Available at www.sciencedirect.com

INFORMATION PROCESSING IN AGRICULTURE 7 (2020) 535–545

journal homepage: www.elsevier.com/locate/inpa

Weed detection in canola fields using maximum likelihood classification and deep convolutional neural network

Muhammad Hamza Asad, Abdul Bais *

Electronic Systems Engineering, Faculty of Engineering and Applied Science, University of Regina, Saskatchewan, Canada

ARTICLE INFO

Article history:
Received 28 August 2019
Received in revised form
26 November 2019
Accepted 19 December 2019
Available online 24 December 2019

Keywords:
Weed detection
Semantic segmentation
Variable rate herbicide
Maximum likelihood classification

ABSTRACT

Herbicide use is rising globally to enhance food production, causing harm to environment and the ecosystem. Precision agriculture suggests variable-rate herbicide application based on weed densities to mitigate adverse effects of herbicides. Accurate weed density estimation using advanced computer vision techniques like deep learning requires large labelled agriculture data. Labelling large agriculture data at pixel level is a time-consuming and tedious job. In this paper, a methodology is developed to accelerate manual labelling of pixels using a two-step procedure. In the first step, the background and foreground are segmented using maximum likelihood classification, and in the second step, the weed pixels are manually labelled. Such labelled data is used to train semantic segmentation models, which classify crop and background pixels as one class, and all other vegetation as the second class. This paper evaluates the proposed methodology on high-resolution colour images of canola fields and makes performance comparison of deep learning meta-architectures like SegNet and UNET and encoder blocks like VGG16 and ResNet-50. ResNet-50 based SegNet model has shown the best results with mean intersection over union value of 0.8288 and frequency weighted intersection over union value of 0.9869.

1. Introduction

According to estimates, crops lose 20–80% of production to diseases, pests, and undesired weeds [1]. Use of chemicals is a standard agricultural practice to increase crop yield. Despite changing chemical composition and modern application techniques, the overall global trend of chemical application to crops is still on the rise [2]. Today, herbicides constitute two-thirds of the total chemical application to agriculture fields in USA, increasing from less than a quarter in 1964 [3]. There are growing environmental and biological concerns for wholesale application of chemicals to crops. Recent studies have shown that glyphosate (a common herbicide used since 1974) has carcinogenic toxicities for human beings [4]. It necessitates that a balanced approach must be adopted while applying herbicides to crop fields.

In conventional agriculture practices, herbicides are applied at uniform rate to the whole field even though weed distribution is patchy. It results in high input costs for farmers and also causes environmental pollution [5]. Contrary to this, Precision Agriculture (PA) aims need-based site-specific application [6]. Adopting PA practices for use of herbicides requires accurate weed mapping by classifying weeds and host plants.

* Corresponding author.
E-mail addresses: maq541@uregina.ca (M.H. Asad), abdul.bais@uregina.ca (A. Bais).
Peer review under responsibility of China Agricultural University.
https://doi.org/10.1016/j.inpa.2019.12.002

Classification of plants is a challenging task because of spectral similarities between different plant types. Existing field mapping methods assume that host plants are seeded in rows. Line detection techniques are used to classify plants in a row as host plants, and plants falling out of seeding lines as weeds. This interline approach fails to detect weed plants located within crop lines and host plants falling out of crop lines. To address these issues, Hemming et al. employ intra-line approach for classification of weeds and host plants [7].

Deep learning has simplified object detection due to its ability to extract features from images automatically. There are two categories of methods found in deep learning for object detection. The first is drawing bounding boxes around images, and the second is classifying object pixels. From labelling point of view, drawing rectangular bounding boxes around objects is easier than labelling object pixels by drawing contours. However, from mapping perspective detecting objects at pixel level is more accurate than bounding box technique. Pixel wise classification of objects named semantic segmentation is not fully exploited in agriculture images due to unavailability of labelled images at pixel level. In this paper, the objective is to map weeds to estimate weed densities for variable rate herbicide application. For this purpose, we apply semantic segmentation on high-resolution imagery collected from canola fields. We train semantic segmentation models on large real data for better generalisation under real field conditions [8]. This paper has the following contributions:

1. Weed detection and weed mapping in canola field using semantic segmentation.
2. Two-step approach for manual labelling of weeds in agricultural images.
3. Performance comparison of VGG16 and ResNet-50 as feature extractors and UNET and SegNet as meta-architectures on canola field imagery.

The proposed methodology has shown promising results in detecting weeds. The remainder of the paper is organised as follows: Section 2 presents related work, Section 3 describes the methodology, Section 4 is about result discussion, and Section 5 is conclusion and recommendations for future work.

2. Related work

Weed classification and detection involve the following four steps [9]:

1. Image Acquisition: Red, Green and Blue (RGB) or multi-spectral images are collected using ordinary camera or special sensors.
2. Image Segmentation: Foreground and background are separated based on pixel values. Normalized Difference Vegetation Index and edge detection based techniques are commonly used for segmentation.
3. Feature Extraction: After segmentation, shape and skeleton-based features are extracted from plants.
4. Image Classification: The images are classified based on extracted features if they have weeds or not.

