2136. https://doi.org/10.3390/rs12132136

Martinez, D. A., Loening, U. E., & Graham, M. C. (2018). Impacts of glyphosate-based herbicides on disease resistance and health of crops: a review. Environmental Sciences Europe, 30(1), 1-14.

Feyaerts, F., & Van Gool, L. (2001). Multi-spectral vision system for weed detection. Pattern Recognition Letters, 22(6-7), 667-674. https://doi.org/10.1016/S0167-8655(01)00006-X

Bakhshipour, A., & Jafari, A. (2018). Evaluation of support vector machine and artificial neural networks in weed detection using shape features. Computers and Electronics in Agriculture, 145, 153-160. https://doi.org/10.1016/j.compag.2017.12.032

Girshick, R. (2015). Fast r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 1440-1448). https://doi.org/10.1109/ICCV.2015.169

Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, 28. https://doi.org/10.1109/tpami.2016.2577031

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016, October). Ssd: Single shot multibox detector. In European conference on computer vision (pp. 21-37). Springer, Cham. https://doi.org/10.1007/978-3-319-46448-0_2

Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham. https://doi.org/10.48550/arXiv.1505.04597

He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 2961-2969). https://doi.org/10.48550/arXiv.1703.06870

Hao, S., Zhou, Y., & Guo, Y. (2020). A brief survey on semantic segmentation with deep learning. Neurocomputing, 406, 302-321. https://doi.org/10.1016/j.neucom.2019.11.118

Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., & Garcia-Rodriguez, J. (2017). A review on deep learning techniques applied to semantic segmentation. arXiv preprint arXiv:1704.06857. https://doi.org/10.48550/arXiv.1704.06857

Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. Sensors, 18(8), 2674. https://doi.org/10.3390/s18082674

Hasan, A. M., Sohel, F., Diepeveen, D., Laga, H., & Jones, M. G. (2021). A survey of deep learning techniques for weed detection from images. Computers and Electronics in Agriculture, 184, 106067. https://doi.org/10.1016/j.compag.2021.106067

Wu, Z., Chen, Y., Zhao, B., Kang, X., & Ding, Y. (2021). Review of weed detection methods based on computer vision. Sensors, 21(11), 3647. https://doi.org/10.3390/s21113647

Chapter 2: Computer Vision techniques for weed detection in agricultural production and natural resource management: A Review

This manuscript has been prepared for journal submission

2.1 Introduction

Agriculture in the 21st century faces multiple challenges: the major concerns are to mass-produce abundant food and fiber to feed a growing population with a smaller rural labor force. The need of the hour is to embrace more efficient and sustainable production methods and adapt to climate change, which could contribute to overall development in the many agricultural-dependent developing countries in the world. The current global population of 8 billion is expected to reach over 9.8 billion by 2050 [3]. To feed this rise in population the overall food production would need to be increased by an estimate of around 70-100%. This implies a significant increase in the production of several key commodities such as cereal production and meat production, considering the required nutritional content [1]. Accordingly, in addition to breeding a higher-yielding variety of crops, it is vital to address the numerous factors affecting crop yield loss in the agriculture community.

Computer vision technologies have been used to resolve many Precision Agriculture (PA) tasks, such as crop yield prediction, water and soil conservation, plant disease detection, weed detection, quality evaluation of crops, and species identification [31-48] to help fetch insights that directly or indirectly helps in increase of crop yield. Table 1 lists various computer vision-based applications in agriculture that were developed recently. PA has proposed site-specific spraying techniques that could prevent waste and herbicide residual problems caused by traditional spraying practices [25]. PA is an approach to farm management that uses Information Technology (IT) to ensure that crops and soil receive exactly what they need for optimum health and productivity [27].

At present, the development of robust real-time applications performing accurate identification of crops and weeds is among the major focuses in Site-Specific Weed Management (SSWM). Training weed detection models using traditional image-processing methods involves extraction of useful features, such as color, texture, and shape, and incorporating them with various image analyses [61-71], and ML [46-49] algorithms such as SVM [28], RF [29], ANN [44], K-Means [30] etc. These methods are trained on features fed manually by the user and have a high dependency on image acquisition, quality of feature extraction, and pre-processing methods. Improvement in computational power, increase in data volume, and advancement in technology, has led to the increase in use of DL algorithms to solve several Computer Vision oriented applications in the agriculture domain. DL methods [31-38] consisting of CNN have been extensively used for detection and classification predictions because of their feature extraction and self-learning capabilities. Due to their enhanced data expression capabilities for images, these models have helped to develop more robust detection models with better accuracy results. avoiding the disadvantages of traditional extraction methods, these methods have gained much attention in the recent past.

Several reviews have already been published which emphasize comparing weed detection models developed using various computer vision approaches [2, 3, 46, 49]. These reviews provide an extensive overview of the strategies used to develop these models along with the pros and cons associated with them. Numerous real-time weed management applications have been developed in recent past, which make use of unique strategies to develop these models. Novel strategies that could be used to develop real-time site-specific weed management applications and the pros and cons associated with existing real-time weed management applications are yet to be discussed in a review. This review provides a comprehensive overview of various CV methods used for developing weed detection in the past three decades, including real-time CV methods developed in more recent years. Lastly, the challenges of current CV techniques for weed detection and some opportunities and trend of future research prospects are discussed that could be used to strategically plan and improve PA in weed management for their real-time site-specific weed management applications.

This review paper is divided into seven sections. The first section briefly explains about the world population trend, the required increase in demand for agriculture production for upcoming decades, the major causes of crop yield loss, and a few solutions to increase crop production in the 21st century. The second section explains the strategy that was adopted to fetch relevant articles to our research domain from available scientific databases. Section 3 gives a brief description of weeds, crop yield loss due to weeds, prevention and management strategies, and some common categories of weeds found in different crops. Section 4 gives a detailed description of traditional approaches that were used to develop weed detection models. ML and DL approaches that were used in the development of weed detection models and their advantages and disadvantages are also discussed along in Section 4 as well. Section 5 reviews the real-time applications of weed detection models that have been developed and tested in agriculture fields for weed detection and discusses probable future directions in utilizing Computer Vision techniques for real-time applications in the development of weed detection models to surge agriculture production and conserve natural resources. Lastly, section 6 presents the summary and conclusions of our review.

2.2 Methodology

This section explains the research questions, search strategy, and selection criteria based on which certain research articles were selected, and a comprehensive review was done on various weed management practices developed over the years.

The analysis of these research papers involved two main steps:

1. Collection of related research articles in the domain, in-depth analysis, and
2. Review of the research performed in their work

Firstly, to collect associated articles involving the development of weed identification models using various Computer Vision approaches, a keyword-based search was performed to gather both conference papers and journal articles on kindred subjects in PA. The main research question is: What is the role of Computer Vision techniques to form a weed map and real-time applications developed for SSWM using the weed map? Weedmaps could be referred to as the classification and detection of weed

Table 2. 1 Computer vision applications in Agriculture

Crop	Input data	Application	Algorithm used	Reference
Cotton	Satellite Spectra data	Cotton yield estimation	ANN	[31]
Apple	RGB images	Apple flowers and fruits counting	DL: CNN	[32]
Maize	Weather and Satellite Spectra data	Silage Maize yield estimation	DL: CNN	[33]
Wheat, Barley	Multispectral images from UAV	Prediction of Wheat and Barley yields	DL: CNN	[34]
Turmeric	Weather data	Predicting oil yield from turmeric rhizomes	ANN	[35]
Pistachio	Soil characteristics	Estimation of pistachio yield in orchards	ANN	[36]
Sunflower	Plant height, SPAD	Sunflower seed yield estimation	PLSR, ANN	[38]
Rice	Satellite Spectral data	Prediction of rice crop yield	ML: RF, SVM	[42]

Mango	Multispectral data from UGV	Prediction of mango maturity level	ML: SVM	[39]
Rice	Weather data, Irrigation, Planting area, fertilization	Prediction of the paddy field yield	ANN, ML: RF, KNN, MLR	[40]
Soybean	Weather and Satellite spectral data	Prediction of soybean in 15 states of USA	DL: CNN, LSTM	[41]
Sugarcane	Satellite spectral data	Prediction of sugarcane yield	MLR	[43]
Rice, Maize, Millet, Ragi	Weather data	Prediction of various Kharif crop yields	MANN, SVR	[44]
Various crops	RGB images	Detection and diagnosis of plant disease	DL: CNN	[45]
Avocado	Hyperspectral images	Detection of nitrogen and iron deficiencies in avocado	ML: DT, MLP	[46]
Peach	Hyperspectral images	Estimation of soluble solid content	ML: SAE-RF	[47]
Jatropa Curcas	X-ray imaging	Prediction of vigor and germination	LDA	[48]

Table 2. 2 Abbreviations table

Abbreviation	Full form
ML	Machine Learning
DL	Deep Learning
UAV	Unmanned Aerial Vehicles
ANN	Artificial Neural Network
LSTM	Long Short-Term Memory
SPAD	Soil and Plant Analyzer Development
PLSR	Partial Least Square Regression
UGV	Unmanned Ground Vehicle
DT	Decision Tree
USL	Unsupervised Learning
SL	Supervised Learning
MLR	Multiple Linear Regression
SVR	Support Vector Regression
MANN	Modular Artificial Neural Networks
LR	Linear Regression
RGB	Red-Green-Blue
XGBoost	Extreme Gradient Boosting
MLP	Multi-Layer Perceptron
KNN	K-Nearest Neighbors
DNN	Deep Neural Networks
MLNLM	Multi-layer Neural Network with No-Linear Mapping
IR	Infra-Red
NIR	Near Infra-Red

CART	Classification and Regression Tree
OD	Object Detection
SS	Semantic Segmentation
CNN	Convolutional Neural Network
LSTM	Long-Short-Term Memory
GAN	General Adversarial Network
RNN	Recurrent Neural Networks
FCN	Fully Convolutional Network
RPN	Region Proposal Network
GPU	Graphical Processing Unit
ResNet	Residual Networks
LoR	Logistic Regression
PA	Precision Agriculture
SSD	Single Shot Detector
YOLO	You Only Look Once
USA	United States of America
IoT	Internet of Things
COCO	Common Objects in Context
PSPNet	Pyramid Scene Parsing Network
RCNN	Regional Convolutional Neural Network
IoU	Intersection over Union
mAP	Mean Average Precision
LiDAR	Light Detection and Ranging
ILSVRC	ImageNet Large Scale Visual Recognition Challenge
TL	Transfer Learning
AWS	Amazon Web Services

patches with respective coordinates information in the agriculture fields. The articles were searched from 7 bibliographic databases: *Google Scholar, Research Gate, Science Direct, MDPI, Wiley Online Library ProQuest, and the University of Nebraska-Lincoln: Library databases*. The primary keyword 'Computer Vision' was paired with the secondary keyword 'weed detection. These keywords were inserted into the databases and related articles were selected that were included in this review. To ensure that no related articles were missed, a hand search was also performed to find related articles in the same domain.

2.3 What are weeds? A few Prevention and Management Strategies

The various constraints to agricultural production or crop yield loss could be classified into the following basic categories: biotic and abiotic constraints in addition to socioeconomic and those related to crop management [14]. Plant breeders and researchers have continuously produced new innovative, resilient, and novel ideas to overcome these constraints in the past few decades [15, 16]. Amongst the biotic factors, weeds represent the most significant constituent that leads to the highest potential crop yield loss along with other pathogens such as bacteria, fungi, etc. Weeds are unwanted plants that grow in crops and compete with them for water, sunlight, nutrients, and space. In addition, they also harbor insects and pathogens which attack crop plants, destroying their native habitats, which could lead to detrimental effects on the quality of crops if uncontrolled. Potential crop yield loss without weed control is about 43% on a global scale [10]. In the

USA alone, crop yield losses due to weeds sum up to an estimated cost of around 33 billion USD [9].

Pesticides such as herbicides have revolutionized and contributed significantly to weed management practices for the last 65 years and have helped in increasing crop yield [52]. Traditionally, weeds were controlled by uniform rate spraying of herbicides in the fields. This practice has a few disadvantages such as safety and environmental pollution, harm to humans, and the evolution of herbicide-resistant weeds which is a major issue of concern [26]. SSWM techniques could bridle the issue of uniform spraying of herbicides in fields to spraying only at desired locations where weeds were detected by machinery or equipment embedded with technologies that could detect weeds in crops [13]. This practice could help overcome the challenges associated with the uniform spraying of herbicides in the fields. Glyphosate has been commercially available since 1974 and is the most effective herbicide ever discovered [52]. Recurrent use of herbicides to control weeds has led to the emergence of legion Herbicide Resistant (HR) weeds in different crops. The first discovery of triazine-resistant weeds was found in western Washington in the 1960s [49] and as per [weedscience.org] 513 unique cases of HR weeds were discovered globally in 2022. Table 3 gives a few examples of common HR weeds that grow in different crops. Also, researchers have discovered various weed biotypes which are resistant to more than one herbicide (i.e., cross-resistance).

Continuous improvements need to be made to improve crop production. The traditional practices including burning, hand sowing, manual spot spraying, considering

Table 2. 3 Glyphosate Resistant weeds worldwide (Reproduced from weedscience.org)

Weed Species	**Crops found**	**Country**
Amaranthus Palmeri	Cotton, Corn, Soybean, Alfalfa, Roadsides, Grapes	USA
Amaranthus tuberculatus	Corn, Soybean	USA
Ambrosia trifida	Corn, Soybean, Cotton	USA
Bromus diandrous	Fence lines	Australia
Chloris truncata	Cropland – cereals	Australia
Conyza canadensis	Corn, Cotton, Rice. Soybean, Wheat, Maize	USA, China, Czech Republic, Greece, Canada, Poland, and Brazil. Spain, Italy
Eleusine indica	Coffee, Cotton, Soybean	Malaysia, Colombia, USA, China
Kochia scoparia	Cereals, Corn, Cotton, Soybean	USA, Canada
Lolium perenne	Barley, Cropland, Grapes, Soybean,	Argentina, New Zealand
Plantago lanceolata	Grapes, Orchards	South Africa
Poa annua	Golf courses, turf	USA

primary tillage, regularly scouting of fields to identify weed presence, rotation in herbicide usage, herbicide pre-emergence or post-emergence application, and repetitive blade hoeing are not practical anymore. These practices impacted the ecosystem rather than benefiting production. The evolution in technology using computer vision applications has led to the development of multiple smart agriculture practices that have made the lives of farmers easier and more comfortable and are discussed in the following sections. The transformation in SSWM applications using various computer vision techniques has been discussed in the following sections of this study.

2.4 Computer Vision for Weed Detection

Weed management is a complex, information-intensive task. Important factors such as the emergence of weeds, the crop's ability to suppress weed growth, effects of different weed species on crop yield need to be taken into consideration to understand the effects of weeds on crops [53]. Large quantities of herbicides are being used to control weeds in fields, gardens, ponds, golf courses, roadsides, and sports fields annually, which leads to environmental pollution and economic concerns. Traditional practices of removing weeds such as manual weeding and the usage of weeding tools to remove weeds are still being practiced for centuries and are still being used by farmers in owning small fields in developing countries [55]. Inefficiency, labor intensive, and high labor-cost are a few major challenges associated with manual weeding practices which make them impractical for large farmlands [54]. Advancement in the field of technology has led to the development of numerous Computer Vision based SSWM techniques that could be used as precise and real-time weed detection applications. These promising tools could replace and overcome the challenges posed by traditional weed management practices. In comparison with manual weeding, these approaches are much more efficient, decrease human labor and cost associated with it, and reduce the usage of herbicides in the fields. The result of a weed detection model is a weed map that could be used for site-specific treatment of weeds [58].

The main intuition behind SSWM is spraying of required herbicide composition at only specific weed patches that are detected. Typically, the development of SSWM practices usually requires 5 important steps to be followed [56, 57]:

1. Collection of data using various remote sensing techniques
2. Pre-processing and utilization of remote sensing data in preparation of Weed detection model and generating a real-time weed map.
3. Decision-making to decide on an action for weeding on the previously detected information and farmer experience
4. Weeding application based on the decision via an actuator
5. Performance evaluation of the precision operation

Amongst these 5 processes, weed detection plays a critical role in SSWM as it involves the algorithm development for the weed detection model and provides information to the successive processes that help in decision making.

RS techniques have been intensively used for numerous PA applications such as crop monitoring, irrigation management, predicting weather conditions, observing air quality, disease and pest management, and yield prediction in the past few decades. RS is the process of detecting and monitoring the physical characteristics of an object without coming into direct contact with the object. Several sensing methods have been used to fetch data using airborne, ground-based, and satellite techniques. These methods can be classified into ground-based, and airborne sensing techniques. Airplanes, UAVs (also drones), and satellites are used for data acquisition of airborne remote sensing data. The data from these pieces of equipment were collected and after proper examination and pre-processing, the data were used as an input for training weed detection models. Multi/Hyper-spectral imaging [39, 46, 47], Visible and near Infra-red (Vis-NIR) spectroscopy [57], X-ray imaging [48], Satellite Spectral data [31, 41-43], distance sensing techniques (Light Detection and Ranging -LiDAR) [29], and RGB images [32, 45] were used to extract important features such as plant spectral characteristics, leaf color, height, texture, shape, etc. and were utilized to train weed detection algorithms using computer vision and image analysis technology. The development of numerous weed detection algorithms using these techniques were carried out to perform classification tasks between different vegetation species.

2.4.1 Traditional Computer Vision approaches

Detecting broad-leaved weeds in cereal crops under actual field conditions were performed using color information and shape features from the images. Two different pattern recognition methods: KNN and Bayes rule were trained on color information and shape features extracted from various cereal crops and weeds and their performance was compared [61]. Average classification success rates (crop/weed) and corresponding variances were evaluated and the performance of 5 different classification algorithms was compared and rained on Spectral range, slit width, and slit length features extracted from 6 weeds and sugar beet crops respectively. Performance comparisons between 5 different classification algorithms namely: ED, MD, KNN, CART, and MLNLM were done. MLNLM performed the best classification tasks between crops and weeds with a classification success rate of 85.8% [62]. Reflectance spectra of crop and weed canopies were used to evaluate the possibilities of weed detection with reflection measurements in laboratory circumstances. Sugar beet and maize and 7 weed species were included in the measurements and their wavelength bands were compared to obtain 97% accuracy in performing classification between weeds and crops respectively [63]. A computer vision system comprising an ANN was trained using shape features extracted from binary images collected from images of radish and weeds. The neural network developed was able to achieve a successful recognition rate of 93.3% for Radish and 93.8% for weeds [64]. Soil-crop segmentation was done with two spectral channels, chosen from 100 channels available from the hyperspectral sensor, and weed detection was based on texture features extracted from segmented images [65]. Image pixels of the crop (Sugar beet) and weeds (4 species) were classified using the differences in spectral characteristics of plant species [66]. A combination of plant height and spectral reflectance features was used to develop an algorithm to perform classification between crops and weeds in organic carrots [67].

Some successful real-time weed detection models were also developed, and prototype testing was performed in agriculture fields to check and compare their performance in performing site-specific spraying of herbicides in the fields. Robotic weed controllers were developed by Blasco et al., [68] in Spain, which included two vision systems. The primary vision system helped the robot to distinguish between crops and weeds and the second vision system helped to retrieve the exact coordinates of weed presence, where an electrode powered by a set of batteries killed weeds by giving an electric discharge of 15000V [68]. Another robotic platform that could adapt and operate in between row crops of 0.25m and 0.5 m was developed by [69] with cameras for row guidance and weed detection. The robot included improved modules such as four-wheel steering and propulsion which helped to attain better mobility from adjustments in orientation.

Development of weed detection model using various spectral features and physical features such as shape, color, texture, and height encountered numerous challenges such as slow processing speeds, proper illumination during data collection, large memory requirements, and high costs of systems hardware. These were a few difficulties that were faced by researchers in the development of these detection models. The detection algorithms developed were not able to achieve accurate real-time applications during prototype testing. The presence of variable soil backgrounds and residue cover complicates the spectral response and hinders the response of vegetative presence [70]. This approach was used by plenty of detection algorithms that compared the vegetative spectral reflectance of the weed areas in the fields. Detection algorithms' performance differed in detection during different times of the day with variation in the lighting in which the remote sensing data was collected which was also a drawback of these approaches [71]. Advancements in the field of technology led to the development of better systems that could perform complex calculations and operations that led to the development of better architectures and algorithms that could perform arduous tasks of object classification and detection. Also, improvements and inventions of various Machine Learning and Deep Learning techniques led to solving challenges and tasks that seemed impossible in the past.

2.4.2 Machine Learning-based Computer Vision approaches for weed detection

ML has contributed significantly to resolving Computer Vision tasks in the agriculture domain such as yield prediction, disease detection, weed detection, and species identification, and has emerged significantly with big data technologies and high-performance computing. ML is defined as an application of AI that empowers a system to learn and improve its performance from experience. It majorly focuses on developing programs that can access data and use it to perform certain tasks such as classification, regression, or clustering. In this section, we present a review of ML applications developed in agriculture specifically for weed detection. The performance of an ML task is evaluated using an evaluation metric that is calculated after regular intervals while training an ML model. ML approaches have been extensively used in the previous decade to perform tasks such as Identification, Classification, Quantification, and Prediction in the agriculture domain [4, 79].

ML tasks could be broadly classified into 2 categories based on the learning approach that is used to train the model: Supervised [76-87] and Unsupervised Learning [72-74]. Also, ML models could be trained using different learning algorithms such as KNN, DT, RF, SVM, ANN, DL. Table 5 presents a detailed description of the weed detection models developed using the ML approach in recent past.

2.4.2.1 Supervised vs Unsupervised ML approaches

ML tasks are classified into two main categories: Supervised and Unsupervised Learning. The major difference between these methods is the approach in which learning is performed to improve the performance of the ML. In supervised learning, data is presented to the model in a manner that explicitly defines both input and desired output associated with it. In simple words, it could be stated that a well-defined labeled dataset is provided for training and the model is expected to learn the ability to distinguish an output given input during testing from knowledge gained during training. Also, another category in supervised learning is reinforcement learning where training is performed based on the feedback that is provided by the user training the model. Positive feedback is given whenever the model performs as expected and negative feedback is given whenever a model deviates from expectations during the training of a reinforcement learning model. To compare performance on how well an ML performs on unseen data, the original dataset available for training is split into two separate independent folders to be used for training and testing. Training of the ML model is performed only using the training set and the performance is evaluated by comparing the predicted output with the actual output during testing. This helps to evaluate how well a model is generalizing on unseen data. Unsupervised learning is a process in which the ML model predicts by itself how to process and label the unlabeled or raw data that is fed to the model for training. No information about the relationship between data values is provided during training unsupervised learning ML models. Supervised Learning is the most used approach for training ML models [71].

2.4.2.2 Clustering Algorithms

Clustering is an unsupervised classification ML algorithm. K-mean clustering like nearest neighbors could be referred to as a method of clustering objects with similar feature information/characteristics in a cluster. This approach has been used by various researchers to perform tasks of weed detection. [72] used an unsupervised ML approach of clustering plants with similar shapes in one group. Data points with similar feature information formed clusters and these were used for the task of weed detection. Zhang et al., [73] presented a weed detection algorithm consisting of three sequentially linked phases that involved, image segmentation, feature extraction, and crop-row detection. Crop-row detection in this approach used made use of a clustering algorithm to form separate clustered feature point sets of crops and row spacing. Clustering has been used to perform classification between weeds and crops using 16 features that were extracted using image processing techniques from the images that were collected from fields. Also, this approach was used to generate weed maps in less complex cases using GPS data [74]. KNN classifiers were used to perform weed identification tasks in low infestation level imagery collected using remote pilot aircraft on sugarcane plantations and overall accuracy of 83.1% was achieved using a kappa coefficient of 0.775 [78].

2.4.2.3 Regression Algorithms

Regression models supervised ML models that perform the prediction of detecting weeds in an image using a set of independent features that are extracted from the images. Major regression algorithms are Linear Regression (LR), Multiple Regression, and Logistic Regression (LoR). Multiple regression [76] analyses were used to find a relationship between ultrasonic readings and the corresponding coverage of crops and weeds. LR and LoR ML models were trained using weed density information and performance comparisons were made on their predictions to detect the presence/absence of weeds in the field [75]. These models were used to generate weed maps of annual grass weeds in the later stages of cereals. Feature extraction methods in combination with various ML models such as LoR [77] models were used to perform classification tasks between crops and weeds.

2.4.2.4 Support Vector Machines

SVMs are supervised ML algorithms that perform the mapping of high-dimensional data points and divide them into classes to form a Maximum Marginal Hyperplane. SVMs help to resolve the overfitting problems that appear in high-dimensional spaces, which makes them a popular ML algorithm used by researchers to perform various classification, clustering, and regression tasks [4]. SVM classifiers using features extracted from Gabor and FFT filters have been used to perform the task of weed detection to detect broad and narrow weeds [80]. Fourier descriptors and moment invariant features together with several shape features were used to train the supervised SVM classifier to classify four species of weeds in sugar beet that achieved an overall accuracy of 95% [81]. An overall accuracy score of 97% was achieved on the testing performance of the SVM classifier trained on 14 optimal features extracted from digital images of crops and weeds [82]. The SVM classifier was able to perform complex classification between avena sterilis weed and cereal crops that shares similar spectral signature successfully even with minimum system memory requirements and computation power [83].

2.4.2.5 Decision Trees

DTs are supervised ML algorithms that could be referred to as probability tree-like consisting of the root node, branches, and internal node that continuously split data to categorize or make predictions using the set of features that are used for training the model. During training DT gradually tries to fit in the objects with similar characteristics under a common root node. DT algorithms were used to extract the shape and texture features of eight species of plant leaves using hyperspectral images to perform classification between corn and weeds under laboratory conditions that achieved an accuracy of over 95% [84]. Hyperspectral data of corn plots were used to classify into 3 separate categories namely: weeds presence, water stress, and nitrogen application rates using DT algorithm [85].

2.4.2.6 Ensemble Learning

Ensemble learning refers to training multiple ML models on different samples from training datasets and aggregating responses from all the models to generate a final output. Ensemble learning generate better ML models that could generalize well on unseen dataset and have proven to overcome overfitting problems. The two common categories under which ensemble learning techniques can be classified are: Bagging and Boosting. Bagging is an ensemble learning technique in which several training samples from a dataset are independently trained on multiple ML algorithms and are aggregated via an appropriate combination technique to perform final prediction task. Boosting is an ensemble learning technique in which multiple ML algorithms are trained sequentially with an intuition that the successor learns from the miss-classifications or errors generated during training phase. This approach has resulted in improving generalizability during prediction from the trained model. One of the powerful supervised ML algorithms that uses ensemble learning techniques are RFs. RF constructs multiple DTs during training and uses an aggregation of outputs from all DTs in a combination to perform the final classification and regression tasks during prediction. RF classifiers were used to perform weed species recognition tasks on the imagery that were collected at low altitudes using UAVs in presence of lower infestation levels in sugarcane fields [86]. The research conducted by [87] has developed RF-based ML algorithms to detect alligator weeds that form dense infestations in aquatic environments using remote sensing data to improve biosecurity supervision and monitoring efforts in Australia. ML algorithms that use gradient boosting techniques such as XGBoost have been used to train state-of art weed detection models to distinguish between Buffel grass and spinifex with accuracy scores of 97% have already been developed in recent past. Predictions performed in different case scenarios such as object rotation, illumination changes, background cluttering suggest the robustness of the ML models developed using gradient boosting techniques [106].

2.4.3 Deep Learning based Computer Vision approaches for weed detection

Evolution in the field of developing better computer hardware and software has helped in the amelioration of the computing power of GPUs that play a vital role in performing multiple, simultaneous computations during the training of ML and DL models. DL algorithms have been used to perform various tasks in the agriculture and farming sector [34, 37, 41, 45], and in the development of various weed detection models lately. Traditional ML approaches have been extensively developed to perform the task of weed detection due to ease in both understanding ML algorithms and the availability

Table 2. 4 ML techniques used to develop weed detection models

Dataset used as Input	ML algorithms used	Results	Reference
Bi-spectral images	Unsupervised Clustering	Overall weed detection accuracy of 75%	[72]
RGB images	Linear Regression	Accuracy achieved in 6 different fields: 75-83%	[75]
RGB images	Multiple Regression	Classification accuracy of 92.8% between infested and non-infested regions	[76]
Bi-spectral images	Supervised and Unsupervised Clustering	Classification of crops and weeds categories	[74]
Pattern recognition features used	KNN classifier	Overall accuracy of 83.1% with a kappa coefficient of 0.775	[78]
Features extracted from Gabor and FFT filters	SVM	Classification of narrow and broad-leaved weeds, Overall accuracy: 100%	[80]

Three types of shape features used for training	SVM and ANN	Overall classification accuracy of crops and weeds, ANN: 92.92% and SVM 92.5%	[81]
Optimal features extracted from RGB images	SVM	Overall classification accuracy score of 97%	[82]
Hyper-spectral images	DTs developed with boosting	Global accuracy scores of 95% were achieved using spectral and shape features	[84]
Hyper-spectral images	CART and DT	Classification accuracy of 96% for early growth stages of weeds in corn crops	[85]
Remote sensing imagery	RF	Overall pixel-based classification of 98.2%	[87]
Image pattern recognition features from UAV	RF with 5 DTs	Overall classification accuracy of 82% and Kappa coefficient of 0.73	[86]
RGB images	ANN	Overall classification accuracy score of 98.1%	[136]
Grayscale and RGB images	RF	Overall classification accuracy of 94%	[137]
RGB images	SVM, XGBoost, LR	SVM classifier obtained a F1-score of 99.29%	[135]

of features that could be used to train these models. DL models usually require large datasets to efficiently train and perform tasks such as prediction and classification.

Traditional ML algorithms were limited in their ability to extract useful insights from raw natural raw data. DL algorithms such as CNNs, LSTM, GANs, FCNs, and RNNs have been used to develop both supervised and unsupervised DL models. CNN-based supervised DL approaches have gained much popularity in the recent past due to their feature generation and self-learning capabilities. CNN consists of multiple layers: convolutional filters, pooling layers, and fully connected layers. Convolutional filters in CNN help in extracting important generic feature maps from raw natural data. The combination of generic feature maps from original data helps in training a DL model to detect several motifs from raw data that further help in performing complex tasks such as pattern recognition and object recognition. Pooling layers perform the task of dimensionality reduction while preserving the feature maps generated by convolutional filters in the DL network. The fully connected layers consist of adjustable weight parameters that represent the magnitude of certain feature outputs generated from convolutional filters that in turn play a major role in deciding the final prediction during training of a supervised DL model. The concept of Backpropagation performs the task of training in a DL algorithm. During the training phase, an objective function is computed that measures the deviation in between the predicted and desired output. Backpropagation performs the task of regularly computing the objective function and updating adjustable weight parameters in fully connected layers such that the predicted output matches the desired output (Lecun et al., 2015) [88].

The ILSVRC was a worldwide contest conducted between 2010-2017 to measure the progress of Computer Vision field to solve the complex tasks of object detection. The ILSVRC dataset consisted of 1.2 million annotated images of different objects (such as Person, airplanes, cars, etc.) across 1000 categories and the goal was to develop a ML architecture that could perform prediction task with higher accuracy scores. DL techniques gained much importance after the deep neural network Alexnet consisting of CNNs won the ILSVRC contest also popularly known as the ImageNet competition in 2012 [111] (https://image-net.org/index.php, access on November 1, 2022). Optimal neural networks have been developed continuously since than to achieve better accuracy scores during prediction on ImageNet dataset. Popular CNN architectures such as VGG [91], ResNet [129], Inception [130], ResNeXt, GoogleNet [130] were developed in the following years to solve the ImageNet challenge and these architectures have been used for tasks such as object recognition, image classification, and automatic object clustering. Also, these architectures have been largely used as backbone (also called encoder or feature extractor) in popular classification algorithms of DL namely OD and SS.

DL approaches have been used to develop weed detection models that could be used to perform the task of SSWM and use those applications for real-time sensing and spraying. Reviews on DL approaches that have been used to develop weed detection models [89], [90] are already published. Several CNN architectures such as GoogleNet

[131], MobileNet [132], ShuffleNet [133], VGGNet [91], ResNet [129], Inception [130], DenseNet [92], and ExceptionNet have been used to develop weed detection models. These CNN architectures differ in the arrangement and number of convolutional filters, pooling layers, and fully connected layers that have been used to create that specific architecture. These architectural arrangements are all unique and have contributed to solving diverse kinds of classification and prediction tasks. Deep CNN architectures (consisting of a higher number of convolutional and pooling layers) such as Inception and ResNet have been successfully used to solve complex classification tasks and achieve higher accuracy scores and shallow CNN architectures (consisting of lower convolutional and pooling layers) such as MobileNet [132], ExceptionNet [134] architectures have been used on edge computing devices to perform real-time applications [92].

2.4.3.1 Transfer Learning and Data Augmentation

A major challenge associated with training DL models is that they usually require large amounts of data during training and an increase in the number of images used during training have helped DL models to achieve higher accuracy scores. A common solution to handle the unavailability of large datasets to train DL models is the usage of TL and data augmentation techniques. Techniques such as TL and Data Augmentation have been popularly used to alleviate the problem during availability of insufficient data to train a DL model.

TL refers to the use of existing knowledge that has already been used to perform one task, usually where large amounts of data is present for training to solve a similar learning task, where there is limited data to perform the training of ML model. Training ML models especially. DL models using this approach have shown amazing results and helped in achieving state-of-art accuracy scores during prediction. Several applications using this approach have already been developed in the recent past [93] and this approach is widely used to develop weed detection models. [105], [107], [108] presents studies in which DL models were developed using TL techniques to overcome the limitation of training data. CNN architectures such as VGG, ResNet, Inception, DenseNet, MobileNet networks that have already been trained on ImageNet [94, 95], Microsoft COCO [96] datasets and are easily available and could be used for developing various detection, classification, and segmentation tasks using TL techniques. TL techniques were used by [30] to develop weed detection model and performance comparisons were made on the models that were developed using MobileNet and Inception architectures already trained on ImageNet datasets. [93] proposes a study in which a weed detection models are developed using a combination of various pre-trained convolutional architectures such as Xception, Inception-ResNet, MobileNet, and DenseNet with traditional ML models such SVMs, and LoR. A section in the review developed by [113] discusses various detection and classification models developed in agriculture domain using TL techniques in detail.

Data Augmentation techniques in ML domain refers to the artificial increase in the amount of training data my adding moderately modified copies of original data or newly created synthetic data from existing data. Various strategies such as adjustment in color intensities, random rotations, flipping, etc. are commonly used to perform data augmentation. DL models usually performs the training on features extracted from the images used for training. Increasing images using these strategies helps to increase the datasets by manifolds and helps in increasing the dataset used for training DL model. These techniques have also proven to escalate the robustness of CNN architectures used for feature extraction process in DL models. [30] performed horizontal flip and crop rotations to increase the images in training dataset. Data augmentation techniques such as pre-processing images by altering brightness, random rotations, flips and color variations were performed on original dataset to increase 920 images in original dataset to 9200 images to be used for training DL models [107].

Recent advancements in the field of PA and need for sustainable practices such as SSWM has led to a significant increase in the use of DL techniques for development of weed detection models. A large number of studies conducted in this domain are conducted to develop neural networks solving weed detection problem on independent and unpublished datasets, this approach makes it exceedingly hard to compare and evaluate the performance of these neural networks in general. In order to perform performance comparisons of neural networks developed in this domain, it is crucial to prepare sufficiently large datasets comprising of weeds and crops images that could be used by researchers. An increase in the number of labeled datasets that are currently being developed in this domain would help to develop more robust and accurate weed detection models in near future.

DL generally has 2 classification algorithms based on the results generated during output. The following two sub-sections explain these 2 approaches in detail along with the weed detection applications developed using these techniques recently are discussed along. Table 7 presents a detailed description of the weed detection models developed using the DL approach in recent past.

2.4.3.2 Object-Detection

DL models perform the classification or detection task in which bounding boxes are formed with probability scores around the predictions as a final output belongs to the category of OD in DL. OD models that are quite popular and have been used to train DL-based weed detection models are Fast RCNN [97], Faster RCNN [98], YOLO [99], SSD [100], and MaskRCNN [101]. These models [97, 98, 101] consist of RPNs that help the model predict certain locations in an image where there is a higher probability of finding an object, and then a convolutional network is used to perform the task of object detection and localization whereas in [99, 100] directly a convolutional network is used for both object detection and localization during prediction. The advantages associated with the former approaches discussed are that they can predict even the small objects that are present in an image whereas the later approaches usually miss predicting small objects. The disadvantages associated with former approaches are that due to their complex architecture they have high processing time and take more time to perform predictions as compared to later approaches. This has led to the use of [99, 100] approaches on edge-computing devices to perform real-time applications [105].

(Sivakumar et al., 2020) [30] developed a weed detection model using OD-based FasterRCNN and SSD models and made performance comparisons of these models based on IoU scores and inference speed. The research concluded that the FasterRCNN model could achieve a higher optimized confidence threshold as compared to SSD which depicts that FasterRCNN generalizes better during testing. [107] conducted research and developed a weed detection model using OD-based DL approaches in vegetables. Performance comparisons between YOLO-v3 and FasterRCNN were made, and the results depicted that YOLO-v3 was able to achieve higher accuracy scores and had a significantly shorter inference time as compared to the latter.

2.4.3.3 Semantic Segmentation

DL models that perform the classification or detection tasks in which pixel-wise labeling is performed around the predictions as a final output belongs to the category of SS in DL. SS models that are quite popular and have been extensively used to develop weed detection models using the DL approach are SegNet [102], UNet [103], LinkNet [127], BiSeNet [128], and PSPNet [104]. SS techniques are gaining much importance as they can achieve higher accuracy results to train state-of-art DL models even with less amount of data [103]. These tasks have helped elucidate complex tasks of scene understanding. There has been a surge in the research related to performing tasks such as scene understanding due to rise in developing real-time applications for self-driving cars [149, 150, 151], augmented reality [152, 153], etc. in recent past as performing pixel-wise classifications is an important task in these techniques. These models consist of encoder-decoder architecture [102, 103] that consists of convolutional, pooling, and up-convolutional layers. Encoder architecture performs the feature extraction process using the convolutional filters and the pooling layers lead to the condensed representation of the original image whereas in the decoder architecture the pooling indices in SegNet and both feature maps and pooling indices in UNet are concatenated along the up-convolutional layers that help in converting the condensed representation back to the original input size of the image during prediction. Different CNN architectures such as DenseNet, VGG, ResNet, MobileNet, Inception, etc. using TL techniques can be used for the generation of feature maps in the encoder part of these models. These architectures usually have high processing time due to their complex architecture. The research conducted by (Asad & Bais, 2020) [108] made performance comparisons between SS models of UNet and SegNet that were used to perform tasks of weed detection in canola fields. ResNet-50 and VGG16 architectures were used in place of the encoder section to perform the task of feature extraction for both the models. It was concluded that the SegNet SS model in which ResNet50 architecture was used in encoder section gave the best results with a mean IoU score of 0.8288. [109] conducted research and conclude that NIR information helped improve the robustness in segmentation against different lighting conditions and obtained a mean IoU score of 88.91%. InceptionV3 was used as an encoder to develop a supervised SS-based UNet model with data augmentation techniques to achieve accuracy scores of 90% for broad-leaves weed classification. Image enhancement methods such as the inclusion of NIR information used to train a supervised SS-based UNet model for weed detection purposes significantly improve the

Table 2. 5 DL techniques used to develop weed detection models

Dataset used	DL model	Results	Reference
400 RGB images from UAV	5 DL models: MobileNetV2, ResNet50 and 3 custom CNNs	5-layer CNN achieves a detection accuracy of 95%	[92]
RGB images under natural light condition	DenseNet model combined with SVM	The proposed model achieves a F1-score of 99.29%	[93]
RGB images captured at an altitude of 0.6m	Object detection models: Faster R-CNN, Yolo-v3, and CenterNet	Overall Yolo-v3 achieved the highest accuracy and computational accuracy amongst others	[107]
RGB imagery from UAV	Semantic Segmentation models: UNet and SegNet	SegNet model performed best with an IoU score of 0.8288	[108]
RGB imagery from UAV	UNet model with InceptionV3 as feature extractor	Accuracy of weed detection > 90%	[109]
RGB imagery and NIR information	Encoder-decoder based Deep CNN	Mean IoU score for pixel-wise segmentation score of 88.91%	[110]

Table 2. 6 Publicly available weed detection datasets

Dataset name	Dataset information	Annotation type	Total no. of images	URL
Plant Seedling [140]	RGB	Images per class category	407	https://vision.eng.au.dk/plant-seedlings-dataset/
Sugar Beets 2016 [141]	Available in multiple formats	Images per class category	>10000	https://www.ipb.uni-bonn.de/datasets_IJRR2017/annotations/
Carrot-Weed [142]	RGB	Pixel level	39	https://github.com/lameski/rgbweeddetection
WeedNet [143]	Multispectral	Images per class category	465	https://github.com/inkyusa/weedNet
CWFI dataset [144]	Multispectral	Pixel level	60	https://github.com/cwfid/dataset
Open Plant Phenotype database [145]	RGB	Bounding box	7590	https://gitlab.au.dk/AUENG-Vision/OPPD/-/tree/master/
Early crop weed [146]	RGB	Images per class category	508	https://github.com/AUAgroup/early-crop-weed
Leaf counting [147]	RGB	Images per class category	9372	https://vision.eng.au.dk/leaf-counting-dataset/

segmentation accuracy and the model were able to achieve a mean IoU score of 88.21% [110]. Publicly available weed detection datasets are mentioned in Table 2.6 . these datasets are available online and can be used to increase datasets while developing weed detection or to check performance comparison on developed DL models on these datasets.