First step in weed detection is image acquisition. Agricultural images are acquired through sensors mounted on four types of platforms. These are satellites, planes, Unmanned Aerial Vehicles (UAVs), and ground moving field vehicles (e.g., quad, sprayers). High-resolution imagery through satellite and plane is expensive and depend upon weather conditions. UAV and field vehicles based imagery is widely adopted due to its high spatial resolution, operational flexibility and economic affordability compared to the former two sources of agriculture images [10]. Second procedural step is segmentation of images by separating vegetation and background soil, stone/rocks and dead plants. Spectral based segmentation techniques are used mostly as vegetation has different RGB and near infra-red ranges compared to soil and dead plants [11]. To minimise the effect of varying conditions, Bah et al. use Otsu-adaptive thresholding for background segmentation [12]. Other segmentation techniques involve clustering algorithms and principle component analysis [13–15].

There are two main strategies for weed detection in the segmented images. One is interline, and the other is intra-line approach. Interline approach assumes that most of the crop plants are in lines [16]. All the vegetation outside those lines is weeds. Whereas, intra-line approach detects weeds both inside and outside the host plant lines [17]. Bah et al. use normalized Hough transform for detection of host plants in crop lines. They classify any interline vegetation as weeds [12]. Another crop line detection method involves using geometric constraints for mapping regions of interest [18]. Intra-line approaches attempt to classify weeds and crops irrespective of their position in the field. Due to identical spectral signature of weeds and host plants, spectral features are insufficient to distinguish between the two. Therefore, additional features like shape and texture are extracted from high resolution images for classification [19]. Okamoto et al. use wavelet transform for feature extraction [20]. Deep learning has made feature extraction easier by extracting features out of data automatically [21].

Machine learning methods are normally applied for the final step of classification. Examples are Bayesian classification, support vector machines, self-organising maps, random forest algorithms, and artificial neural networks [22–26]. With the advent of deep learning techniques, feature extraction step and classification step are merged. Convolutional Neural Networks (CNN) perform feature extraction automatically before classification and detection tasks. Olsen et al. use off the shelf deep learning architectures to classify images of sixteen different types of weeds [27]. Contrary to classification of individual weed images, weed density estimation requires semantic segmentation [28]. The challenge of this method is manual labelling of data at pixel level, which is time-consuming. To overcome this challenge, Dryrmann et al. synthesize data by putting different weeds close to host plant for training semantic segmentation models [29]. Semantic model trained on synthesized data does not generalize well to real scenarios. Potenza et al. perform multilevel classification on

real dataset. In the first level, background and vegetation are separated, and in the second level of classification, weeds and crops are separated. They train a neural network using a small number of representative images to label the rest of the dataset. They then retrain their models using the entire dataset. The bottleneck of this technique is pixel labelling for generating labels for the rest of the data. The number of images is not large enough to build a general model [30]. Moreover, errors of this step propagate to the second step.

3. Methodology

In this paper, we use semantic segmentation for accurate mapping of weeds. Manual labelling of pixels is a major bottleneck in agriculture images for semantic segmentation. We develop a novel two-step method to label data manually. Procedural steps of the proposed methodology are shown in Fig. 1. Rest of this section provides details of each step.

3.1. Data description

High resolution RGB images of canola fields at two different growth stages of crop are selected for this study. The geo-tagged images are acquired with a Quad mounted Camera (Nikon D610) from canola fields in Manitoba Canada. The dataset consists of 906 images which are manually labelled into two classes at pixel level. We use different augmentation techniques to add effects in the images which may occur in the actual environment like light and contrast variations, blurring of images due to equipment vibrations and changes in crop lines orientations. The techniques used to enhance the dataset are horizontal and vertical flips, rotation, colour variations, shear, scaling, cropping, blurring and zero padding. These are applied in random combinations. Fig. 2b to f show examples of data augmentation applied to Fig. 2a. The original image dimension is 4288 × 2848. Due to memory constraint, these images are split into tiles of size 1440 × 960.

3.2. Two-step manual labelling

In the data preprocessing step of weed detection, we propose a two-step manual labelling process. In the first step, vegetation is separated from background soil, stones and dead plants and in the second step, minority class pixels are labelled on background-segmented image through a labelling tool. Background removal helps efficiently label background pixels. Also, it reduces the time for manual labelling of minority class pixels. With background, there are chances to miss some of minority class pixels in manual labelling process. We are segmenting background for the sake of constructing labels. The original images and corresponding labels are used for training. For background segmentation, we employed different rule-based colour thresholding schemes, but these did not work for images with high colour variation, changing lighting condition and shadows. We use statistical technique named Maximum Likelihood Classification (MLC) [31] which has demonstrated better performance than thresholding techniques.

Prior to background segmentation, similar geo-tagged images are stitched together to make a mosaic. Images are simply stacked if there is no overlap between them. Thereafter, training samples are taken from this larger image which are representative of both classes. Spectral signatures are calculated for these training samples. MLC is performed based on the two principles: Bayes' theorem and the assumption that class samples are normally distributed [31]. If c_i represents class i and v represents the feature vector then according to Bayes' theorem [32]:

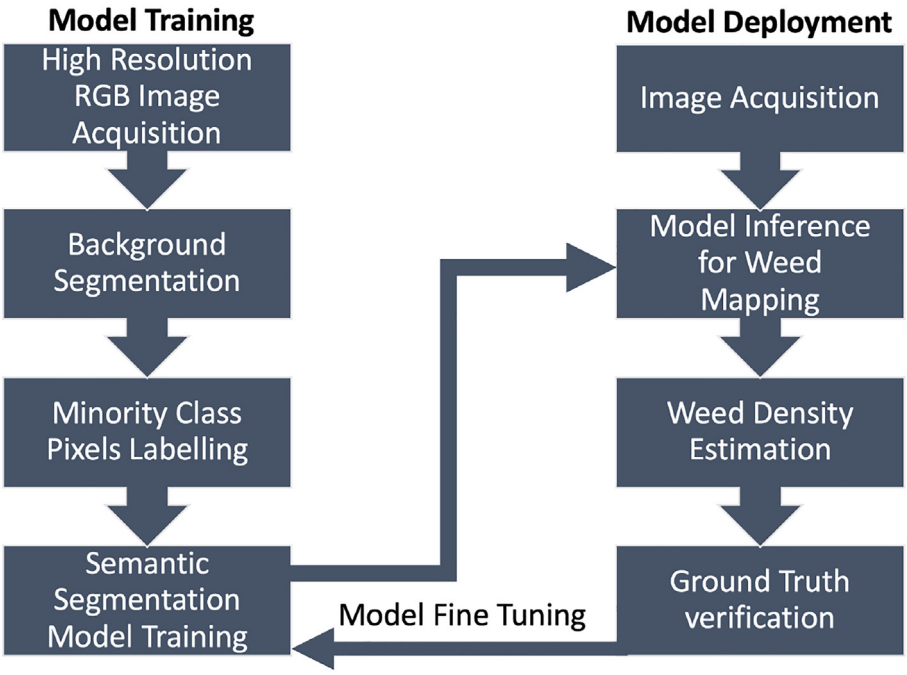


Fig. 1 – Flow diagram for proposed weed detection methodology.

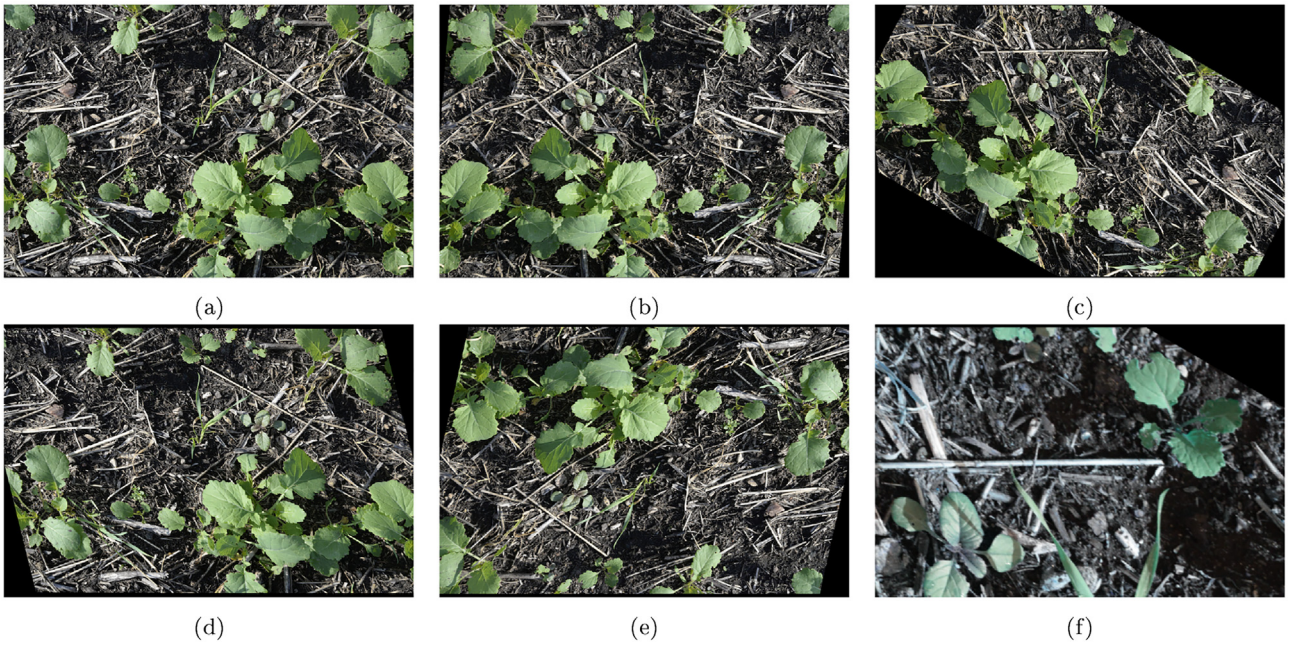


Fig. 2 – (a) Is the original image and (b-f) are different combinations of flipping, shearing, rotation, zero-padding and colour variations applied to (a).