2.5 Review of real-time weed detection applications: Challenges and opportunities

Overall increase in food production and methods to prevent crop losses is a matter of serious concern due to increasing world population. Major crop yield loss, decrease in crop quality, evolution of herbicide-resistant weeds, and the ill-effects associated with increased use of herbicides have motivated scientists and engineers to perform continuous research to handle the menace that arise due to weeds in agriculture fields. SSWM applications could help overcome these challenges and weed detection models developed using various techniques are discussed in previous sections of this study. The final goal of developing these weed detection models is to use them for performing real-time on-field weeding operations. Numerous real-time SSWM applications have already been developed that make use of both chemical-based weeding strategies such as herbicide spraying using various ground-based [114] and aerial based equipment's [115], and non-chemical weeding strategies such as using high electric discharge [68], mechanical actuators to remove weeds [113] etc. in recent past. This section discusses various approaches that have been used in various studies performed in relation to this domain and their associated advantages and disadvantages are described.

Embedded systems have played a vital role in performing real-time SSWM applications in agriculture fields. Embedded systems could be referred to as minicomputers consisting of both hardware and a software that could perform a specific function or sets of functions. Some common examples of embedded systems are Raspberry Pi, Orange Pi, and Nvidia Jetson. These embedded systems could be employed on various agricultural equipment's such as tractors [113], UGVs including robotic models [116], UAVs [115], and could be trained to perform weeding operations. To perform weeding operations, the embedded systems are responsible for pre-processing the sensory information collected from fields, use this information to generate weed maps, and initiate actuators such as nozzle sprays, releasing electrical discharge, or mechanical actuators attached with agricultural equipment's to eliminate weeds from fields.

DL methods have shown prominent results in tasks such as object detection and classification, albeit rather their complex black-box structure, higher computational power requirements for processing, power supply issues, and higher memory storage requirements makes implementation of these DL models on embedded systems hard as they possess only limited processing and storage capabilities. One of the potential solutions to meet the high computational requirements and to perform millions of computations during training and inference phase of DL models is to use Cloud technologies. According to [120], *Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.* This definition includes the five essential characteristics associated with cloud-computing namely: on-demand self-service, broad network access, resource pooling – location independence, rapid elasticity, and measurable service. Availability of on-demand cloud computing platforms such as AWS, Azure, GCP, Alibaba Cloud etc. from companies such as Amazon, Microsoft, Google, and Alibaba have popularized use of cloud-based technologies for PA applications [118, 119].

Developing a real-time weed detection model involves five steps as discussed in Section 4 of this study. After data collection using various RS techniques the following two operations which involves high computational requirements such as pre-processing of the data and generation of a weed map using weed detection model have been performed on cloud platforms in recent past. In order to perform real-time weeding applications, data sharing across cloud platforms and embedded systems is performed via IoT. The sharing of data includes sending and receiving of instructions on where to perform weeding operations after processing of information at cloud locations. In order to overcome the high space and time complexity of computer vision algorithms that perform weed detection tasks [121], a cloud-based architecture was used to leverage high computational resources and additional aggregated knowledge to generate weed maps for real-time SSWM applications. [122] developed a cloud and AI based application that uses RS data such as RGB imagery collected from airplanes, UAVs, and satellites. The application rapidly processes the data to perform various applications such as measuring plant height and canopy, detecting, and locating plant gaps, and developing plant heat maps with higher accuracy scores. Google cloud platform were used to perform real-time detection of parthenium weed plants in pulse crops by [123]. In this research, a LinkNet model (SS based DL model) was used with ResNet34 as feature extractor to develop a weed detection model that achieved a mean accuracy score of 0.598 in 0.217s. In the study proposed by [124], a weed detection model was developed in which an IoT device collects images from the fields and transmits them to a cloud server where YOLO-v5 DL model performs the task of weed detection on the images. The research conducted by [125] proposed a model that integrates technologies such as IoT, data analytics, and cloud computing to detect nutrient deficiency, weeds, and disease detection in chili crops. There are a few issues associated with the use of cloud-based technologies [126] to perform real-time applications in weed detection domain and these are as follows:

1. *Latency*: The weed detection models developed using cloud technologies requires sharing of both the input data generated locally in the fields and the weed maps generated after pre-processing the information in between cloud and sensors in the fields to perform real-time SSWM applications. Strong network connectivity plays a vital role in using these technologies as there needs to be a strong and robust communication between sensors and cloud

locations. There needs to be a constant access of internet to maintain proper connectivity between both platforms. Also, exploiting resources available at cloud may face additional queuing. These issues could lead to network latency and delay decision making for real-time weeding applications.

2. *Privacy*: This is a major issue of concern as there might be risks associated with data leakage or compromising of personal data from the cloud locations. There exist other issues associated with the use of cloud technologies such as misuse of sensitive information already uploaded to the cloud by the cloud companies. User needs to be wary about the privacy concerns of the information shared with cloud.
3. *Scalability*: Data generated from sensors in the fields needs to be shared with cloud regularly and sharing large loads of data in a short time could be a challenge. Uploading higher resolution imagery or video streaming with cloud would excessive bandwidth consumptions and could lead to scalability issues in sharing data. Also, scalability issues would increase if multiple cameras shared data concurrently with the cloud.

Edge computing has served as a practical solution to overcome the centralization issues such as Latency, Scalability, and Privacy associated with using cloud computing technologies. Recent advancements in developing powerful GPU-accelerated parallel processing embedded devices such as Nvidia Jetson, Google Coral, and Intel NCS etc., have led to their increase in running various DL related applications such as image classification, object detection, segmentation, and speech recognition. According to [138], *Edge computing refers to the enabling technologies allowing computation to be performed at the edge of the network, on downstream data on behalf of cloud services and upstream data on behalf of IoT services.* Also, the development of various Software Developments Kits (SDK) such as Tensor RT by Nvidia and Tensorflow Lite by Google have helped in developing optimized DL models that delivers low latency and high throughput for inference applications. SDKs like TensorRT perform several important transformations and optimizations to the neural network graph, including constant folding, pruning unnecessary graph nodes, layer fusion, and more (https://developer.nvidia.com/tensorrt). These optimization strategies help in decreasing the computational and memory requirements of DL models while preserving their performance capabilities. TensorRT is a SDK specifically designed to optimize DL models and run them on Nvidia Jetson boards whereas optimization performed using Tensorflow Lite could be used on various other GPUs. [148] performed a study in which performance comparisons of the Mean IoU score and inference time were made in between two SS based weed detection models developed with and without TensorRT optimizations on Nvidia Jetson Nano. In performing comparisons between the developed models, decrease in mean IoU score was observed by a factor of 14.8, but inference time accelerated by a factor of 14.7% in the model developed using TensorRT optimizations. These results show that there is a trade-off between overall accuracy and inference time, performing optimizations on these DL models using various SDKs helps in accelerating the inference time while performing predictions using the developed model but this leads to a decrease in performing accurate predictions by some extent which could be referred to as a disadvantage of performing optimizations on DL models. Overall, the use of embedded GPUs to perform real-time applications is still in its initial stages. Several studies have shown results for improved performance using optimized DL models on the energy efficient embedded GPUs. Further study in this domain could focus on developing state-of-art models using these techniques to perform real-time on-board applications for various smart agriculture applications like weed detection, livestock management, and plant disease detections etc. in agriculture domain with better prediction performance.

2.6 Summary and Conclusions

Weed management is one of the most important crop production practices. The global increase in herbicides use to control weeds has led to various issues such as decrease in crop quality, economic losses due to excessive off-target herbicide movement, and evolution of herbicide resistant weeds. Site-Specific Weed Management strategies refer to the use of right and precise amounts of herbicide only at locations where weed infestation are detected. There have been several studies in relation to development of these SSWM applications in agriculture domain, this review article discusses various image processing, machine vision, machine learning, and deep learning-based computer vision techniques in detail that have been used to develop weed detection models in recent years. Also, a detailed description on various ML algorithms and DL models used to develop weed detection models and their associated advantages and disadvantages are also discussed.

The use of CNN in DL methods have revolutionized the tasks of image detection, classification, and segmentation tasks as these models have the advantages of feature extraction and self-learning capabilities over other conventional ML models where an optimal set of features extracted from the data are used for training. Overall, development of weed detection models using DL techniques have helped in generation of better weed maps during prediction and achieve higher accuracy scores as compared to weed detection models developed using traditional machine learning, machine vison and image-processing techniques. Several weed detection models developed using various ML and DL techniques along with their corresponding results are presented in this study. Also, a list of publicly available datasets that have been developed in recent past that could be used to check the performance of new DL models or increase the amount of dataset being used to train DL models are also mentioned along with their respective URLs in this study. In the final section of this review, the popular approaches that have been used to perform real-time SSWM applications are discussed. An overview of various techniques and their associated advantages and disadvantages are discussed in the section.

We conclude that DL models could be further explored by tuning and tweaking hyperparameters used for training these models, as this could lead to generation of weed detection models that could perform the detection tasks more efficiently and robustly. It is known that training of DL model requires large dataset for training, there exists an absence of a datasets like ImageNet, COCO, CamVid etc., in weed detection domain which could be used by researchers to test performance of various DL models and architectures. Several datasets have already been prepared and are available publicly, but they include less images as compared to the datasets mentioned earlier. The preparation of a common generic dataset would help to check and compare performance comparisons between different DL models and architectures used by researchers while preparing these models. It has been observed that there exists a class imbalance issue amongst various datasets that have been used to develop a weed detection model. Future work could involve developing robust weed detection models that could address these class imbalance problems.

There is a wide scope of research that could be conducted in performing optimizations of the DL models using various SDKs such as Tensorflow Lite and TensorRT techniques so that these trained models could be used on edge devices for real-time SSWM applications. It has been proven that performing predictions on optimized DL models have an accelerated inference time, but these models have a trade-off with performing accurate predictions. Further research could be conducted in developing models that could achieve the tasks of optimizing DL models while preserving their performance capabilities in performing accurate predictions on edge devices such as Nvidia Jetson, Google Coral etc. to achieve the tasks of SSWM applications in agriculture fields.

2.7 References:

1. Savelli, C. J., Bradshaw, A., Ben Embarek, P., & Mateus, C. (2019). The FAO/WHO international food safety authorities' network in review, 2004–2018: learning from the past and looking to the future. Foodborne pathogens and disease, 16(7), 480-488. https://doi.org/10.1089/fpd.2018.2582

2. Rosle, R., Che'Ya, N. N., Ang, Y., Rahmat, F., Wayayok, A., Berahim, Z., Fazlil Ilahi, W. F., Ismail, M. R., & Omar, M. H. (2021). Weed Detection in Rice Fields Using Remote Sensing Technique: A Review. Applied Sciences, 11(22), 10701. https://doi.org/10.3390/app112210701

3. Wang, D., Cao, W., Zhang, F., Li, Z., Xu, S., & Wu, X. (2022). A review of deep learning in multiscale agricultural sensing. Remote Sensing, 14(3), 559. https://doi.org/10.3390/rs14030559

4. Liakos, K., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine Learning in Agriculture: A Review. Sensors, 18(8), 2674. https://doi.org/10.3390/s18082674

5. Benos, L., Tagarakis, A. C., Dolias, G., Berruto, R., Kateris, D., & Bochtis, D. (2021). Machine Learning in Agriculture: A Comprehensive Updated Review. Sensors, 21(11), 3758. https://doi.org/10.3390/s21113758

6. Hasan, A. M., Sohel, F., Diepeveen, D., Laga, H., & Jones, M. G. (2021). A survey of deep learning techniques for weed detection from images.

Computers and Electronics in Agriculture, 184, 106067. https://doi.org/10.1016/j.compag.2021.106067

7. Sagi, O., & Rokach, L. (2018). Ensemble learning: A survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(4), e1249. https://doi.org/10.1002/widm.1249

8. World Bank. (2007). World development report 2008: Agriculture for development. The World Bank. https://doi.org/10.1596/978-0-8213-6807-7

9. Pimentel, D., Zuniga, R., & Morrison, D. (2005). Update on the environmental and economic costs associated with alien-invasive species in the United States. Ecological economics, 52(3), 273-288. https://doi.org/10.1016/j.ecolecon.2004.10.002

10. Oerke, E. C. (2006). Crop losses to pests. The Journal of Agricultural Science, 144(1), 31-43. https://doi.org/10.1017/S0021859605005708

11. Lamb, D. W., & Brown, R. B. (2001). Pa—precision agriculture: Remote-sensing and mapping of weeds in crops. Journal of Agricultural Engineering Research, 78(2), 117-125. https://doi.org/10.1006/jaer.2000.0630

12. Thorp, K. R., & Tian, L. F. (2004). A review on remote sensing of weeds in agriculture. Precision Agriculture, 5(5), 477-508. https://doi.org/10.1007/s11119-004-5321-1

13. Christensen, S., Søgaard, H. T., Kudsk, P., Nørremark, M., Lund, I., Nadimi, E. S., & Jørgensen, R. (2009). Site-specific weed control technologies. Weed Research, 49(3), 233-241. https://doi.org/10.1111/j.1365-3180.2009.00696.x

14. Ghersa, C. M. (2013). "Agroecological basis for managing biotic constraints," in Sustainable Food Production, eds P. Christou, R. Savin, B. A. Costa-Pierce, I. Misztal, and C. B. A.Whitelaw (New York, NY: Springer), 18–30. https://doi-org.libproxy.unl.edu/10.1007/978-1-4614-5797-8_196

15. Lowenberg-DeBoer, J. (2015). The Precision agriculture revolution. Foreign Aff., 94, 105. https://heinonline.org/HOL/P?h=hein.journals/fora94&i=667

16. Chakravarti, A. K. (1973). Green revolution in India. Annals of the Association of American Geographers, 63(3), 319-330. https://doi.org/10.1111/j.1467-8306.1973.tb00929.x

17. Shi, Y., Thomasson, J. A., Murray, S. C., Pugh, N. A., Rooney, W. L., Shafian, S., ... & Yang, C. (2016). Unmanned aerial vehicles for high-throughput phenotyping and agronomic research. PloS one, 11(7), e0159781. https://doi.org/10.1371/journal.pone.0159781

18. Tsouros, D. C., Bibi, S., & Sarigiannidis, P. G. (2019). A review on UAV-based applications for precision agriculture. Information, 10(11), 349. https://doi.org/10.3390/info10110349

19. Chauhan, B. S. (2020). Grand challenges in weed management. Frontiers in Agronomy, 1, 3. (https://doi.org/10.3389/fagro.2019.00003)

20. Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., & Garcia-Rodriguez, J. (2017). A review on deep learning techniques applied to semantic segmentation. arXiv preprint arXiv:1704.06857. (https://doi.org/10.48550/arXiv.1704.06857)

21. Gomes, J. F. S., & Leta, F. R. (2012). Applications of computer vision techniques in the agriculture and food industry: a review. European Food Research and Technology, 235(6), 989-1000. https://doi.org/10.1007/s00217-012-1844-2

22. Tripathi, M. K., & Maktedar, D. D. (2020). A role of computer vision in fruits and vegetables among various horticulture products of agriculture fields: A survey. Information Processing in Agriculture, 7(2), 183-203. https://doi.org/10.1016/j.inpa.2019.07.003

23. Kuricheti, G., & Supriya, P. (2019, April). Computer vision based turmeric leaf disease detection and classification: a step to smart agriculture. In 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI) (pp. 545-549). IEEE. https://doi.org/10.1109/ICOEI.2019.8862706

24. Donmez, C., Villi, O., Berberoglu, S., & Cilek, A. (2021). Computer vision-based citrus tree detection in a cultivated environment using UAV imagery. Computers and Electronics in Agriculture, 187, 106273. https://doi.org/10.1016/j.compag.2021.106273

25. Bàrberi, P. A. O. L. O. (2002). Weed management in organic agriculture: are we addressing the right issues? Weed research, 42(3), 177-193. https://doi.org/10.1046/j.1365-3180.2002.00277.x

26. Heap, I. (2014). Global perspective of herbicide-resistant weeds. Pest management science, 70(9), 1306-1315. https://doi.org/10.1002/ps.3696

27. Zhang, N., Wang, M., & Wang, N. (2002). Precision agriculture—a worldwide overview. Computers and electronics in agriculture, 36(2-3), 113-132. https://doi.org/10.1016/S0168-1699(02)00096-0

28. Kok, Z. H., Shariff, A. R. M., Alfatni, M. S. M., & Khairunniza-Bejo, S. (2021). Support vector machine in precision agriculture: a review. Computers and Electronics in Agriculture, 191, 106546. https://doi.org/10.1016/j.compag.2021.106546

29. Tarsha Kurdi, F., Amakhchan, W., & Gharineiat, Z. (2021). Random forest machine learning technique for automatic vegetation detection and modeling in LiDAR data. International Journal of Environmental Sciences and Natural Resources, 28(2). http://eprints.usq.edu.au/id/eprint/42809

30. Veeranampalayam Sivakumar, A. N., Li, J., Scott, S., Psota, E., J Jhala, A., Luck, J. D., & Shi, Y. (2020). Comparison of object detection and patch-based classification deep learning models on mid-to late-season weed detection in UAV imagery. Remote Sensing, 12(13), 2136. https://doi.org/10.3390/rs12132136

31. Haghverdi, A., Washington-Allen, R. A., & Leib, B. G. (2018). Prediction of cotton lint yield from phenology of crop indices using artificial neural networks. Computers and Electronics in Agriculture, 152, 186-197. https://doi.org/10.1016/j.compag.2018.07.021

32. Bhattarai, U., & Karkee, M. (2022). A weakly-supervised approach for flower/fruit counting in apple orchards. Computers in Industry, 138, 103635. https://doi.org/10.1016/j.compind.2022.103635

33. Aghighi, H., Azadbakht, M., Ashourloo, D., Shahrabi, H. S., & Radiom, S. (2018). Machine learning regression techniques for the silage maize yield prediction using time-series images of Landsat 8 OLI. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11(12), 4563-4577. https://doi.org/10.1109/JSTARS.2018.2823361

34. Nevavuori, P., Narra, N., & Lipping, T. (2019). Crop yield prediction with deep convolutional neural networks. Computers and electronics in agriculture, 163, 104859. https://doi.org/10.1016/j.compag.2019.104859

35. Akbar, A., Kuanar, A., Patnaik, J., Mishra, A., & Nayak, S. (2018). Application of artificial neural network modeling for optimization and prediction of essential oil yield in turmeric (Curcuma longa L.). Computers and Electronics in Agriculture, 148, 160-178. https://doi.org/10.1016/j.compag.2018.03.002

36. Pourmohammadali, B., Hosseinifard, S. J., Salehi, M. H., Shirani, H., & Boroujeni, I. E. (2019). Effects of soil properties, water quality, and management practices on pistachio yield in Rafsanjan region, southeast of Iran. Agricultural water management, 213, 894-902. https://doi.org/10.1016/j.agwat.2018.12.005

37. Chen, Y., Lee, W. S., Gan, H., Peres, N., Fraisse, C., Zhang, Y., & He, Y. (2019). Strawberry yield prediction is based on a deep neural network using high-resolution aerial orthoimages. Remote Sensing, 11(13), 1584. https://doi.org/10.3390/rs11131584

38. Wenzhi, Z. E. N. G., Chi, X. U., Gang, Z. H. A. O., Jingwei, W. U., & Huang, J. (2018). Estimation of sunflower seed yield using partial least squares regression and artificial neural network models. Pedosphere, 28(5), 764-774. https://doi.org/10.1016/S1002-0160(17)60336-9