$$P(c_i \mid \boldsymbol{v}) = \frac{P(\boldsymbol{v} \mid c_i) \times P(c_i)}{P(\boldsymbol{v})} \quad (1)$$

The feature vector υ belongs to c_i if the probability given by (1) is greatest among all classes. For the binary classification we can write it as [32]:

$$P(c_i \mid \boldsymbol{v}) > P(c_j \mid \boldsymbol{v}) \quad (2)$$

where:

$$P(\boldsymbol{v}) = \sum_{i=1}^{N} P(\boldsymbol{v} \mid c_i) \times P(c_i) \quad (3)$$

Substituting (1) in (2), we obtain [32]:

$$P(\boldsymbol{v} \mid c_i) \times P(c_i) > P(\boldsymbol{v} \mid c_j) \times P(c_j) \quad (4)$$

where $P(\boldsymbol{v}|c_i)$ is likelihood probability. In our case, the class probability is same for both classes. As $\boldsymbol{v}$ is normally distributed in all spectral bands, $P(\boldsymbol{v}|c_i)$ can be given by the following relationship [32]:

$$P(\boldsymbol{v} \mid c_i) = (2\pi)^{-0.5} \times |\mathbf{Y}_i|^{0.5} \times exp(-\frac{1}{2}(\boldsymbol{v} - \mathbf{z}_i)^T) \times \mathbf{Y}_i^{-1}(\boldsymbol{v} - \mathbf{z}_i) \quad (5)$$

$\mathbf{Y}_i$ and $\mathbf{z}_i$ are the co-variance matrix and mean value vector which are known from training samples.

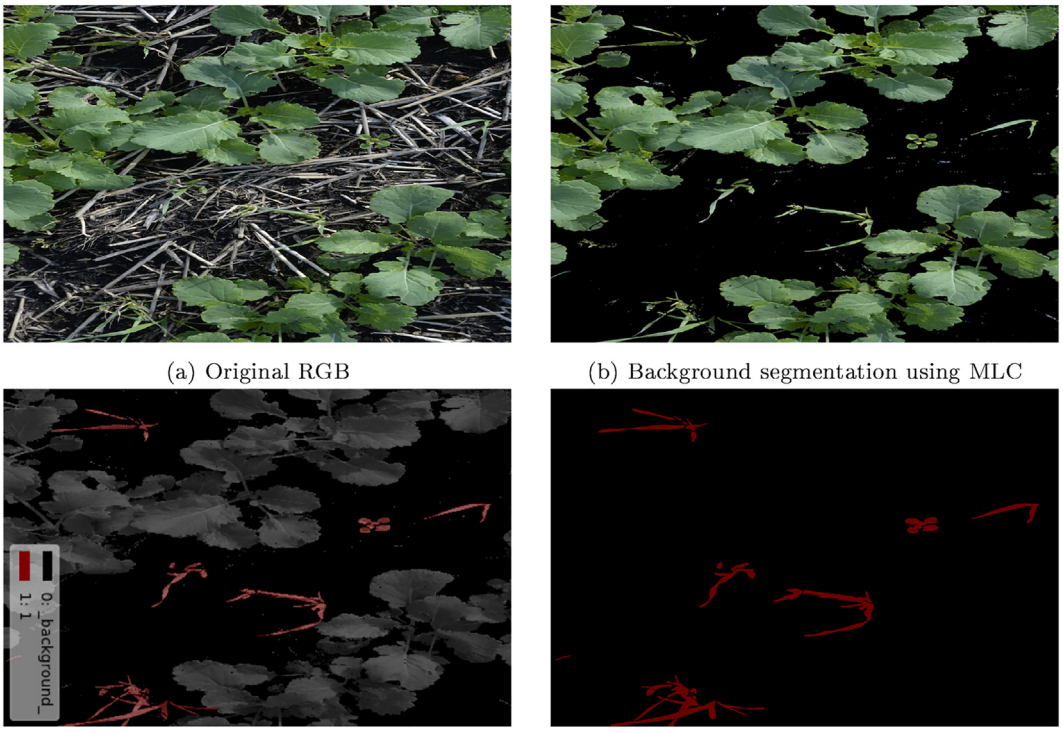
MLC is performing equally good for weedy and non-weedy images. We also applied MLC for separating weeds and host plants, but it did not work due to spectral similarity. In the second step of manual labelling, weeds are labelled in background segmented images. If the image contains both weeds and crop plants, we use labelling tool to only label minority class pixels which are weeds in this case. Polygons are drawn around weed pixels to mark exact boundaries of weed and non-weed pixels. By adopting this two-step procedure, time for manual labelling is significantly decreased. Fig. 3 shows the procedure of creating labels. The background segmented image is shown in Fig. 3a. Weeds in this image are labelled using the Labelme tool (Fig. 3c). The final labels are shown in Fig. 3d. Fig. 3a and s make the image-label pair for the network.

3.3. *Semantic segmentation*

We use semantic segmentation for weed detection and mapping. Deep learning based semantic segmentation consists of two main blocks, encoding and decoding blocks. Encoding block is a downsampling block that extracts features out of images whereas decoding block up sample feature space to image dimensions. Semantic segmentation networks like UNET [33] and SegNet [34] employ fully convolutional networks whose top layers are also convolutional instead of fully connected dense layers as in typical CNNs. Fig. 4 shows semantic segmentation arrangement.

3.3.1. *Feature extraction*

CNN has revolutionised image classification and segmentation due to its ability to automatically extract features. A typical CNN architecture consists of alternating convolutional and max-pooling layers. Convolutional layer extracts features out of images while max-pooling layers reduce the dimension of feature space. After these layers, there are fully connected layers, and topmost layer has usually softmax as an activation function for multi-class classification. Addition of more layers increases the ability to learn complex features in the image, but it also increases the complexity of the model and requires more computational power. Numerous CNN architectures have been developed like VGG16 [35], ResNet-50



(a) Original RGB (b) Background segmentation using MLC

(c) Manually labelled weeds (d) Label for model training

Fig. 3 – (a–d) Explains two-step manual labelling process.

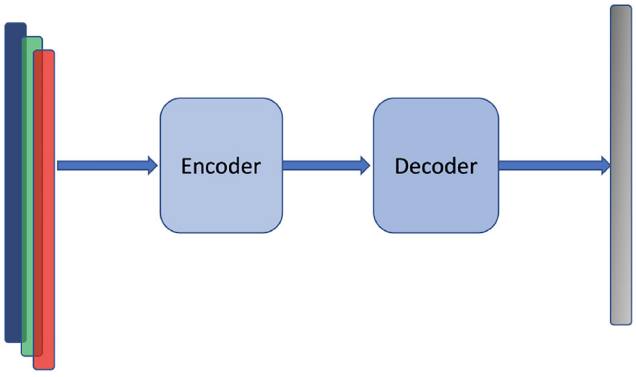


Fig. 4 – Semantic segmentation architecture.

[36], Inception V3 [37] and Xception [38]. In this paper, we are using VGG16 and ResNet-50 for feature extraction. The feature extractor is the encoder block in the meta-architecture.

VGG16 uses simple 3 × 3 size kernels in convolutional layers and 2 × 2 size in max pooling layers. There are two fully connected layers of size 4096 but in semantic segmentation meta architectures these layers are stripped and replaced with fully convolutional layers. Rectified Linear Unit (ReLU) is commonly used activation function in VGG16. Suffix 16 in VGG16 shows the number of weight layers. Fig. 5 shows the VGG16 architecture.

In ResNet-50, instead of using sequential layers as in VGG16, micro-architectures are used where residues flow forward in the layers. Deeper networks face the problem of vanishing gradients and they do not converge well. To address this problem, an identity function skip connection is added in ResNet that passes the information from previous layer to next layer [36]. Fig. 6 shows a typical residue-based micro-architecture with skip connections.

ResNet-50 models can be more deeper than VGG models but require less space for weights storage due to global pooling layer instead of fully connected layers.

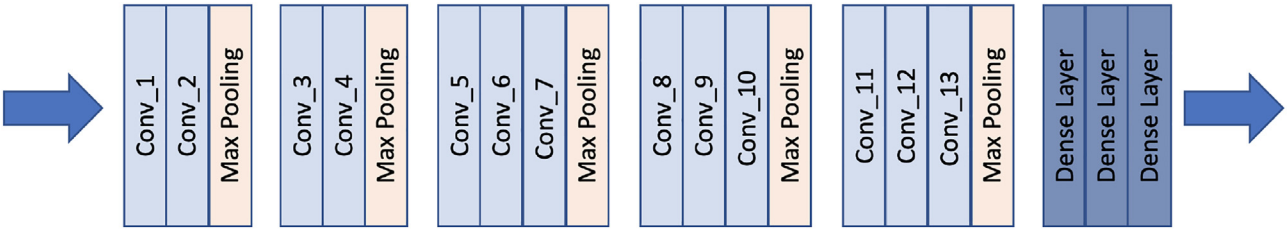


Fig. 5 – VGG16 architecture.