39. Gutiérrez, S., Wendel, A., & Underwood, J. (2019). Spectral filter design based on in-field hyperspectral imaging and machine learning for mango ripeness estimation. Computers and Electronics in Agriculture, 164, 104890. https://doi.org/10.1016/j.compag.2019.104890

40. Gopal, P. M., & Bhargavi, R. (2019). A novel approach for efficient crop yield prediction. Computers and Electronics in Agriculture, 165, 104968. https://doi.org/10.1016/j.compag.2019.104968

41. Sun, J., Di, L., Sun, Z., Shen, Y., & Lai, Z. (2019). County-level soybean yield prediction using deep CNN-LSTM model. Sensors, 19(20), 4363. https://doi.org/10.3390/s19204363

42. Son, N. T., Chen, C. F., Chen, C. R., Guo, H. Y., Cheng, Y. S., Chen, S. L., ... & Chen, S. H. (2020). Machine learning approaches for rice crop yield predictions using time-series satellite data in Taiwan. International Journal of

Remote Sensing, 41(20), 7868-7888.
https://doi.org/10.1080/01431161.2020.1766148

43. Rahman, M. M., & Robson, A. (2020). Integrating Landsat-8 and sentinel-2 time series data for yield prediction of sugarcane crops at the block level. Remote Sensing, 12(8), 1313. https://doi.org/10.3390/rs12081313

44. Khosla, E., Dharavath, R., & Priya, R. (2020). Crop yield prediction using aggregated rainfall-based modular artificial neural networks and support vector regression. Environment, Development, and Sustainability, 22(6), 5687-5708. https://doi.org/10.1007/s10668-019-00445-x

45. Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. Computers and electronics in agriculture, 145, 311-318. https://doi.org/10.1016/j.compag.2018.01.009\

46. Abdulridha, J., Ampatzidis, Y., Ehsani, R., & de Castro, A. I. (2018). Evaluating the performance of spectral features and multivariate analysis tools to detect laurel wilt disease and nutritional deficiency in avocado. Computers and electronics in agriculture, 155, 203-211. https://doi.org/10.1016/j.compag.2018.10.016

47. Yang, B., Gao, Y., Yan, Q., Qi, L., Zhu, Y., & Wang, B. (2020). Estimation method of soluble solid content in peach based on deep features of hyperspectral imagery. Sensors, 20(18), 5021. https://doi.org/10.3390/s20185021

48. de Medeiros, A. D., Pinheiro, D. T., Xavier, W. A., da Silva, L. J., & dos Santos Dias, D. C. F. (2020). Quality classification of Jatropha curcas seeds using radiographic images and machine learning. Industrial Crops and Products, 146, 112162. https://doi.org/10.1016/j.indcrop.2020.112162

49. Gressel, J. (1992). Addressing Real Weed Science Needs with Innovations. Weed Technology, 6(3), 509-525. https://doi.org/10.1017/S0890037X00035739

50. Beckie, H. J., & Tardif, F. J. (2012). Herbicide cross resistance in weeds. Crop Protection, 35, 15-28. https://doi.org/10.1016/j.cropro.2011.12.018

51. Buhler, D. D. (1996). Development of alternative weed management strategies. Journal of Production Agriculture, 9(4), 501-505. https://doi.org/10.2134/jpa1996.0501

52. Heap, I. (2014). Global perspective of herbicide-resistant weeds. Pest management science, 70(9), 1306-1315. https://doi.org/10.1002/ps.3696

53. Wiles, L. J., King, R. P., Schweizer, E. E., Lybecker, D. W., & Swinton, S. M. (1996). GWM: general weed management model. Agricultural Systems, 50(4), 355-376. https://doi.org/10.1016/0308-521X(95)00016-X

54. Wang, A., Zhang, W., & Wei, X. (2019). A review on weed detection using ground-based machine vision and image processing techniques. Computers and electronics in agriculture, 158, 226-240. https://doi.org/10.1016/j.compag.2019.02.005

55. Slaughter, D. C., Giles, D. K., & Downey, D. (2008). Autonomous robotic weed control systems: A review. Computers and electronics in agriculture, 61(1), 63-78. https://doi.org/10.1016/j.compag.2007.05.008

56. Wang, A., Zhang, W., & Wei, X. (2019). A review on weed detection using ground-based machine vision and image processing techniques. Computers and electronics in agriculture, 158, 226-240. https://doi.org/10.1016/j.compag.2019.02.005

57. López-Granados, F. (2011). Weed detection for site-specific weed management: mapping and real-time approaches. Weed Research, 51(1), 1-11. https://doi.org/10.1111/j.1365-3180.2010.00829.x

58. Vrindts, E., De Baerdemaeker, J., & Ramon, H. (2002). Weed detection using canopy reflection. Precision agriculture, 3(1), 63-80. https://doi.org/10.1023/A:1013326304427

59. Sishodia, R. P., Ray, R. L., & Singh, S. K. (2020). Applications of remote sensing in precision agriculture: A review. Remote Sensing, 12(19), 3136. https://doi.org/10.3390/rs12193136

60. Lamb, D. W., & Brown, R. B. (2001). Pa—precision agriculture: Remote-sensing and mapping of weeds in crops. Journal of Agricultural Engineering Research, 78(2), 117-125. https://doi.org/10.1006/jaer.2000.0630

61. Forero, M. G., Herrera-Rivera, S., Ávila-Navarro, J., Franco, C. A., Rasmussen, J., & Nielsen, J. (2018, November). Color classification methods for perennial weed detection in cereal crops. In Iberoamerican Congress on

Pattern Recognition (pp. 117-123). Springer, Cham. https://doi.org/10.1007/978-3-030-13469-3_14

62. Feyaerts, F., & Van Gool, L. (2001). Multi-spectral vision system for weed detection. Pattern Recognition Letters, 22(6-7), 667-674. https://doi.org/10.1016/S0167-8655(01)00006-X

63. Vrindts, E., De Baerdemaeker, J., & Ramon, H. (2002). Weed detection using canopy reflection. Precision agriculture, 3(1), 63-80. https://doi.org/10.1023/A:1013326304427

64. Cho, S., Lee, D. S., & Jeong, J. Y. (2002). AE—automation and emerging technologies: Weed–plant discrimination by machine vision and artificial neural network. Biosystems Engineering, 83(3), 275-280. https://doi.org/10.1006/bioe.2002.0117

65. Alchanatis, V., Ridel, L., Hetzroni, A., & Yaroslavsky, L. (2005). Weed detection in multi-spectral images of cotton fields. Computers and Electronics in Agriculture, 47(3), 243-260. https://doi.org/10.1016/j.compag.2004.11.019

66. Piron, A., Leemans, V., Lebeau, F., & Destain, M. F. (2009). Improving in-row weed detection in multispectral stereoscopic images. Computers and electronics in agriculture, 69(1), 73-79. https://doi.org/10.1016/j.compag.2009.07.001

67. Okamoto, H., Murata, T., Kataoka, T., & HATA, S. I. (2007). Plant classification for weed detection using hyperspectral imaging with wavelet

analysis. Weed Biology and Management, 7(1), 31-37. https://doi.org/10.1111/j.1445-6664.2006.00234.x

68. Blasco, J., Aleixos, N., Roger, J. M., Rabatel, G., & Moltó, E. (2002). AE—Automation and emerging technologies: Robotic weed control using machine vision. Biosystems Engineering, 83(2), 149-157. https://doi.org/10.1006/bioe.2002.0109

69. Bak, T., & Jakobsen, H. (2004). Agricultural robotic platform with four-wheel steering for weed detection. Biosystems Engineering, 87(2), 125-136. https://doi.org/10.1016/j.biosystemseng.2003.10.009

70. Thorp, K. R., & Tian, L. F. (2004). A review on remote sensing of weeds in agriculture. Precision Agriculture, 5(5), 477-508. https://doi.org/10.1007/s11119-004-5321-1

71. Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255-260. https://doi.org/10.1126/science.aaa8415

72. Sukumar, P., & Ravi, D. S. (2016). Weed detection using image processing by clustering analysis. International Journal of Emerging Technologies in Engineering Research (IJETER), 4(5), 14-18.

73. Zhang, X., Li, X., Zhang, B., Zhou, J., Tian, G., Xiong, Y., & Gu, B. (2018). Automated robust crop-row detection in maize fields based on position clustering algorithm and shortest path method. Computers and electronics in agriculture, 154, 165-175. https://doi.org/10.1016/j.compag.2018.09.014

74. Weis, M., & Gerhards, R. (2008). Detection of weeds using image processing and clustering. Bornimer Agrartechnische Berichte, 69(138), e144

75. Jensen, S. M., Akhter, M. J., Azim, S., & Rasmussen, J. (2021). The Predictive Power of Regression Models to Determine Grass Weed Infestations in Cereals Based on Drone Imagery—Statistical and Practical Aspects. Agronomy, 11(11), 2277. https://doi.org/10.3390/agronomy11112277

76. Andújar, D., Weis, M., & Gerhards, R. (2012). An ultrasonic system for weed detection in cereal crops. Sensors, 12(12), 17343-17357. https://doi.org/10.3390/s121217343

77. Rani, S. V., Kumar, P. S., Priyadharsini, R., Srividya, S. J., & Harshana, S. (2022). An automated weed detection system in smart farming for developing sustainable agriculture. International Journal of Environmental Science and Technology, 19(9), 9083-9094. https://doi.org/10.1007/s13762-021-03606-6

78. Yano, I. H., Mesa, N. F. O., Santiago, W. E., Aguiar, R. H., & Teruel, B. (2017). Weed identification in sugarcane plantations through images taken from remotely piloted aircraft (RPA) and KNN classifier. J. Food Nutr. Sci, 5, 211. https://doi.org/10.11648/j.jfns.20170506.11

79. Singh, A., Ganapathysubramanian, B., Singh, A. K., & Sarkar, S. (2016). Machine learning for high-throughput stress phenotyping in plants. Trends in plant science, 21(2), 110-124. https://doi.org/10.1016/j.tplants.2015.10.015

80. Ishak, A. J., Mustafa, M. M., Tahir, N. M., & Hussain, A. (2008, December). Weed detection system using support vector machine. In 2008 International

Symposium on Information Theory and Its Applications (pp. 1-4). IEEE. https://doi.org/10.1109/ISITA.2008.4895454

81. Bakhshipour, A., & Jafari, A. (2018). Evaluation of support vector machine and artificial neural networks in weed detection using shape features. Computers and Electronics in Agriculture, 145, 153-160. https://doi.org/10.1016/j.compag.2017.12.032

82. Ahmed, F., Al-Mamun, H. A., Bari, A. H., Hossain, E., & Kwan, P. (2012). Classification of crops and weeds from digital images: A support vector machine approach. Crop Protection, 40, 98-104. https://doi.org/10.1016/j.cropro.2012.04.024

83. Tellaeche, A., Pajares, G., Burgos-Artizzu, X. P., & Ribeiro, A. (2011). A computer vision approach for weeds identification through Support Vector Machines. Applied Soft Computing, 11(1), 908-915. https://doi.org/10.1016/j.eswa.2012.03.040

84. Lin, F., Zhang, D., Huang, Y., Wang, X., & Chen, X. (2017). Detection of corn and weed species by the combination of spectral, shape and textural features. Sustainability, 9(8), 1335. https://doi.org/10.3390/su9081335

85. Waheed, T., Bonnell, R. B., Prasher, S. O., & Paulet, E. (2006). Measuring performance in precision agriculture: CART—A decision tree approach. Agricultural water management, 84(1-2), 173-185. https://doi.org/10.1016/j.agwat.2005.12.003

86. Yano, I. H., Alves, J. R., Santiago, W. E., & Mederos, B. J. (2016). Identification of weeds in sugarcane fields through images taken by UAV and Random Forest classifier. IFAC-PapersOnLine, 49(16), 415-420. https://doi.org/10.1016/j.ifacol.2016.10.076

87. Sheffield, K. J., Clements, D., Clune, D. J., Constantine, A., & Dugdale, T. M. (2022). Detection of Aquatic Alligator Weed (Alternanthera philoxeroides) from Aerial Imagery Using Random Forest Classification. Remote Sensing, 14(11), 2674. https://doi.org/10.3390/rs14112674

88. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature, 521(7553), 436-444. https://doi.org/10.1038/nature14539

89. Hasan, A. S. M. M., Sohel, F., Diepeveen, D., Laga, H., & Jones, M. G. K. (2021). A survey of deep learning techniques for weed detection from images. Computers and Electronics in Agriculture, 184(March), 106067. https://doi.org/10.1016/j.compag.2021.106067

90. Wu, Z., Chen, Y., Zhao, B., Kang, X., & Ding, Y. (2021). Review of weed detection methods based on computer vision. Sensors, 21(11), 1–23. https://doi.org/10.3390/s21113647

91. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. https://doi.org/10.48550/arXiv.1409.1556

92. Razfar, N., True, J., Bassiouny, R., Venkatesh, V., & Kashef, R. (2022). Weed detection in soybean crops using custom lightweight deep learning models.

Journal of Agriculture and Food Research, 8, 100308. https://doi.org/10.1016/j.jafr.2022.100308

93. Espejo-Garcia, B., Mylonas, N., Athanasakos, L., Fountas, S., & Vasilakoglou, I. (2020). Towards weeds identification assistance through transfer learning. Computers and Electronics in Agriculture, 171, 105306. https://doi.org/10.1016/j.compag.2020.105306

94. Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255). Ieee. https://doi.org/10.1109/CVPR.2009.5206848

95. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. Communications of the ACM, 60(6), 84-90. https://doi.org/10.1145/3065386

96. Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. (2014, September). Microsoft coco: Common objects in context. In European conference on computer vision (pp. 740-755). Springer, Cham. https://doi.org/10.1007/978-3-030-01424-7_27

97. Girshick, R. (2015). Fast r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 1440-1448). https://doi.org/10.1109/ICCV.2015.169

98. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural

information processing systems, 28. https://doi.org/10.1109/tpami.2016.2577031

99. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788). https://doi.org/10.1109/CVPR.2016.91

100. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016, October). Ssd: Single shot multibox detector. In European conference on computer vision (pp. 21-37). Springer, Cham. https://doi.org/10.1007/978-3-319-46448-0_2

101. He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 2961-2969). https://doi.org/10.48550/arXiv.1703.06870

102. Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE transactions on pattern analysis and machine intelligence, 39(12), 2481-2495. https://doi.org/10.1109/TPAMI.2016.2644615

103. Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham. https://doi.org/10.48550/arXiv.1505.04597

104. Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). Pyramid scene parsing network. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2881-2890). https://doi.org/10.1109/CVPR.2017.660

105. Deepa, R., Tamilselvan, E., Abrar, E. S., & Sampath, S. (2019, April). Comparison of yolo, ssd, faster rcnn for real time tennis ball tracking for action decision networks. In 2019 International conference on advances in computing and communication engineering (ICACCE) (pp. 1-4). IEEE. https://doi.org/10.1109/ICACCE46606.2019.9079965

106. Sandino, J., Gonzalez, F., Mengersen, K., & Gaston, K. J. (2018). UAVs and machine learning revolutionising invasive grass and vegetation surveys in remote arid lands. Sensors, 18(2), 605. https://doi.org/10.3390/s18020605

107. Jin, X., Sun, Y., Che, J., Bagavathiannan, M., Yu, J., & Chen, Y. (2022). A novel deep learning-based method for detection of weeds in vegetables. Pest Management Science, 78(5), 1861-1869. https://doi.org/10.1002/ps.6804

108. Asad, M. H., & Bais, A. (2020). Weed detection in canola fields using maximum likelihood classification and deep convolutional neural network. Information Processing in Agriculture, 7(4), 535–545. https://doi.org/10.1016/j.inpa.2019.12.002

109. Boyina, L., Sandhya, G., Vasavi, S., Koneru, L., & Koushik, V. (2021, June). Weed Detection in Broad Leaves using Invariant U-Net Model. In 2021 International Conference on Communication, Control and Information

Sciences (ICCISc) (Vol. 1, pp. 1-4). IEEE. https://doi.org/10.1109/ICCISc52257.2021.9485001

110. Wang, A., Xu, Y., Wei, X., & Cui, B. (2020). Semantic segmentation of crop and weed using an encoder-decoder network and image enhancement method under uncontrolled outdoor illumination. IEEE Access, 8, 81724-81734. https://doi.org/10.1109/ACCESS.2020.2991354

111. Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., ... & Asari, V. K. (2018). The history began from alexnet: A comprehensive survey on deep learning approaches. arXiv preprint arXiv:1803.01164. https://doi.org/10.48550/arXiv.1803.01164

112. Lemtiri, A., Colinet, G., Alabi, T., Cluzeau, D., Zirbes, L., Haubruge, É., & Francis, F. (2014). Impacts of earthworms on soil components and dynamics. A review. Biotechnologie, Agronomie, Société et Environnement, 18. https://hdl.handle.net/2268/163467

113. Raja, R., Nguyen, T. T., Slaughter, D. C., & Fennimore, S. A. (2020). Real-time robotic weed knife control system for tomato and lettuce based on geometric appearance of plant labels. Biosystems engineering, 194, 152-164. https://doi.org/10.1016/j.biosystemseng.2020.03.022\

114. Bini, D., Pamela, D., & Prince, S. (2020, March). Machine vision and machine learning for intelligent agrobots: A review. In 2020 5th International Conference on Devices, Circuits and Systems (ICDCS) (pp. 12-16). IEEE. https://doi.org/10.1109/ICDCS48716.2020.243538

115. Shankar, R. H., Veeraraghavan, A. K., Sivaraman, K., & Ramachandran, S. S. (2018, December). Application of UAV for pest, weeds and disease detection using open computer vision. In 2018 International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 287-292). IEEE. https://doi.org/10.1109/ICSSIT.2018.8748404

116. Sujaritha, M., Annadurai, S., Satheeshkumar, J., Sharan, S. K., & Mahesh, L. (2017). Weed detecting robot in sugarcane fields using fuzzy real time classifier. Computers and electronics in agriculture, 134, 160-171. https://doi.org/10.1016/j.compag.2017.01.008

117. Rainville, D., Durand, A., Fortin, F. A., Tanguy, K., Maldague, X., Panneton, B., & Simard, M. J. (2014). Bayesian classification and unsupervised learning for isolating weeds in row crops. Pattern Analysis and Applications, 17(2), 401-414. https://doi.org/10.1007/s10044-012-0307-5

118. Ampatzidis, Y., Partel, V., & Costa, L. (2020). Agroview: Cloud-based application to process, analyze and visualize UAV-collected data for precision agriculture applications utilizing artificial intelligence. Computers and Electronics in Agriculture, 174, 105457. https://doi.org/10.1016/j.compag.2020.105457

119. Tan, L., Hou, H., & Zhang, Q. (2016, July). An extensible software platform for cloud-based decision support and automation in precision agriculture. In 2016 ieee 17th international conference on information reuse and integration (iri) (pp. 218-225). IEEE. https://doi.org/10.1109/IRI.2016.35

120. Mell, P., & Grance, T. (2009). Effectively and securely using the cloud computing paradigm. NIST, Information Technology Laboratory, 2(8), 304-311. https://zxr.io/nsac/ccsw09/slides/mell.pdf

121. Lameski, P., Zdravevski, E., Trajkovik, V., & Kulakov, A. (2017, July). Cloud-based architecture for automated weed control. In IEEE EUROCON 2017-17th International Conference on Smart Technologies (pp. 757-762). IEEE. https://doi.org/10.1109/EUROCON.2017.8011212

122. Ampatzidis, Y., Partel, V., & Costa, L. (2020). Agroview: Cloud-based application to process, analyze and visualize UAV-collected data for precision agriculture applications utilizing artificial intelligence. Computers and Electronics in Agriculture, 174, 105457. https://doi.org/10.1016/j.compag.2020.105457

123. Revanasiddappa, B., Arvind, C., & Swamy, S. (2020). Real-time early detection of weed plants in pulse crop field using drone with IoT. Technology, 16(5), 1227-1242.