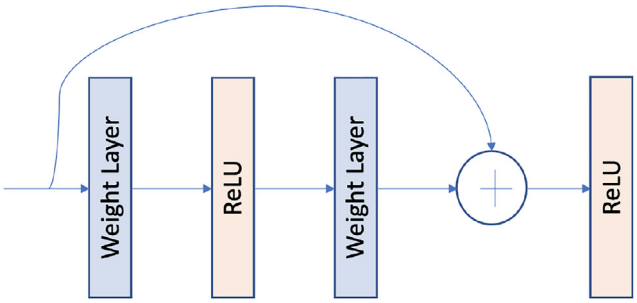


Fig. 6 – Residue based micro structure in a simple ResNet.

3.3.2. *SegNet*

SegNet is based on the concept of Fully Convolutional Network (FCN). In FCN, fully connected top layers are also replaced with convolutional layers [28,39]. SegNet has staged upsampling decoder. Deconvolutional layers are transpose of convolutional layers. Indexes of max-pooling layer are transferred from encoding blocks to decoding blocks to help decoding blocks upsample. SegNet performs well on small datasets [40]. Fig. 7 shows network architecture of SegNet.

3.3.3. *UNET*

UNET is also a FCN. Unlike SegNet, in UNET whole feature map is transferred from encoder block to decoder block which requires more memory. The skip connections between encoder and decoder blocks perform concatenation operation. The objective of concatenating skip connections is to combine local information extracted from encoding block with global spatial information. UNET also performs well on small datasets [41]. Fig. 8 refers to typical UNET architecture.

4. Results discussion and analysis

The proposed methodology is evaluated on RGB images described in Section 3.1. Preprocessing of images like background segmentation and manual labelling are performed using ARCGIS [42] and Labelme [43] tool box respectively. The dataset is divided into training and test data of sizes 85% and 15% respectively. After train-test split, images are augmented and split into smaller tiles for training. To speed up the model training, NVIDIA GPU GeForce GTX 1080 Ti [44] with 11 GB memory is used.

4.1. *Evaluation metrics*

In semantic segmentation models simple accuracy, precision and recall score do not give full picture due to high class imbalance [45]. To address this problem, we are using Intersection Over Union (IOU), Mean Intersection Over Union (MIOU) and Frequency Weighted Intersection Over Union (FWIOU). IOU is measure of overlap of each predicted class with ground truth. MIOU is average of all class IOUs. FWIOU calculates weighted average of IOUs based on pixel classes. IOU, MIOU and FWIOU are given by following equations:

$$IOU_i = \frac{\text{Area of Overlap}}{\text{Area of Union}} \tag{7}$$

where IOU_i is intersection over union for pixel class i.

$$MIOU = \frac{IOU_i + IOU_j}{k} \tag{8}$$

where i and j are pixel classes and k is number of classes which is 2 in our case.

$$FWIOU = w_i \times IOU_i + w_j \times IOU_j \tag{9}$$

where w_i and w_j are the weights of each class.

4.2. *Results discussion*

Four semantic segmentation models of UNET and SegNet each with VGG16 and ResNet-50 as base models are trained. The training data is further split into training and validation sets of 85% and 15% respectively. All models are trained for 100 epochs with Adadelta [46] as optimizer. Fig. 9 shows the

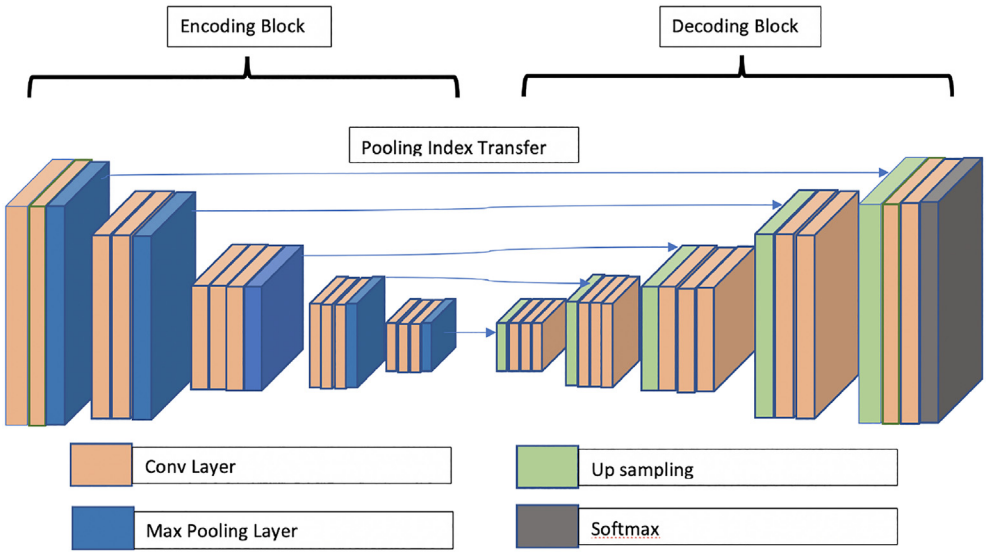


Fig. 7 – SegNet meta-architecture [34].