124. Alrowais, F., Asiri, M. M., Alabdan, R., Marzouk, R., Hilal, A. M., & Gupta, D. (2022). Hybrid leader based optimization with deep learning driven weed detection on internet of things enabled smart agriculture environment. Computers and Electrical Engineering, 104, 108411. https://doi.org/10.1016/j.compeleceng.2022.108411

125. Elijah, O., Rahim, S. K., Abioye, E. A., Salihu, Y. O., & Oremeyi, A. A. (2022, April). Decision Support Platform for Production of Chili using IoT,

Cloud Computing, and Machine Learning Approach. In 2022 IEEE Nigeria 4th International Conference on Disruptive Technologies for Sustainable Development (NIGERCON) (pp. 1-5). IEEE. https://doi.org/10.1109/NIGERCON54645.2022.9803077

126. Chen, J., & Ran, X. (2019). Deep learning with edge computing: A review. Proceedings of the IEEE, 107(8), 1655-1674. https://doi.org/10.1109/JPROC.2019.2921977

127. Chaurasia, A., & Culurciello, E. (2017, December). Linknet: Exploiting encoder representations for efficient semantic segmentation. In 2017 IEEE Visual Communications and Image Processing (VCIP) (pp. 1-4). IEEE. https://doi.org10.1109/VCIP.2017.8305148

128. Yu, C., Wang, J., Peng, C., Gao, C., Yu, G., & Sang, N. (2018). Bisenet: Bilateral segmentation network for real-time semantic segmentation. In Proceedings of the European conference on computer vision (ECCV) (pp. 325-341). https://doi.org/10.48550/arXiv.1808.00897

129. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778). https://doi.org/10.48550/arXiv.1512.03385

130. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2818-2826). https://doi.org/10.48550/arXiv.1512.00567

131. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9). https://doi.org/10.48550/arXiv.1409.4842

132. Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., ... & Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.

133. Zhang, X., Zhou, X., Lin, M., & Sun, J. (2018). Shufflenet: An extremely efficient convolutional neural network for mobile devices. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 6848-6856). https://doi.org/10.48550/arXiv.1707.01083

134. Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1251-1258). https://doi.org/10.48550/arXiv.1610.02357

135. Dadashzadeh, M., Abbaspour-Gilandeh, Y., Mesri-Gundoshmian, T., Sabzi, S., Hernández-Hernández, J. L., Hernández-Hernández, M., & Arribas, J. I. (2020). Weed classification for site-specific weed management using an automated stereo computer-vision machine-learning system in rice fields. Plants, 9(5), 559. https://doi.org/10.3390/plants9050559

136. Sabzi, S., & Abbaspour-Gilandeh, Y. (2018). Using video processing to classify potato plant and three types of weed using hybrid of artificial neural

network and partincle swarm algorithm. Measurement, 126, 22-36. https://doi.org/10.1016/j.measurement.2018.05.037

137. Kamath, R., Balachandra, M., & Prabhu, S. (2020). Crop and weed discrimination using Laws' texture masks. International Journal of Agricultural and Biological Engineering, 13(1), 191-197. https://doi.org/10.25165/j.ijabe.20201301.4920

138. Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. IEEE internet of things journal, 3(5), 637-646. https://doi.org/10.1109/JIOT.2016.2579198.

139. Assunção, E., Gaspar, P. D., Mesquita, R., Simões, M. P., Alibabaei, K., Veiros, A., & Proença, H. (2022). Real-Time Weed Control Application Using a Jetson Nano Edge Device and a Spray Mechanism. Remote Sensing, 14(17), 4217. https://doi.org/10.3390/rs14174217

140. Giselsson, T. M., Jørgensen, R. N., Jensen, P. K., Dyrmann, M., & Midtiby, H. S. (2017). A public image database for benchmark of plant seedling classification algorithms. arXiv preprint arXiv:1711.05458. https://doi.org/10.48550/arXiv.1711.05458

141. Chebrolu, N., Lottes, P., Schaefer, A., Winterhalter, W., Burgard, W., & Stachniss, C. (2017). Agricultural robot dataset for plant classification, localization and mapping on sugar beet fields. The International Journal of Robotics Research, 36(10), 1045-1052. https://doi.org/10.1177/0278364917720510

142. Lameski, P., Zdravevski, E., Trajkovik, V., & Kulakov, A. (2017, September). Weed detection dataset with RGB images taken under variable light conditions. In International Conference on ICT Innovations (pp. 112-119). Springer, Cham. https://doi.org/10.1109/ACCESS.2017.2684913.

143. Sa, I., Chen, Z., Popović, M., Khanna, R., Liebisch, F., Nieto, J., & Siegwart, R. (2017). weednet: Dense semantic weed classification using multispectral images and mav for smart farming. IEEE robotics and automation letters, 3(1), 588-595. https://doi.org/10.1109/LRA.2017.2774979

144. 144.Haug, S., & Ostermann, J. (2014, September). A crop/weed field image dataset for the evaluation of computer vision based precision agriculture tasks. In European conference on computer vision (pp. 105-116). Springer, Cham. https://doi.org/10.5194/isprsannals-II-3-W5-311-2015

145. Leminen Madsen, S., Mathiassen, S. K., Dyrmann, M., Laursen, M. S., Paz, L. C., & Jørgensen, R. N. (2020). Open plant phenotype database of common weeds in Denmark. Remote Sensing, 12(8), 1246. https://doi.org/10.3390/rs12081246

146. Espejo-Garcia, B., Mylonas, N., Athanasakos, L., Fountas, S., & Vasilakoglou, I. (2020). Towards weeds identification assistance through transfer learning. Computers and Electronics in Agriculture, 171, 105306. https://doi.org/10.1016/j.compag.2020.105306

147. Teimouri, N., Dyrmann, M., Nielsen, P. R., Mathiassen, S. K., Somerville, G. J., & Jørgensen, R. N. (2018). Weed growth stage estimator using deep

convolutional neural networks. Sensors, 18(5), 1580. https://doi.org/10.3390/s18051580

148. Assunção, E., Gaspar, P. D., Mesquita, R., Simões, M. P., Alibabaei, K., Veiros, A., & Proença, H. (2022). Real-Time Weed Control Application Using a Jetson Nano Edge Device and a Spray Mechanism. Remote Sensing, 14(17), 4217. https://doi.org/10.3390/rs14174217

149. Ramos, S., Gehrig, S., Pinggera, P., Franke, U., & Rother, C. (2017, June). Detecting unexpected obstacles for self-driving cars: Fusing deep learning and geometric modeling. In 2017 IEEE Intelligent Vehicles Symposium (IV) (pp. 1025-1032). IEEE.

150. Treml, M., Arjona-Medina, J., Unterthiner, T., Durgesh, R., Friedmann, F., Schuberth, P., ... & Hochreiter, S. (2016). Speeding up semantic segmentation for autonomous driving.

151. Badue, C., Guidolini, R., Carneiro, R. V., Azevedo, P., Cardoso, V. B., Forechi, A., ... & De Souza, A. F. (2021). Self-driving cars: A survey. Expert Systems with Applications, 165, 113816.

152. Azuma, R. T. (1997). A survey of augmented reality. Presence: teleoperators & virtual environments, 6(4), 355-385.

153. Billinghurst, M., Clark, A., & Lee, G. (2015). A survey of augmented reality. Foundations and Trends® in Human–Computer Interaction, 8(2-3), 73-272.

3 Chapter 3: Performance comparison of semantic segmentation based deep learning models on weed detection in maize fields

This manuscript has been prepared for journal submission

3.1 Introduction

According to the latest surveys conducted by the Department of Economic and Social Affairs, United Nations – the world population is expected to reach at 8.5 billion people by 2030 and an additional 1.8 billion in the following two decades reaching to 9.7 billion by 2050 [68]. A major concern associated with the increasing population trends is related to the mass production of food and fiber along with a smaller rural labor force. An estimated increase in global food production of around 70-100% is required to meet the demands of the population. These trends demand an increase for research and development in agriculture domain bringing out much attention to various sectors such as generation of high-seed quality to increase crop yield, improving soil health for better crop growth, and prevention of crop yield losses. One of the major concerns that leads to crop yield losses in agriculture are associated with the growth of weeds, pests, and crop-diseases in the fields which are responsible for a global crop yield loss of around 40%. Weeds are unwanted plants that grow in fields and compete with crops for water, light, nutrients, and space leading to causing crop yield losses [41, 42]. Traditional practices associated with handling weeds in fields include various manual techniques such as hoeing, tilling, burning, and spraying of chemicals have already been used in agriculture fields. The most used techniques to handle these weed infestations is the use of uniform rate spraying of herbicides in the fields. Weed infestations are usually observed at a few locations in fields and the need to perform uniform spraying activities could be avoided. Practicing herbicide spraying activities at a uniform rate across the fields are associated with a few disadvantages such as economic loss due to excessive usage, environmental pollution, contamination of non-target vegetation and evolution of herbicide-resistant weeds [24].

Precision agriculture proposes site-specific weed management techniques which perform the task of detecting weed infestations and using different weeding techniques to remove weeds from the fields. Computer vision techniques have been used to develop weed detection models that make use of data collected from various remote sensing techniques such satellites, LIDAR, UAVs, UGVs to generate weed maps that could be used to perform the tasks of SSWM. Several features such as shape, color, band information, texture, height were collected and used to develop weed detection models. A two step-approach to develop an inter-line weed detection model was developed by [62] using image processing techniques. Inter-line crop row detections were made using Hough transformation algorithm and Normalized Difference Vegetation Indices were used to detect weed infestation that existed between crop rows in the study. A study was conducted by [59], developed a weed detection model in which color-based machine vision techniques were used to perform detection of reddish weed stems in wheat and soybean fields which have greenish stems. Spectral characteristics such as texture, color, shape was used to differentiate between crops, weeds, and soil [52] and a real-time weed detection model were developed and deployed in agriculture fields. Real-time weed detection applications were developed using Hough transform and simple linear iterative clustering techniques to detect inter and intra line weeds in crops [53]. A Discriminant analysis and an Artificial Neural network were designed in the research where misclassification rates in performing weed detection were below 3 percent. The measurements of reflectance characteristics of crops and weed in visual and near-infra-red rays were utilized to generate a weed detection model in laboratory conditions in the research performed by [60]. A weed detection model to classify broad-leaved weed were developed by [61] using physical characteristics extracted from imagery collected from ground vehicles. Color based information were used to perform elimination between vegetation and soil, and spatial information such as shape and texture-based features were used to perform classifications between crops and weeds in the study. Most of the traditional weed detection models developed using machine vision and image processing techniques were trained to perform classifications tasks under constrained laboratory conditions and didn’t perform well in natural circumstances in fields.

The collection of multi-spectral images from the UAV remote sensing techniques have helped in extraction of useful features which have been used with various ML algorithms such as SVMs, DTs, RFs to develop weed detection models in recent past. Due to their better learning capabilities on features extracted from these images, ML algorithms have helped achieve higher accuracy scores to perform the classifications tasks between crops and weeds. However, training a weed detection has been a complex task due to the homogenous crop-weed characteristics such as color similarity, high occlusions, similar reflective index, indistinguishable shape, and texture features. Performance comparisons between weed detection models developed using Logistic Regression and SVMs based ML algorithms were made in detecting weeds in carrot crops [63]. The histogram of gradients features were extracted from the data and it was observed that SVMs performed better classification tasks as compared to Logistic regression based weed detection model. Pattern based recognition techniques were used to train weed detection model using several ML algorithms such as KNN, SVM, and ANN in sugarcane crops. ANN based ML models performed the best and were able to achieve accuracy scores of 63.1% [64]. The model proposed by [65], used an ANN that trained on a combination of Red, Green, NIR, and texture features extracted using the high-quality imagery from UAVs to generate weed maps in leguminous crops which achieved an accuracy score of 98.87 %. Shape vectors along with Fourier descriptors and moment invariant features were used to train a weed detection model in classifying four different weed categories in Sugar beet plants. The weed detection model helped achieved an overall classification accuracy of 94.5% [55]. The use of ensemble learning techniques have helped improved the training of ML algorithms. Ensemble learning refers to training multiple ML models on different samples from training datasets and aggregating responses from all the models to generate a final output. ML models developed using ensemble learning techniques have proven to overcome overfitting and have helped generate models that generalize well on unseen dataset. One of the powerful supervised ML algorithms that uses ensemble learning techniques is RFs. RF constructs multiple DTs during training and uses an aggregation of outputs from all DTs in a combination to perform the final classification and regression tasks during prediction. The research conducted by [66] has developed RF-based ML algorithms to detect alligator weeds that form dense infestations in aquatic environments using remote sensing data to improve biosecurity supervision and monitoring efforts in Australia.

DL techniques using CNN have revolutionized computer vision tasks by creating state-of-art models to perform the complex tasks of image classification, object detection, semantic segmentation, and instance segmentation [9, 10, 12, 14, 18]. CNN based deep learning approach have gained much importance over the past few years for detection and classifications tasks because of their feature generation and better self-learning capabilities as compared to conventional image processing or vegetation index-based approaches. The training of a DL models requires systems with high configurational capabilities to perform the complex computations required during the training phase. The availability of systems having better computer hardware, software capabilities and use of GPUs have helped in training DL models to achieve higher accuracy scores. Several weed detection models have been developed using Object Detection and Semantic Segmentation techniques. Object detection perform the task of generating bounding boxes around detected objects during prediction and Semantic segmentation models generate pixel-wise mapping of different class categories detected during predictions. The tasks of SS have gained much importance in recent past as they better perform the tasks of scene understanding and helps in generating better accuracy scores during predictions. SS based DL approaches are still not fully exploited to develop weed detection models due to the unavailability of labelled data required to train these models. A major challenge associated in developing weed detection models using these approaches is the class imbalance issue that arise due to the availability of unequal proportions of pixel labels for different class categories used for training of the model. This imbalance problem prevents the model to train equally well in predicting pixel values that are underrepresented and are not present in equal proportions with other classes. This issue leads to the problem of underfitting in which the trained model does not perform well in detecting underrepresented pixel values during prediction.

The objective of this study was to develop Semantic Segmentation based Deep Learning models that perform weed detection in maize. This is a complex task due to the crop-weed color similarity and high occlusions between crops and weeds found in the maize fields. Amaranthus Palmeri is an important and a common weed category usually found in maize and soybean fields.. Also, these weeds come under the category of herbicide-resistant weeds and are resistant to glyphosate, which is the most commonly used herbicide. Extraction of Semantic segmentation model results could be used with weeding actuators in fields to perform the complex task of Site-Specific Weed Management and remove weed patches. The benefits associated with Site-Specific Weed Management practices ends up being two-folds: Firstly, it prevents environmental pollution linked with excessive uniform spraying of herbicides in fields and performs weeding only at locations where weed patches are found. Secondly, it decreases the economic burden on farmers as these practices leads to reduction of both herbicides' quantities being used and human labor costs associated with spraying activities in agriculture fields. The paper has the following contributions:

1. Performing pixel-wise classifications of weeds, soil, and crops in RGB images from maize fields using Semantic Segmentation based Deep Learning approaches.
2. Two-step approach in optimizing the training of weed detection models to resolve the challenges associated with class-imbalance problem are explained
3. Performance comparisons between two Semantic Segmentation models of LinkNet and UNet developed using various architecture backbones such as VGG, Inception, and ResNet in the encoder section

Table 3. 1 Abbreviations table

Abbreviation	Full form
GB	Giga Bytes
RAM	Random Access Memory
GHz	Giga Hertz
RGB	Red, Green Blue
DL	Deep Learning
ML	Machine Learning
GPU	Graphical Processing Unit
AGL	Above Ground Level
UAV	Unmanned Aerial Vehicles
CMOS	Complementary Metal Oxide Semiconductor
OD	Object Detection
CVAT	Computer Vision Annotation Toolkit
SS	Semantic Segmentation
SSWM	Site-Specific Weed Management

The remaining research article is organized as follows. Section 2 explain the encoder-decoder architecture of SS models and related works of weed detection models developed using these approaches are discussed. Section 3 gives a detailed description of the proposed method discussed. Section 4 presents the experimental results obtained after training state-of-art weed detection models in this study. Also, performance comparisons between weed detection models developed based on various evaluation metrics obtained during testing are discussed. Section 5 presents the conclusion of our study and discusses possible future outcomes associated with developing real-time weed detection models to perform tasks of SSWM.

3.2 Semantic Segmentation

SS based DL models perform the tasks of pixel-wise classification of each and every pixel in an image to a separate class category, and the pixel values corresponding to similar class categories are labelled as a single entity during prediction. These tasks have helped elucidate complex tasks of scene understanding. There has been a surge in the research related to performing tasks such as scene understanding due to rise in developing real-time applications for self-driving cars [3, 4, 5], augmented reality [6, 7], etc. in recent past as performing pixel-wise classifications is an important task in these techniques. The task of labelling each and every pixel in an image or a video helps in performing the tasks of precise object recognition and localization during predictions. Since Alexnet [8] won the ImageNet challenge in 2012, DL methods have been extensively researched upon and developed continuously with modifications to perform the complex tasks of image classification, object-detection, semantic segmentation, and instance segmentation tasks. DL methods comprising of CNN architectures have gained much popularity due to their feature extraction and self-learning capabilities. DL approaches have been extensively used to perform the classification and detection tasks, a difference between different classification tasks is shown in Figure 3.6. The image shows the difference between the final predicted outputs using various DL models. The first image performs the task of image classification in which labels are generated for various objects detected in an image, object detection performs the task of detecting various objects and localizing a bounding box around these objects as an output from these models. Several popular DL models such as Fast RCNN [9], Faster RCNN [10], YOLO [11], SSD [12] have been developed to solve the tasks of OD in recent past. As discussed earlier about the SS models, Several SS based DL models have been developed to perform the tasks of pixel-wise classification such as UNet [13], LinkNet [14], PSPNet [15], SegNet [16], BiSeNet [17] have been already developed. Instance segmentation performs a more complex task than SS as it implements an algorithm that classifies each and every pixel to a specific label and it considers objects of similar class category as an individual which is different from SS approach which detects objects of similar class categories as a single entity. DL models such as Mask RCNN [18], Panoptic segmentation [19] perform the tasks of instance segmentation. The success achieved by researchers in performing complex recognition tasks using DL approaches have motivated various studies in agriculture domain perform the complex recognition tasks such as different crops detection [21], fruit yield estimation [22], detecting cotton balls in fields [23], weed detection [24], tassel detection [25], disease detection [26] in agriculture domain.

In our research, we explored SS based DL models to perform the tasks of weed recognition in maize fields. The main intuition behind using SS based DL models in our research were that these models have two-folds benefits of better performance on complex computer vision tasks and these models could be used on embedded devices for real-time weeding applications in fields. Also, these models have not been fully explored to perform the task of weed recognition in agriculture domain for real-time SSWM applications. In our research we made performance comparisons on weed detection models developed using SS based DL models of UNet [13] and LinkNet [14]. These models have an encoder-decoder architecture, and their working are explained in the following two sub-sections:

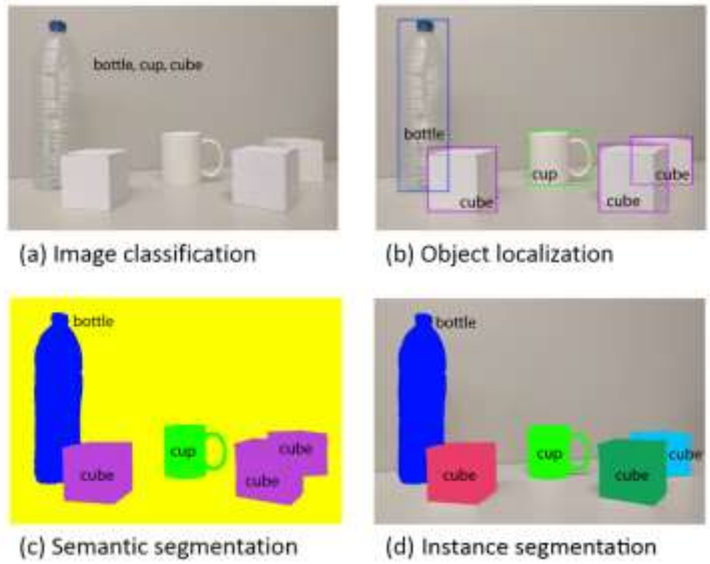

Figure 3.1 An example of different computer vision tasks.

This figure is borrowed from [2]

3.2.1 UNet

UNet is a powerful DL model that performs the SS tasks and was developed by [13] to perform biomedical image segmentation tasks. This architecture won the ISBI cell tracking challenge in 2015 and thereafter, has been extensively used for performing various DL applications. UNet models have been shown great success in training DL models with only a limited dataset and have achieved higher accuracy scores during predictions [27]. UNet model has achieved great success in SS due to its symmetrical encoder-decoder architecture as shown in Figure 3.2. This architecture has various DL tools such as Convolutional filters, max-pool layers, up-convolutional layers, arranged in such a fashion that forms a U like structure from which the model got its name.

The encoder part consists of Convolutional filters and maxpool layers that helps to extract features/textures/patterns from the images that are further used for training DL model. Then the images are down sampled regularly using maxpool layers until it reaches to its base layer where it forms a condensed representation of original image. As the number of convolutional filters increase during consecutive layers in the encoder sections, regular down sampling helps in decreasing spatial dimensions of images and thus leads to decreasing of overall computations performed during training. The decoder part consists of transposed conv layers that perform the up sampling. For localization to be more precise, so-called skip connections are used where the feature maps from the encoder block are concatenated to the output of transposed conv of the same layer. This process leads to the condensed representation of an image back to the original input image.

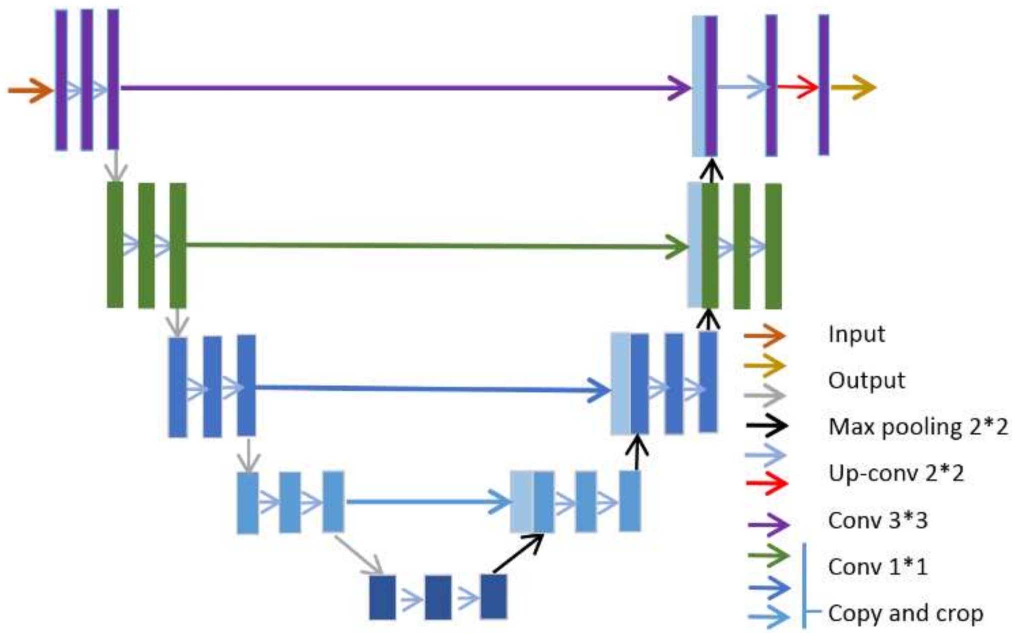

Figure 3.2 UNet architecture [13]

3.2.2 LinkNet

LinkNet [14] is a powerful SS based DL model which is an optimized version of UNet architecture, and this model also consists of various DL tools such as Convolutional filters, max-pool layers, and up-convolutional layers. The major difference between both of these models is that LinkNet makes uses fewer parameters that could result in faster execution during prediction and in turn could be used for real-time application tasks. This model also has an encoder-decoder architecture and the main difference between LinkNet, and other SS based models is the usage of information exchange between encoder-decoder. On comparison with other models instead of sharing entire feature maps between encoder and decoder as in UNet, only the input to the encoder block is shared with the corresponding decoder block in the model. The information share amongst encode-decoder blocks could be depicted from Figure 3 that depicts the LinkNet architecture. In spite of using fewer parameters as compared to UNet, LinkNet has shown better results during predictions on popular datasets such as CityScapes [28] and CamVid [29], and ImageNet [30] as well.

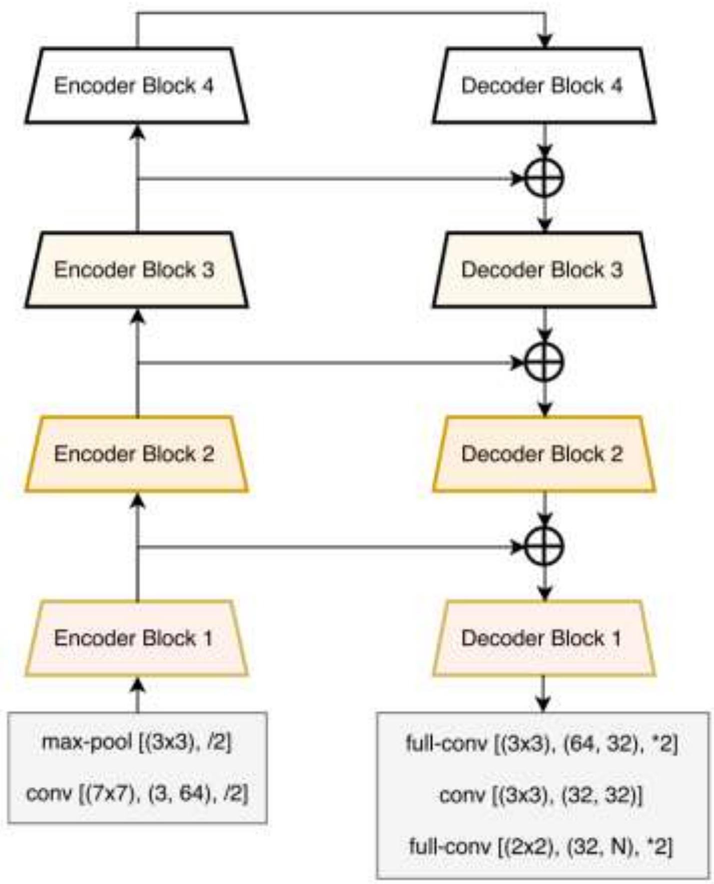

Figure 3.3 LinkNet architecture [14]

3.3 Methodology

3.3.1 Data acquisition

UAV imageries were collected from maize fields in Carleton (Thayer County - Southeastern Nebraska), NE, USA (Figure 3.1) in 2021 growing season. Heavy infestations of Palmer Amaranthus (Amaranthus palmeri) a glyphosate resistant weed was found in these fields and a weed detection model focusing on detecting palmer weed infestations were developed in this research work. DJI Phantom 4 RTK drone (Figure 3.2) was used for the UAV imagery data acquisition. The camera used in the drone had a 1-inch CMOS sensor, and a 20-megapixel camera that captured images with an 84-degree field of view. The dimensions of the images were 5472 × 4648 pixels (3:2) in three bands – Red, Green, and Blue, and only RGB images were used to train various DL models in this research work to develop an economical solution. The spatial resolution of the images collected using the UAV were 0.27 cm/pixel. In order to collect drone imagery of corn crops during different growth stages, multiple drone flights were conducted on June 14, 2021, June 30, 2021, and July 9, 2021, at 10 meter and 25-meter altitude AGL respectively. A total of 2006 RGB images were collected during data collection at the study site. The drone flights were conducted at around solar noon in order to capture images with better illumination intensity. DJI ground station pro software was used to design flight missions and the drone flew at a speed of 2 meter/sec with a 90% front and sidewise overlap during data collection.

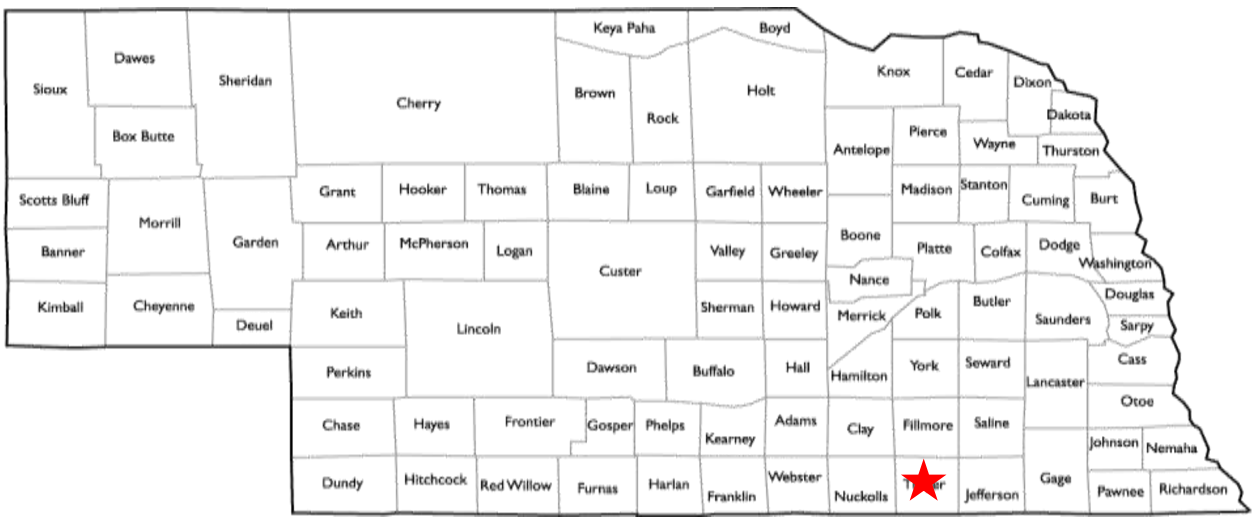

Figure 3.4 UAV data imagery collection location in Nebraska – Carleton (Thayer County), NE

Figure 3.5 DJI Phantom 4 RTK used for data collection

3.3.2 Dataset preparation

Training a SS based DL model requires meticulous pixel-wise labelling of ground truth images that could be used for training purposes. Dataset preparation is one of the most important phases in this approach as it directly contributes towards the training and in turn generalizability of the DL model towards testing. On careful analysis of the UAV imagery captured during data acquisition, a training dataset comprising of high-quality images were formed after an in-depth analysis of the UAV imagery collected during data acquisition for training DL model in this research work after proper pre-processing as described later in this section.

3.3.2.1 Data Annotation

A web-based annotation tool named CVAT was used to perform the task of pixel-wise labelling of images in our dataset. In our research, we manually annotated weeds and soil patches in the original RGB images and all the un-annotated pixels values besides these class categories (pixel values not labelled as either weeds or soil) were automatically labelled as background labels during annotation. Polygon bounding boxes were used to annotate several weeds and soil patches during the process of data annotation, a total of 8320 weed patches and 329 soil patches were annotated manually. All the images in the training dataset were annotated accurately and their corresponding masks were generated. Figure 3.5 represents the tasks of annotating original images and generating corresponding masks of these images. In summary, we had pixel-wise information of all the images in our dataset corresponding to 3 class labels namely: Background, Weeds, and Soil.

There were a few constraints that led to elimination of images collected during data collection and they were as follows -

1. It was observed that the quality of the pixel values from UAV imagery collected at 25-meter AGL were compromised. Due to this reason images captured at 25-meter AGL were skipped and only the images captured at 10-meter AGL were used in this study as they had better pixel level information.
2. During the last data collection i.e., on July 9, 2021, most of the UAV imagery collected had weed infestations hidden under corn canopies. Due to this reason these images could not be used for training our DL model as they only had a few weed patches in the images.
3. During the early growth stage both weeds and crops infestations were quite sparse in the fields. There were large soil patches inter-row spacing between crops which were quite hard to annotate. Due to this reason, they were not included in the dataset used for training.

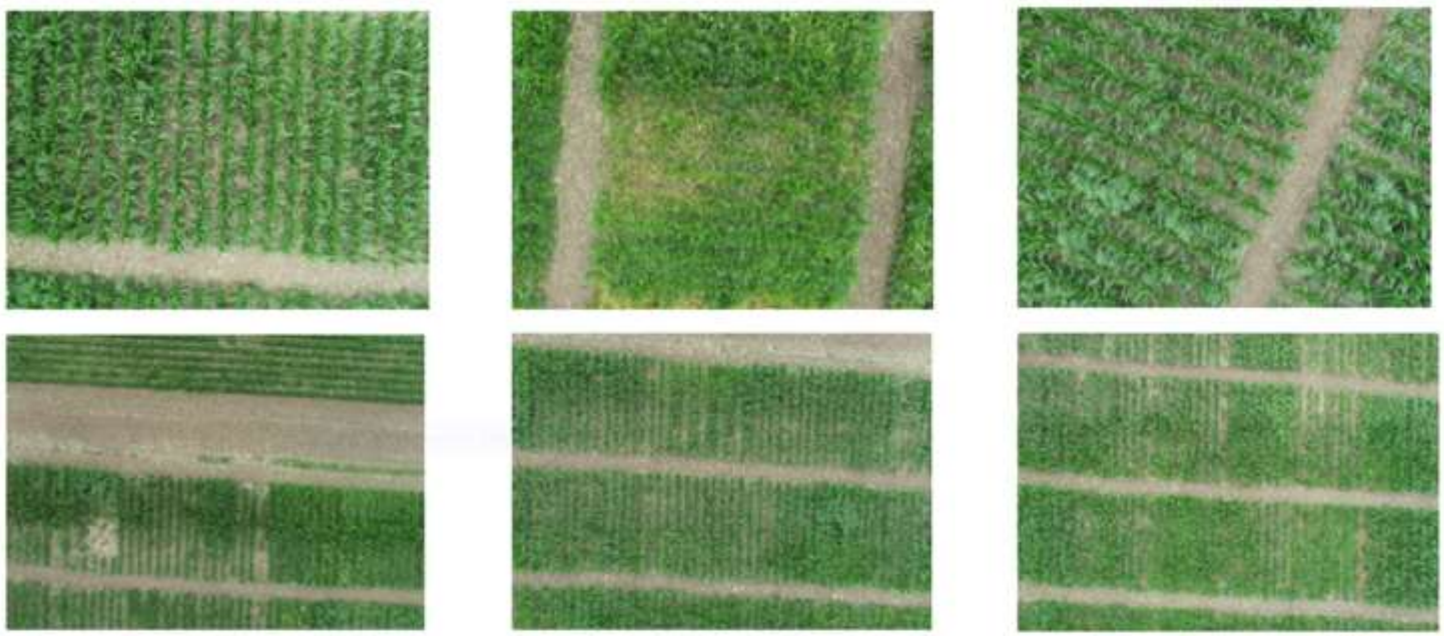

Figure 3.6 Few original UAV imagery that were collected from fields in Carleton, NE and used for developing DL models

a) Images (1-3) collected at 10-meter AGL
b) Images (4-6) collected at 25-meter AGL

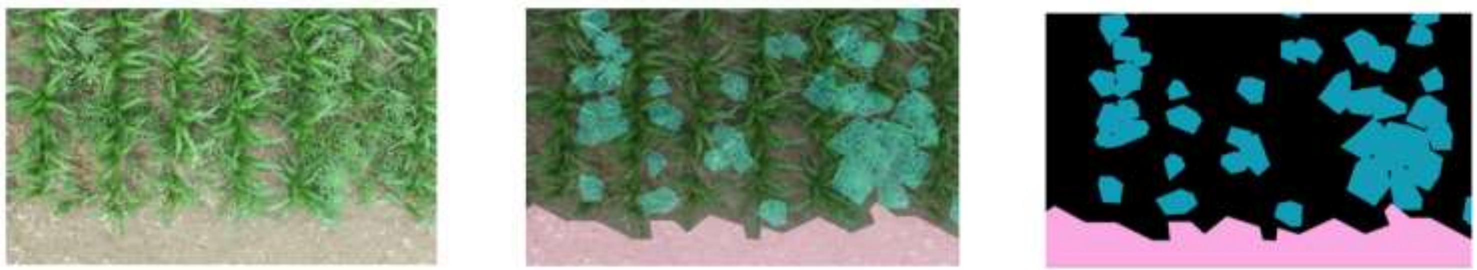

Figure 3.7 Data annotation: Pixel wise labelling of images

Soil; Weeds; Background

(a) Original sliced RGB image; (b) Polygon bounding boxes used for data annotation.

(c) Annotated image used during training of DL model

3.3.2.2 Data Augmentation

Training DL models requires large amounts of data during training, increasing the amount of data used for training has shown generation of better DL models that achieve higher accuracy scores as compared to prior. Data augmentation refers to artificially increasing the amount of data that is used for training DL models by randomly generating new data from already existing data. These techniques have been popularly used in the recent past and have helped the network to learn desired invariance and increase robustness while predictions. Data augmentation techniques that were adopted in this study were: horizontal flip, vertical flip. ImageDataGenrator is a data augmentation module available under Tensorflow library [32] was used to perform the tasks of Data augmentation in our study. These techniques were used only in the training phase and not during the testing phase. Data augmentation techniques were performed on both original RGB images in the dataset along with their corresponding masks during the training of DL models in this study. Figure 3.5 depicts a few augmentation techniques that were used in this study.

3.3.3 Transfer Learning

The SS based DL models usually consists of encoder-decoder architecture. The encoder sections perform the tasks of feature extractions and down-sampling images using Convolutional and MaxPool layers. The uniqueness in the architectures of various SS models such as UNet[13]. LinkNet [14] lies in the method of information exchange between the encoder-decoder layers which helps in recovering the spatial information lost during continuous down sampling in

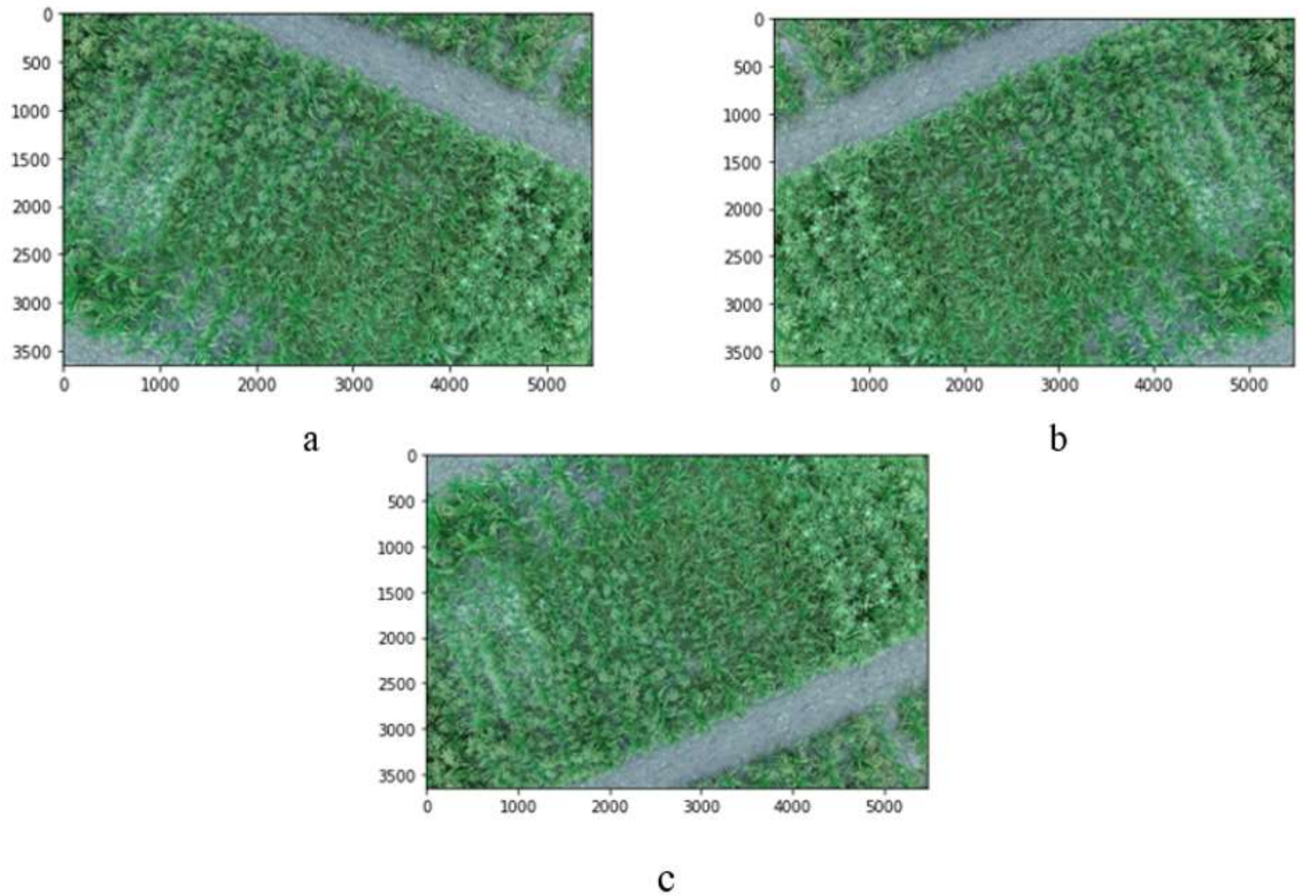

Figure 3.6 Data augmentation performed in this study

a) Original image b) Horizontally flipped image, c) Vertically flipped image encoder sections. This process helps to generate better results during predictions. To better perform the tasks of features extractions from the images used for training SS based DL models, transfer learning techniques could be used to obtain better results from training these models. Encoder section is the most important phase in training SS models as it performs the tasks of feature extractions. Several state-of-art architecture backbones such as VGG [32], ResNet [33], Inception [34] already trained on ImageNet datasets could be used to perform the tasks of feature extraction and better train SS models rather training these models from scratch. Transfer learning techniques makes use of pre-trained weights already trained to extract useful features on large datasets such as CamVid [29], ImageNet [30], etc., this approach helps to transfer the knowledge already acquired to solve a similar OD or SS based DL tasks. Fine-tuning of these models could also be adopted while using transfer learning techniques which helps the model to update the weights in such a manner that they could perform a different detection/classification task optimally. These architectures could also be referred to as backbones that are used while training SS based DL models. We trained and made performance comparisons on both UNet and LinkNet models trained using VGG [32], ResNet [33], and Inception [34] backbones. The main difference in these architectures backbones lies in the orientation and layout of various DL tools such as Convolutional filters, Maxpool layers, used for feature extraction. In this research we made performance comparisons on developing SS based models using three different architectures and these architectures are briefly explained in the following sub-sections.