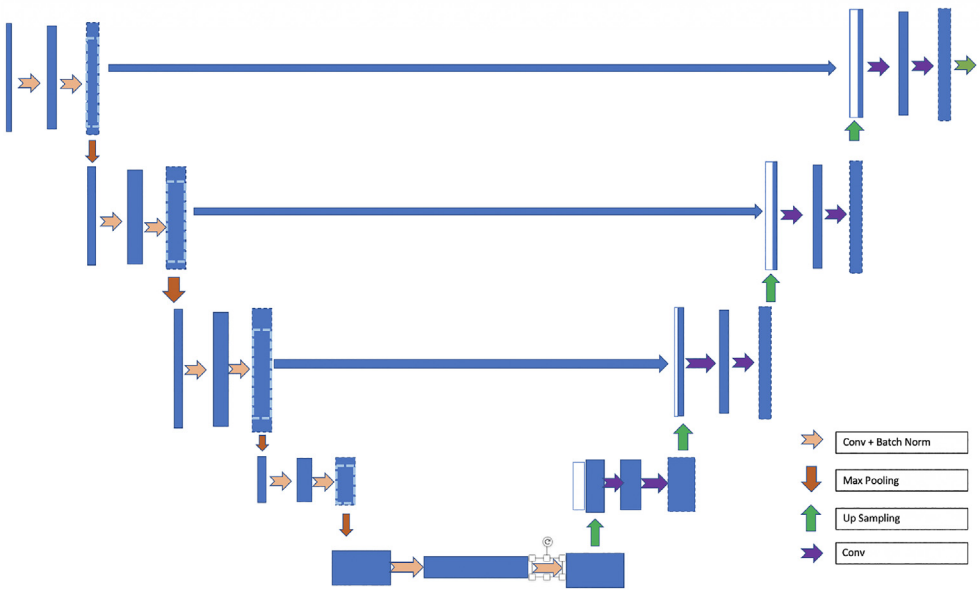
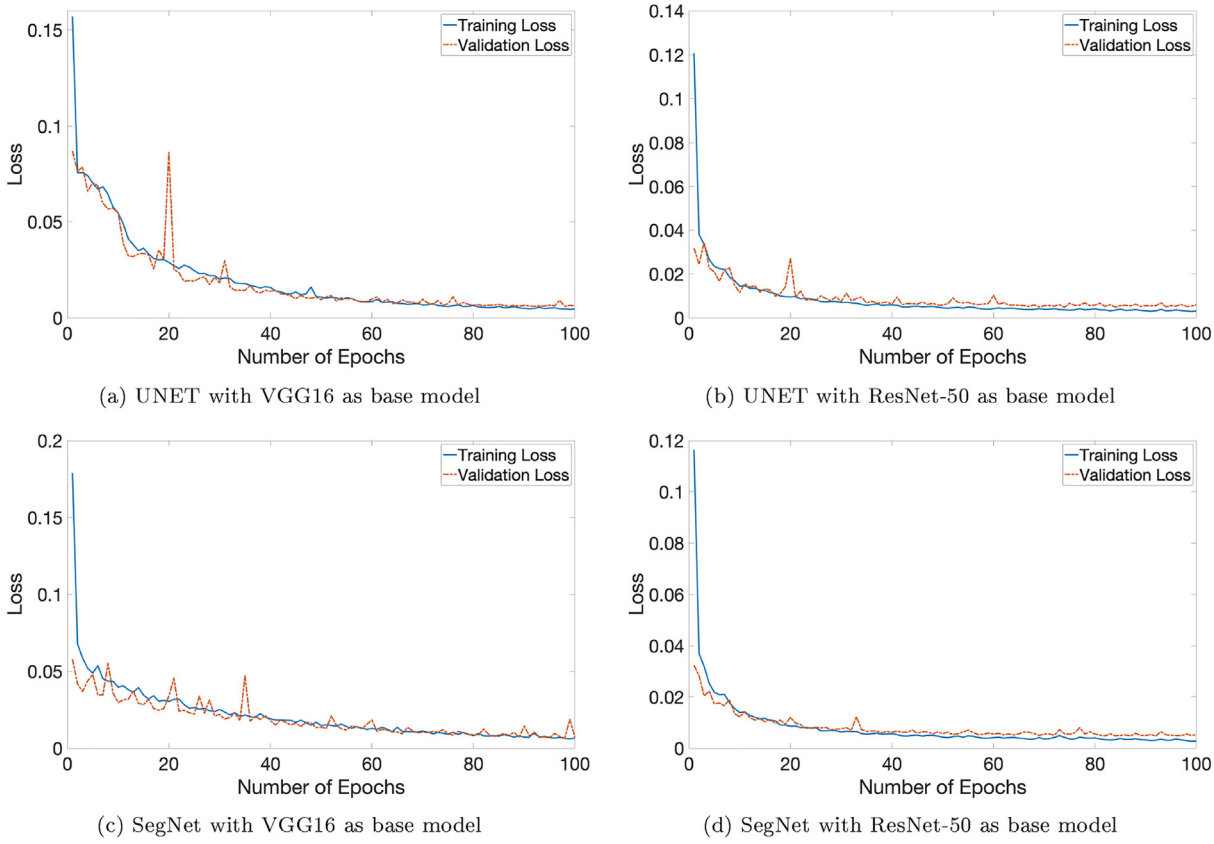


Fig. 8 – UNET meta-architecture [33].