3.3.3.1 VGG

Since VGG won ImageNet contest in 2014, it has been extensively used as a backbone for various OD and SS based DL models for the process of feature extraction. The major difference between VGG network and other architecture backbones popular during that time was in the depth and layout of various convolutional filters and fully connected layers. VGG networks had a depth of around 16-19 layers and the number of convolutional layers in the architecture were around 8-16 [35]. A significant difference between VGG network and other CNN architectures was the use of 3×3 convolutional filter kernels which helped to build deeper models and also generate better feature extractions from the input data fed to the DL model. Also, using 3×3 convolutional filter kernels helped to decrease the number of parameters being used in the training of the model and were able to generate better results with less use of computational power during training as compared to Alexnet [8]. The network had three fully connected layers in the last, 2 layers had fully connected 4096 neurons and a SoftMax layer was used as a last layer to perform and learn the classification tasks during training.

3.3.3.2 Inception

Inception modules were developed and used in Inception architectures that made use of convolutional filters of various kernel sizes in a single layer to perform the feature extraction process. Various filter size kernels such 1×1 , 3×3, 5×5 were used in a single layer and their average were taken in the end to learn features. Rather than using convolutional filters of single size of 3×3 as used in VGG network, Inception modules made use of multiple filter size kernels in a single layer to perform feature extraction process. Making use of inception modules had the following advantages, these modules helped to decrease the number of parameters being used during training and this in turn help to reduce computational efficiency. This was achieved by using multiple lower size convolutional filters as compared to using larger convolutional filters in convolutional layers, it was observed that two 3×3 convolutional filters ($2 * (3 * 3) = 18$ parameters) used less parameters than one 5×5 ($5 * 5 = 25$ parameters) convolutional filter size. These benefits made use of inception backbones popular in developing DL models. Some common Inception backbones used in DL models are: InceptionV2, InceptionV3, InceptionResNetV2.

3.3.3.3 ResNet

The development of architectures such as VGG [32] and Inception [34] popularized development of deeper CNNs after its creation several deeper architectures were developed by researchers thereafter. A rising concern that became popular in deeper CNN architectures was the issue of vanishing gradient problem. Backpropagation techniques are used by DL models to learn and execute the training process and perform complex tasks of OD and SS. In these deeper architectures, certain predictions were made during testing and using the concept of backpropagation weights were updated in the previous layers so as the model learns and performs ideally. An approach was developed by [33] in which certain skip connections were used between convolutional layers to mitigate the performance decay and perform the task of backpropagation optimally during training [35]. ResNet architecture won the ImageNet challenge in 2015 and has been one of the most popular architecture backbones that have been used to solve various complex recognition and detection tasks. Development of ResNet architectures made it possible to use deeper CNN architectures in DL models and resolve the vanishing gradient problem. ResNet18, ResNet34, ResNet50, ResNet101, and ResNet152 architecture backbones. The major difference between these CNNs is the total number of layers being used in these architectures to perform the feature extraction process. We used ResNet 18 and ResNet34 architecture backbones in the encoder sections and made performance comparisons along other architecture backbones in our study. ResNet34 is a deeper architecture that uses a greater number of parameters during training as compared to ResNet18 architecture.

3.3.4 Class-imbalance problem

The original image dimensions of 5472 × 3648 pixels were too large to fit in the memory for processing while training DL model, so the original images were sliced into 384 sub-images of dimensions 228 × 228 pixels. A total of 38,400 sub-images were formed of these original RGB images and were included in the training dataset. Dividing original dataset into sub-images led to an issue of class imbalance problem as most of the sub-images belonged to background class labels or had most of the pixel values associated with background class category. Figure 3.9 depicts a few images from our original dataset that explain the class imbalance problem by depicting weeds, soil, and background class categories in the images included in our dataset. Class imbalance issues is one of the major factors that makes it hard to train DL models to perform the tasks of better predicting the under-represented samples that are being used for training the model. Several techniques such as increasing under-represented samples using Data augmentation techniques, practice of using under sampling (to use less samples during training) for majority class or over sampling (to use more samples during testing) of minority class, using ensemble learning techniques to train a ML model have been adopted in recent past to handle these issues. The novelty in this research lies in the two approaches which were adopted to overcome the class imbalance problem, and these were as follows –

3.3.4.1 Refining training dataset

The main intuition behind training a weed detection model is to train the model that has a higher generalizability over unseen dataset and performs well in accurately detecting weed patches in order for the model to be used for real-time SSWM applications. To perform the refining of the dataset in a manner such the dataset to be used for training has a balance of pixel values of each class category a strategy was adopted. A new training dataset were formed by analyzing each and every image in the original dataset. An algorithm behind the strategy adopted to form a new training dataset is explained in Algorithm 3.1. An image was included in the new training dataset if and only if the total number of background pixels values in that image were < 95% of the total pixel values in that image. This strategy helped to remove the images that pre-dominantly include only background pixel values in them and created a balance in the dataset which had either of weeds or soil pixel values to be used for training DL model. Also, details about the total number of images in each dataset namely: Training, Validation, and Testing before and after refining of the dataset are mentioned in Table 1 respectively.

Algorithm 3.1: *Creation of a new dataset to be used for training*

New Training Dataset = []

for i in range (total number of images):

image = cv2.imread (Read image in the dataset)

if (# of background pixel values in an image) >= (total pixel values in an image):

continue # Image was not included in the new training dataset

else:

New Training Dataset.append(image) # Image included in the new training dataset

Table 3. 2 Number of images in different categories before and after refining original dataset

	Training	Validation	Testing	Total
Before	28800	5760	3840	38400
After	9366	1873	3840	15079

3.3.4.2 Optimizing backpropagation

DL could be referred to as an ability to train computers in the ability to make use of neurons and imitate the working as in human brain. The human brain consists of millions of neurons, and it makes use of these neurons to perform decision making. Supervised DL models consists of multiple layers such as convolutional layers, fully connected layers consisting of neurons that are initialized with random weights during training, these randomized weights are regularly updated while training these DL models in such a manner that they could optimally perform the tasks of classifications or predictions. Backpropagation techniques are used to update these weights in DL models and these techniques perform the most important task in training a DL model. Continuous research to optimize these backpropagation techniques by tuning hyperparameters used in training have been performed in recent past and is a topic of continuous research in the DL domain. Several loss functions have been developed in the recent past that help refining such as Class imbalance problem. In our research, we train our SS based DL

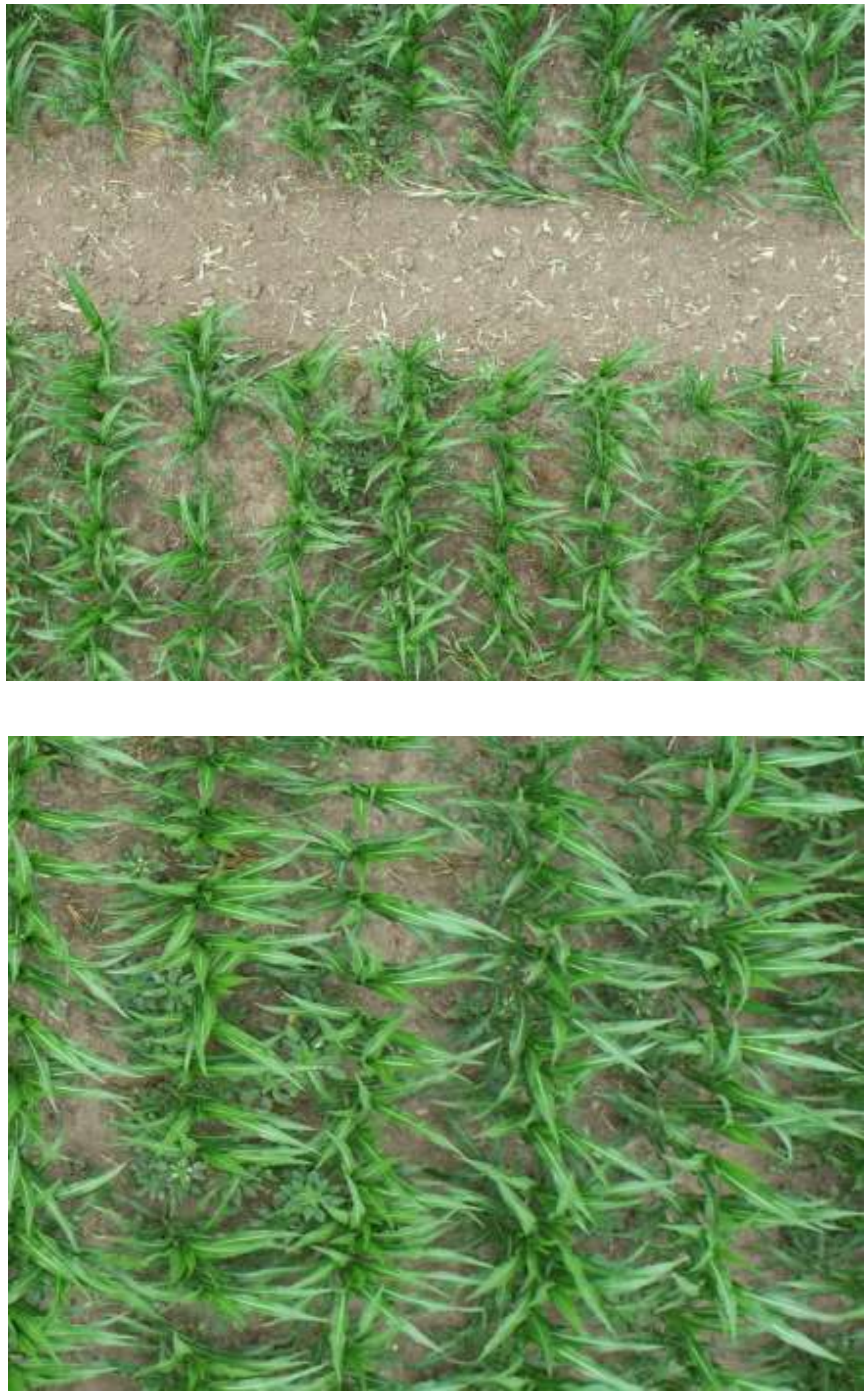

Figure 3.6 Images explaining the class imbalance problem between different class categories in our dataset

models using a fixed sets of hyperparameters and extract the best state-of-art weed detection model. As a next step, to further optimize our weed detection model we make use of several loss functions such as Focal loss [36], Dice Loss [38], and Jaccard Loss [39] while training state-of-art model and performance comparisons were made.

3.3.5 Experimental setup

3.3.5.1 Software and hardware setup

An Alienware 17 R5 laptop with the following specifications 64-bit OS, 2.9 GHz core Intel i9-8590HK processor with 32 GB of RAM was used to train the DL models. An 8 GB NVIDIA GeForce GTX 1080 GPU was also installed on the laptop to achieve optimal performance during training DL models. Windows 10 Enterprise software environment was installed on the machine. DL codes were written in Python 3.9.10 programming language and use of DL libraries such as Tensorflow 2.8.0, Keras 2.8.0, and Segmentation models API were made to train semantic segmentation-based DL models of UNet and LinkNet in this research. Training of DL models were executed in Spyder text editor and matplotlib library were used for visualizations of predictions during testing phase.

3.3.5.2 Hyperparameter tuning

In this study, we tried and tested a few different sets of hyperparameters. All the images used for training DL models in this research had a fixed dimension of 228 × 228 pixels. After a few testings, a suitable set of hyperparameters that helped in an optimum training a state-of-art DL models were set and used for training all the DL models whose performance were evaluated in this study. All the models were trained for 50 epochs, and each epoch ran for 586 iterations as we used a batch size of 16 while training. We used cross validation and data augmentation techniques to overcome the issues of overfitting during training phase. We used transfer learning techniques during training and various architectures such as ResNet, VGG, Inception already trained on ImageNet dataset were used to train the model. Also, fine-tuning of the weights used in these architectures were adopted to help optimize these architectures and learn better to perform a specific task of weed detection. Relu, a non-linear activation function was used as this helps to train the model faster and reliably. Adam optimizer with a learning rate of 1e-3 was used during training phase. A detailed description of hyperparameters used to train DL models are described in Table 2.

3.3.6 Evaluation metrics

In this research, several evaluation metrics such as Precision, Recall, F1-score, and IoU were evaluated during the training phase. The respective validation metric scores were extracted after comparing the trained model performance on the validation dataset. These scores were analyzed regularly so as to note if the model were showing trends of either overfitting or underfitting during training. It was observed while performing the assessment for SS based DL models evaluation metrics such as Precision, Recall, and F1-score did not give a better idea on how well the model was performing due to high class imbalance [31]. IoU score gives an estimation of the overlap between the predicted and the ground-truth class labels and is a better evaluation metric for SS based DL models.

Table 3. 3 Number of images in different categories before and after refining original dataset

Parameters	Values
Learning rate	1e-3
Optimizer	Adam
Activation function	Relu
Training Epochs	50
Focal loss (γ)	2
Batch size	16
Input size	228 × 228
Train, Validation, testing split	75%, 15%, 10%
Data Augmentation	True
Encoder weights	ImageNet
Fine-tuning	True
IoU Threshold	0.5

1. Precision = $\frac{\sum TP}{\sum TP + \sum FP}$

2. Recall = $\frac{\sum T/P}{\sum TP + \sum FP}$

3. F1–score = $2 \times \frac{Precision \times Recall}{Precision + Recall}$

4. IoU = $\frac{\sum TP}{\sum TP + \sum FP + \sum FN}$

Here, TP – True Positive, FP - False Positive, and FN – False Negative.

$\boldsymbol{IoU_i}$ for pixel i, could be described as:

$$\boldsymbol{IoU_i = \frac{Area\ of\ overlap}{Area\ of\ Union}}$$

The predictions of the trained model were performed on testing dataset, which comprised of images which were not used to train the state-of-art weed detection model. The predictions of DL model performed on the testing dataset, so as to check on how well the model generalizes on unseen dataset. This process also helped to measure on how well the model could perform on unseen dataset and also to check if the model was overfitting. IoU is one of the most important metrics that is evaluated during the tasks of SS based DL models as it helps to predict on how well the model is able to perform the task of pixel-wise classification of all pixels. All the predicted pixel values were compared with the ground-truth pixels values during prediction and both class-wise IoU and mean IoU score were evaluated to check the performance of the trained model.

3.4 Results discussion

In this section we evaluate performance comparisons between SS based DL models of LinkNet and UNet developed using various architecture backbones such as VGG, Inception, and ResNet. A detailed description of various architecture backbone that were used as feature extractors is explained in Section 2.4. The main difference between these architecture backbones lies in the layout of various DL tools such as Convolutional layer, MaxPool layers while designing these CNNs. The total number of parameters being used while training decides the training time required to train a DL model. Increase in the total number of parameters used by the model leads to performing a higher number of computations while training a DL model. Thus, this leads to an increase in both longer training and inference time during testing. Table 3 describes the total number of parameters that were being used while training weed detection models. We observe that ResNet18 architecture backbone when used as a feature extractor for training our weed detection models uses minimum number of parameters both for LinkNet and UNet models. These models converge faster and have a shorter training and testing time as compared to others.

A total of eight SS based DL models were developed and their performances were evaluated in this research. Evaluation metrics such as Precision, Recall, IoU, F1 score, and Mean IoU were observed, and the results are presented in this section. We utilized cross-validation techniques while performing the training procedure and divided our dataset into three sets of Training, Validation, and Testing in a ratio 75%, 15%, and 10%.

Table 3. 4 Comparison between number of parameters used during training in LinkNet and UNet

Feature extractor (Encoder section)	LinkNet	UNet
ResNet 18	**11.5 M**	**14.3 M**
VGG 16	20.3 M	23.7 M
ResNet 34	21.6 M	24.5 M
InceptionV3	26.2 M	29.9 M

'Adam' optimizer and 'Categorical_cross_entropy' loss function was used while training DL models. We trained our model for 50 epochs and complete details of the hyperparameters used to train our model are described in Table 2. Training was performed on only the training and validation dataset and testing dataset included a dataset of images which were not used while training. The predictions performed while testing of our developed weed detection models were performed on the testing dataset, so as to observe on how well our models were generalizing on unseen dataset. A detailed description of the evaluation metrics scores that were recorded after the DL models were developed in this study are presented in Table 4, this table depicts the performance comparisons between different weed detection models developed. IoU evaluation metric is considered to be the most important metric while comparing the performances of different SS based DL models developed. IoU scores gives us an estimate on how robustly a model is able to predict each and every pixel during predictions successfully [31].

Performance comparisons between the eight DL models developed shows that the LinkNet model with ResNet18 as feature extractor helped us achieve the best scores amongst all the models. These results are consistent with other similar works in this domain, [37] performed a study in which performance comparisons between LinkNet and UNet SS based DL models were made while developing a weed detection models in pulse crops. In comparison to their work, the proposed LinkNet model developed with ResNet18 architecture has shown an improvement of around 31% in the Mean IoU scores respectively.

The main objective of this study was to develop robust weed detection models that could perform the task of weed detections in maize fields and compare the inference time in which different models could perform the task of generating weed maps. Table 5 presents the inference time of different DL models in performing the predictions on both a single sub-image and on entire original image comprising of 384 sub-images. Complex issues such as crop-weed color similarity and crop-weed occlusions make development of these models quite hard. In order to further improve evaluation metric scores as obtained using previous DL techniques, we optimized the backpropagation techniques as discussed in Section 2.5.2 by using different loss functions. Several loss functions have already been developed in recent past to address the class-imbalance issues during training of DL models. We trained 2 DL models namely: LinkNet model with ResNet18 as a backbone and UNet model with ResNet18 as a backbone with several loss functions such as: Focal Loss [36], Dice Loss [38], Jaccard Loss [39], and Focal Dice Loss [40] and compared their performance by evaluating the Mean IoU and class wise IoU scores of these models. A detailed description of the mean IoU and IoU scores for each class category were recorded while performing predictions on the testing dataset by the weed detection models developed using different loss functions and are presented in Table 6. We observed that Focal loss function helps us improve the Mean IoU and IoU scores for classifying weeds during testing. These results suggests that the LinkNet model with ResNet18 backbone and Focal loss as a loss function leads to development of a robust weed detection model that achieves the best results.

Table 3.5 Performance comparison: Validation metric scores obtained after training different DL models for 50 epochs

		Validation metric scores				
Models	**Feature extractor**	**Precision**	**Recall**	***IoU***	**F1 score**	**Size of model/MB**
LinkNet	ResNet 18	**0.8651**	0.8649	**0.785**	0.865	**135**
	VGG 16	0.8031	0.7971	0.7473	0.81	238
	ResNet 34	0.8498	0.8545	0.762	0.8610	254
	InceptionV3	0.6737	0.6739	0.6875	0.678	309
UNet	ResNet 18	0.8553	**0.8767**	0.779	**0.8715**	168
	VGG 16	0.811	0.7821	0.7315	0.83	278
	ResNet 34	0.8608	0.8689	0.774	0.8700	287
	InceptionV3	0.6859	0.66	0.693	0.672	352

Table 3.6 Performance comparison of inference speed of LinkNet and UNet DL models trained on different backbones

DL model	**Feature extractor**	**Prediction time for a sub-image in FPS**	**Prediction time for an original image in FPS**
LinkNet	ResNet18	**67.1**	**0.175**
	VGG 16	54.3	0.136
	ResNet34	64.9	0.170
	Inception V3	42.5	0.112
UNet	ResNet18	59.9	0.155
	VGG 16	50.5	0.127
	ResNet34	57.2	0.136
	Inception V3	41.8	0.108

Table 3.7 Performance comparison between LinkNet-ResNet18 & UNet-ResNet18 model trained using different loss functions: Class wise and mean IoU scores during prediction

Model	**Loss function**	**IoU Background**	**IoU Weeds**	**IoU Soil**	**Mean IoU**
LinkNet-ResNet18	Categorical cross entropy	0.819	0.677	**0.859**	0.785
	Jaccard Loss	0.827	0.647	0.833	0.769
	Dice Loss	0.828	0.622	0.851	0.767
	Focal Loss	**0.868**	**0.691**	0.845	**0.801**
UNet-ResNet18	Categorical cross entropy	0.824	0.669	0.847	0.779
	Jaccard Loss	0.818	0.631	0.819	0.755
	Dice Loss	0.814	0.620	0.840	0.758
	Focal Loss	0.852	0.684	0.838	0.791

3.5 Future Research and Conclusions

Weeds pose a major threat in agriculture as they compete with crops for water, light, nutrients, and space. Weeds account for about 40% of global yield loss and cause a major threat to crop yield loss worldwide. Growing world population and increase in global food requirement create an alarming need to decrease crop yield. An effective way to handle this menace of crop yield losses due to weeds is developing site-specific weed management techniques. These techniques could help alleviate the crop yield losses by performing spot spraying practices of herbicides at only specific locations where weed infestations are detected leading to decrease in environmental pollution caused by overuse of herbicides. In this work, a study was conducted to develop a weed detection model using SS based DL techniques. Developing a weed detection model is a complex task as their exist multiple challenges such as crop-weed color similarity, crop-weed occlusions, presence of low weed infestations in fields that decrease the number of weed samples that are to be used for training DL models. Class imbalance problem is also a major challenge that needs to be addressed while developing these weed detection models due to over-sampling of other class categories as compared to weeds used for training. We used SS based DL techniques as these models help to perform pixel wise classifications during predictions, execute better scene understanding and in turn help generating better results as compared to other DL models.

We developed and investigated performance comparisons between two SS based DL models namely: LinkNet and UNet which comprise of an encoder-decoder architecture. A detailed explanation of SS based DL techniques and the two models used are discussed. We used different architecture backbones in the encoder sections and transfer learning techniques were used with fine tuning to develop these DL models. IoU score is an important evaluation metric that is used to compare and check the performance of SS based DL developed. Also, IoU metric helps to measure the generalizability of the developed model during testing on unseen dataset.

The novelty of this research lies in the techniques used to handle the class imbalance issues while developing weed detection models. Firstly, we performed an analysis of all the images included in the original dataset and created a refined dataset of images to be used for training. The images consisting of background pixel values of more than 95% were skipped and not used for training, this approach performs the task of balancing the data to be used for training. Thereafter, optimized backpropagation techniques to overcome the class imbalance issue were performed by training the weed detection models using different loss functions. We analyzed that Focal loss function in comparison with other loss functions helps to address the class imbalance problem by some extent and helps the model to train better on hard examples in our case ‘weeds’. Finally, we concluded that training a LinkNet SS model using ResNet18 in the encoder section along with ‘Focal Loss’ loss function helped us achieve state-of -art accuracy scores of Mean IoU score of 0.801 and weeds IoU score of 0.691 respectively. The developed model is quite robust and generalizes well on unseen UAV imagery of maize fields included in the test dataset.

Future work for this study is to further use these developed models on embedded devices to perform real-time Site-Specific Weed Management applications in agriculture fields. This could be achieved by optimizing these models using Software Development Toolkits like TensorRT and Tensorflow Lite as these helps to perform several important transformations and optimizations to the neural network graph, including constant folding, pruning unnecessary graph nodes, layer fusion for these Deep Learning models. This helps to decrease the computational and memory requirements that delivers low latency and high throughput for inference applications. Also, this research gives a basis to develop various weed detection models in different crops and also on detecting multiple weed categories in the fields.

3.6 References

1. Hao, S., Zhou, Y., & Guo, Y. (2020). A Brief Survey on Semantic Segmentation with Deep Learning. *Neurocomputing*, *406*, 302–321. https://doi.org/10.1016/j.neucom.2019.11.118
2. Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., & Garcia-Rodriguez, J. (2017). *A Review on Deep Learning Techniques Applied to Semantic Segmentation*. 1–23. http://arxiv.org/abs/1704.06857
3. Ramos, S., Gehrig, S., Pinggera, P., Franke, U., & Rother, C. (2017, June). Detecting unexpected obstacles for self-driving cars: Fusing deep learning and

geometric modeling. In 2017 IEEE Intelligent Vehicles Symposium (IV) (pp. 1025-1032). IEEE.

4. Treml, M., Arjona-Medina, J., Unterthiner, T., Durgesh, R., Friedmann, F., Schuberth, P., ... & Hochreiter, S. (2016). Speeding up semantic segmentation for autonomous driving.
5. Badue, C., Guidolini, R., Carneiro, R. V., Azevedo, P., Cardoso, V. B., Forechi, A., ... & De Souza, A. F. (2021). Self-driving cars: A survey. Expert Systems with Applications, 165, 113816.
6. Azuma, R. T. (1997). A survey of augmented reality. Presence: teleoperators & virtual environments, 6(4), 355-385.
7. Billinghurst, M., Clark, A., & Lee, G. (2015). A survey of augmented reality. Foundations and Trends® in Human–Computer Interaction, 8(2-3), 73-272.
8. Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., ... & Asari, V. K. (2018). The history began from alexnet: A comprehensive survey on deep learning approaches. arXiv preprint arXiv:1803.01164. https://doi.org/10.48550/arXiv.1803.01164
9. Girshick, R. (2015). Fast r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 1440-1448). https://doi.org/10.1109/ICCV.2015.169
10. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, 28. https://doi.org/10.1109/tpami.2016.2577031

11. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788). https://doi.org/10.1109/CVPR.2016.91

12. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016, October). Ssd: Single shot multibox detector. In European conference on computer vision (pp. 21-37). Springer, Cham. https://doi.org/10.1007/978-3-319-46448-0_2

13. Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE transactions on pattern analysis and machine intelligence, 39(12), 2481-2495. https://doi.org/10.1109/TPAMI.2016.2644615

14. Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham. https://doi.org/10.48550/arXiv.1505.04597

15. Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). Pyramid scene parsing network. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2881-2890). https://doi.org/10.1109/CVPR.2017.660

16. Chaurasia, A., & Culurciello, E. (2017, December). Linknet: Exploiting encoder representations for efficient semantic segmentation. In 2017 IEEE Visual

Communications and Image Processing (VCIP) (pp. 1-4). IEEE. https://doi.org10.1109/VCIP.2017.8305148

17. Yu, C., Wang, J., Peng, C., Gao, C., Yu, G., & Sang, N. (2018). Bisenet: Bilateral segmentation network for real-time semantic segmentation. In Proceedings of the European conference on computer vision (ECCV) (pp. 325-341). https://doi.org/10.48550/arXiv.1808.00897
18. He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 2961-2969). https://doi.org/10.48550/arXiv.1703.06870
19. Kirillov, A., He, K., Girshick, R., Rother, C., & Dollár, P. (2019). Panoptic segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 9404-9413).
20. Bolya, D., Zhou, C., Xiao, F., & Lee, Y. J. (2019). Yolact: Real-time instance segmentation. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 9157-9166).
21. Bosilj, P., Aptoula, E., Duckett, T., & Cielniak, G. (2020). Transfer learning between crop types for semantic segmentation of crops versus weeds in precision agriculture. Journal of Field Robotics, 37(1), 7-19.
22. Maheswari, P., Raja, P., Apolo-Apolo, O. E., & Pérez-Ruiz, M. (2021). Intelligent fruit yield estimation for orchards using deep learning based semantic segmentation techniques—a review. Frontiers in Plant Science, 12, 684328.

23. Singh, N., Tewari, V. K., Biswas, P. K., Dhruw, L. K., Pareek, C. M., & Singh, H. D. (2022). Semantic segmentation of in-field cotton bolls from the sky using deep convolutional neural networks. Smart Agricultural Technology, 2, 100045.
24. Veeranampalayam Sivakumar, A. N., Li, J., Scott, S., Psota, E., J Jhala, A., Luck, J. D., & Shi, Y. (2020). Comparison of object detection and patch-based classification deep learning models on mid-to late-season weed detection in UAV imagery. Remote Sensing, 12(13), 2136. https://doi.org/10.3390/rs12132136
25. Alzadjali, A., Alali, M. H., Veeranampalayam Sivakumar, A. N., Deogun, J. S., Scott, S., Schnable, J. C., & Shi, Y. (2021). Maize tassel detection from uav imagery using deep learning. Frontiers in Robotics and AI, 136.
26. Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. Computers and electronics in agriculture, 145, 311-318.
27. Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., Iwamoto, Y., ... & Wu, J. (2020, May). Unet 3+: A full-scale connected unet for medical image segmentation. In ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 1055-1059). IEEE. https://doi.org/10.1109/ICASSP40776.2020.9053405
28. Brostow, G. J., Fauqueur, J., & Cipolla, R. (2009). Semantic object classes in video: A high-definition ground truth database. Pattern Recognition Letters, 30(2), 88-97.
29. Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., ... & Schiele, B. (2016). The cityscapes dataset for semantic urban scene

understanding. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3213-3223).

30. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. Communications of the ACM, 60(6), 84-90.
31. Asad, M. H., & Bais, A. (2020). Weed detection in canola fields using maximum likelihood classification and deep convolutional neural network. Information Processing in Agriculture, 7(4), 535-545.
32. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. https://doi.org/10.48550/arXiv.1409.1556
33. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778). https://doi.org/10.48550/arXiv.1512.03385
34. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2818-2826). https://doi.org/10.48550/arXiv.1512.00567
35. Zaidi, S. S. A., Ansari, M. S., Aslam, A., Kanwal, N., Asghar, M., & Lee, B. (2022). A survey of modern deep learning based object detection models. Digital Signal Processing, 103514. https://doi.org/10.1016/j.dsp.2022.103514

36. Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In Proceedings of the IEEE international conference on computer vision (pp. 2980-2988). https://doi.org/10.48550/arXiv.1708.0200

37. Revanasiddappa, B., Arvind, C., & Swamy, S. (2020). Real-time early detection of weed plants in pulse crop field using drone with IoT. Technology, 16(5), 1227-1242.

38. Li, X., Sun, X., Meng, Y., Liang, J., Wu, F., & Li, J. (2019). Dice loss for data-imbalanced NLP tasks. arXiv preprint arXiv:1911.02855.

39. Berman, M., Triki, A. R., & Blaschko, M. B. (2018). The lovász-softmax loss: A tractable surrogate for the optimization of the intersection-over-union measure in neural networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4413-4421).

40. Wang, P., & Chung, A. (2018). Focal dice loss and image dilation for brain tumor segmentation. In Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support (pp. 119-127). Springer, Cham.

41. Espejo-Garcia, B., Mylonas, N., Athanasakos, L., Fountas, S., & Vasilakoglou, I. (2020). Towards weeds identification assistance through transfer learning. Computers and Electronics in Agriculture, 171, 105306.

42. Lamb, D. W., & Brown, R. B. (2001). Pa—precision agriculture: Remote-sensing and mapping of weeds in crops. Journal of Agricultural Engineering Research, 78(2), 117-125. https://doi.org/10.1006/jaer.2000.0630

43. Forero, M. G., Herrera-Rivera, S., Ávila-Navarro, J., Franco, C. A., Rasmussen, J., & Nielsen, J. (2018, November). Color classification methods for perennial weed detection in cereal crops. In Iberoamerican Congress on Pattern Recognition (pp. 117-123). Springer, Cham. https://doi.org/10.1007/978-3-030-13469-3_14

44. Feyaerts, F., & Van Gool, L. (2001). Multi-spectral vision system for weed detection. Pattern Recognition Letters, 22(6-7), 667-674. https://doi.org/10.1016/S0167-8655(01)00006-X

45. Vrindts, E., De Baerdemaeker, J., & Ramon, H. (2002). Weed detection using canopy reflection. Precision agriculture, 3(1), 63-80.

46. Jaware, T. H., Badgujar, R. D., & Patil, P. G. (2012). Crop disease detection using image segmentation. World Journal of Science and Technology, 2(4), 190-194.

47. Van Klompenburg, T., Kassahun, A., & Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review. Computers and Electronics in Agriculture, 177, 105709.

48. López-Granados, F. (2011). Weed detection for site-specific weed management: mapping and real-time approaches. Weed Research, 51(1), 1-11.

49. Kong, J., Wang, H., Wang, X., Jin, X., Fang, X., & Lin, S. (2021). Multi-stream hybrid architecture based on cross-level fusion strategy for fine-grained crop species recognition in precision agriculture. Computers and Electronics in Agriculture, 185, 106134.

50. Makowski, D., Naud, C., Jeuffroy, M. H., Barbottin, A., & Monod, H. (2006). Global sensitivity analysis for calculating the contribution of genetic parameters

to the variance of crop model prediction. Reliability Engineering & System Safety, 91(10-11), 1142-1147.

51. Oliveira, D. A. B., Pereira, L. G. R., Bresolin, T., Ferreira, R. E. P., & Dorea, J. R. R. (2021). A review of deep learning algorithms for computer vision systems in livestock. Livestock Science, 253, 104700.
52. Wang, N. I. N. G., Zhang, N., Dowell, F. E., Sun, Y., & Peterson, D. E. (2001). Design of an optical weed sensor usingplant spectral characteristics. Transactions of the ASAE, 44(2), 409. https://doi.org/10.13031/2013.4673
53. Bah, M. D., Hafiane, A., & Canals, R. (2017, November). Weeds detection in UAV imagery using SLIC and the hough transform. In 2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA) (pp. 1-6). IEEE. https://doi.org/10.1109/IPTA.2017.8310102
54. Sankaran, S., Khot, L. R., Espinoza, C. Z., Jarolmasjed, S., Sathuvalli, V. R., Vandemark, G. J., ... & Pavek, M. J. (2015). Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: A review. European Journal of Agronomy, 70, 112-123. https://doi.org/10.1016/j.eja.2015.07.004
55. Bakhshipour, A., & Jafari, A. (2018). Evaluation of support vector machine and artificial neural networks in weed detection using shape features. Computers and Electronics in Agriculture, 145, 153-160. https://doi.org/10.1016/j.compag.2017.12.032
56. Chauhan, B. S. (2020). Grand challenges in weed management. Frontiers in Agronomy, 1, 3.

57. Adeux, G., Vieren, E., Carlesi, S., Bàrberi, P., Munier-Jolain, N., & Cordeau, S. (2019). Mitigating crop yield losses through weed diversity. Nature Sustainability, 2(11), 1018-1026. https://doi.org/10.1038/s41893-019-0415-y

58. Martinez, D. A., Loening, U. E., & Graham, M. C. (2018). Impacts of glyphosate-based herbicides on disease resistance and health of crops: a review. Environmental Sciences Europe, 30(1), 1-14.

59. El-Faki, M. S., Zhang, N., & Peterson, D. E. (2000). Weed detection using color machine vision. Transactions of the ASAE, 43(6), 1969. https://doi.org/10.13031/2013.3103

60. Vrindts, E., De Baerdemaeker, J., & Ramon, H. (2002). Weed detection using canopy reflection. Precision agriculture, 3(1), 63-80. https://doi.org/10.1023/A:1013326304427

61. Perez, A. J., Lopez, F., Benlloch, J. V., & Christensen, S. (2000). Colour and shape analysis techniques for weed detection in cereal fields. Computers and electronics in agriculture, 25(3), 197-212. https://doi.org/10.1016/S0168-1699(99)00068-X

62. Asif, M., Amir, S., Israr, A., & Faraz, M. (2010). A vision system for autonomous weed detection robot. International Journal of Computer and Electrical Engineering, 2(3), 486.

63. Rani, S. V., Kumar, P. S., Priyadharsini, R., Srividya, S. J., & Harshana, S. (2022). An automated weed detection system in smart farming for developing

sustainable agriculture. International Journal of Environmental Science and Technology, 19(9), 9083-9094. https://doi.org/10.1007/s13762-021-03606-6

64. Yano, I. H., Mesa, N. F. O., Santiago, W. E., Aguiar, R. H., & Teruel, B. (2017). Weed identification in sugarcane plantations through images taken from remotely piloted aircraft (RPA) and KNN classifier. J. Food Nutr. Sci, 5, 211. https://doi.org/10.11648/j.jfns.20170506.11

65. Pantazi, X. E., Tamouridou, A. A., Alexandridis, T. K., Lagopodi, A. L., Kashefi, J., & Moshou, D. (2017). Evaluation of hierarchical self-organising maps for weed mapping using UAS multispectral imagery. Computers and Electronics in Agriculture, 139, 224-230.

66. Sheffield, K. J., Clements, D., Clune, D. J., Constantine, A., & Dugdale, T. M. (2022). Detection of Aquatic Alligator Weed (Alternanthera philoxeroides) from Aerial Imagery Using Random Forest Classification. Remote Sensing, 14(11), 2674. https://doi.org/10.3390/rs14112674

Chapter 4: Conclusions

Several studies have already been conducted to develop precision agriculture applications using computer vision techniques in the recent past. In this thesis work, an intensive literature review was conducted to understand various computer vision techniques that have been used to generate robust weed detection models and determine the research gaps that are present in the existing research. Semantic segmentation based deep learning-based techniques were used to train the weed detection models to train a powerful weed detection model that could be used to generate weed maps on drone imagery collected from maize fields.

Even though numerous studies have already been conducted that review various computer vision-based techniques to develop weed detection models. We present a detailed literature review on various approaches that have been used to develop weed detection models in the past few decades. Traditional image processing, machine vision, machine learning and deep learning techniques were explained in detail and their associated advantages and disadvantages were discussed. Also, the techniques used to perform real-time Site-Specific Weed Management applications were discussed. The role of cloud computing and edge computing in developing these real-time applications and their associated advantages and disadvantages were discussed. Lastly, conclusions and some prospects that could be used to perform future studies were described in the final section of the study.

The advancements in Remote Sensing techniques to fetch high quality data using Unmanned Aerial Vehicles have helped in the development of robust weed detection models using deep learning techniques. In this study, we developed a Semantic Segmentation based Deep Learning model to perform the complex task of weed detection in maize fields. These Semantic Segmentation models perform the tasks of pixel-wise labelling of detected objects during predictions which help in developing robust Deep Learning models. Also, these models help to perform tasks like better scene understanding and learn to extract useful features to learn the complex trends of color similarity, inter-intra line weed infestations, high occlusions between crop and weeds in maize fields. The goal of this study was to develop a robust weed detection model using Semantic Segmentation based Deep Learning techniques and address the class imbalance problem. The developed model was expected to generate accurate weed maps on UAV imagery collected from maize fields.

To perform this study, we made performance comparisons between weed detection models developed using two Semantic Segmentation models namely LinkNet and UNet. Transfer learning techniques were used to perform the feature extraction process and these features were used for training weed detection models. The datasets that are frequently used to train deep learning models in weed detection domain usually have a class-imbalance issue that negatively affects the models in robustly identifying weed patches in the fields. To overcome this issue, we performed a two-step approach: Firstly, we refined the dataset to be used for training and developed a balanced dataset to be sued for training. Secondly, we performed optimized backpropagation techniques and made performance comparisons between models trained using different loss functions. We concluded that, the weed detection model trained on LinkNet model with ResNet18 as a feature extractor and ‘Focal Loss’ as a loss function helped us achieve the best Intersection Over Union scores which is an important evaluation metric to compare Semantic Segmentation based Deep Learning models. Also, this model had the minimum inference time amongst others which signifies that this model could be used to perform real-time weeding applications for Site-Specific Weed Management applications. The model was able to achieve a Mean IoU score of 0.801 and IoU score for weed class category of 0.691.

Future work for this study would be to further use these developed models on embedded devices to check their performance for real-time Site-Specific Weed Management applications in agriculture fields. This could be achieved by optimizing these models using Software Development Toolkits like TensorRT and Tensorflow Lite as these helps to perform several important transformations and optimizations to the neural network graph, including constant folding, pruning unnecessary graph nodes, layer fusion for these Deep Learning models. This helps to decrease the computational and memory requirements that delivers low latency and high throughput for inference applications.