(a) UNET with VGG16 as base model

(b) UNET with ResNet-50 as base model

(c) SegNet with VGG16 as base model

(d) SegNet with ResNet-50 as base model

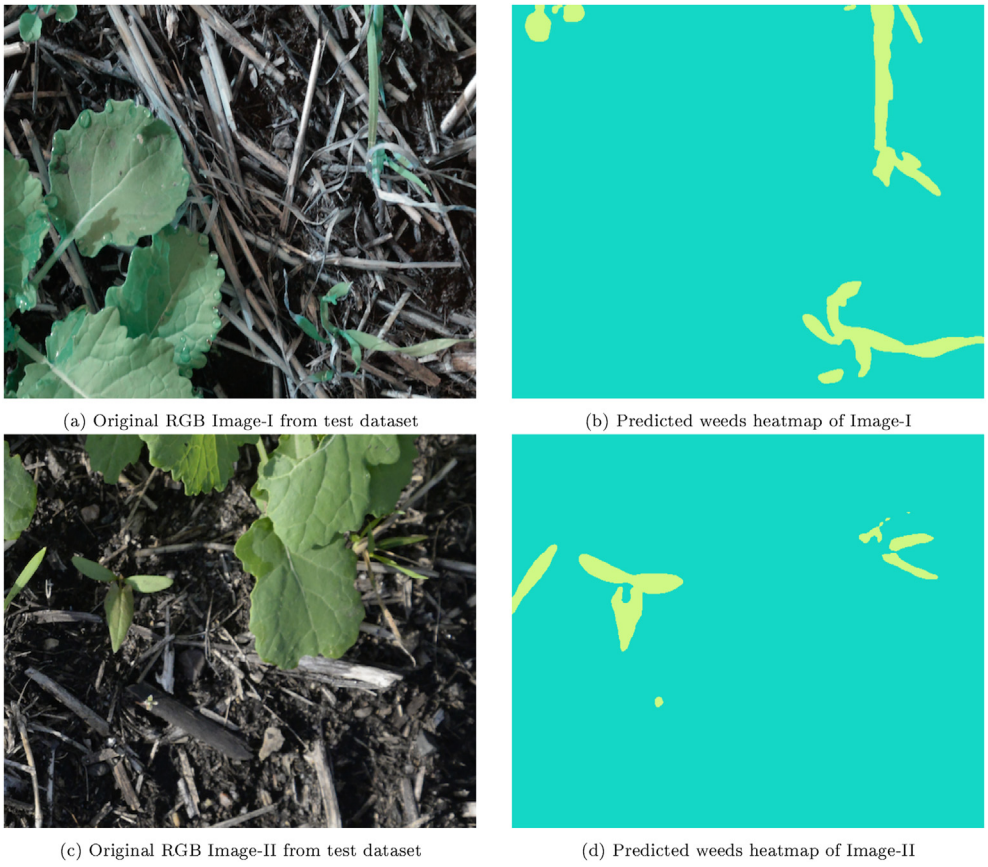
Fig. 9 – Convergence curves of trained models for 100 epochs.

convergence curves for trained models. It is observed that SegNet meta-architecture converges faster than UNET on given dataset. In SegNet, both validation and training loss curves do not vary much beyond 50th epoch.

Table 1 compares models on accuracy, IOU, MIOU and FWIOU metrics for test dataset. SegNet and UNET with ResNet-50 as base model are outperforming VGG16 based models. If we compare SegNet and UNET meta-

Table 1 – Comparing performance of deep learning models. Bold face shows the best performing model for the specific metric.

Models	Accuracy	F1-Score	IOU Non-Weed	IOU Weeds	MIOU	FWIOU
UNET_VGG16	0.9906	0.9952	0.9905	0.57	0.7805	0.9830
UNET_ResNet-50	0.9928	**0.9964**	0.9927	0.6622	0.8274	0.9868
SegNet_VGG16	0.9911	0.9955	0.9910	0.5930	0.7920	0.9839
SegNet_ResNet-50	**0.9948**	0.9929	**0.9928**	**0.6648**	**0.8288**	**0.9869**



(a) Original RGB Image-I from test dataset

(b) Predicted weeds heatmap of Image-I

(c) Original RGB Image-II from test dataset

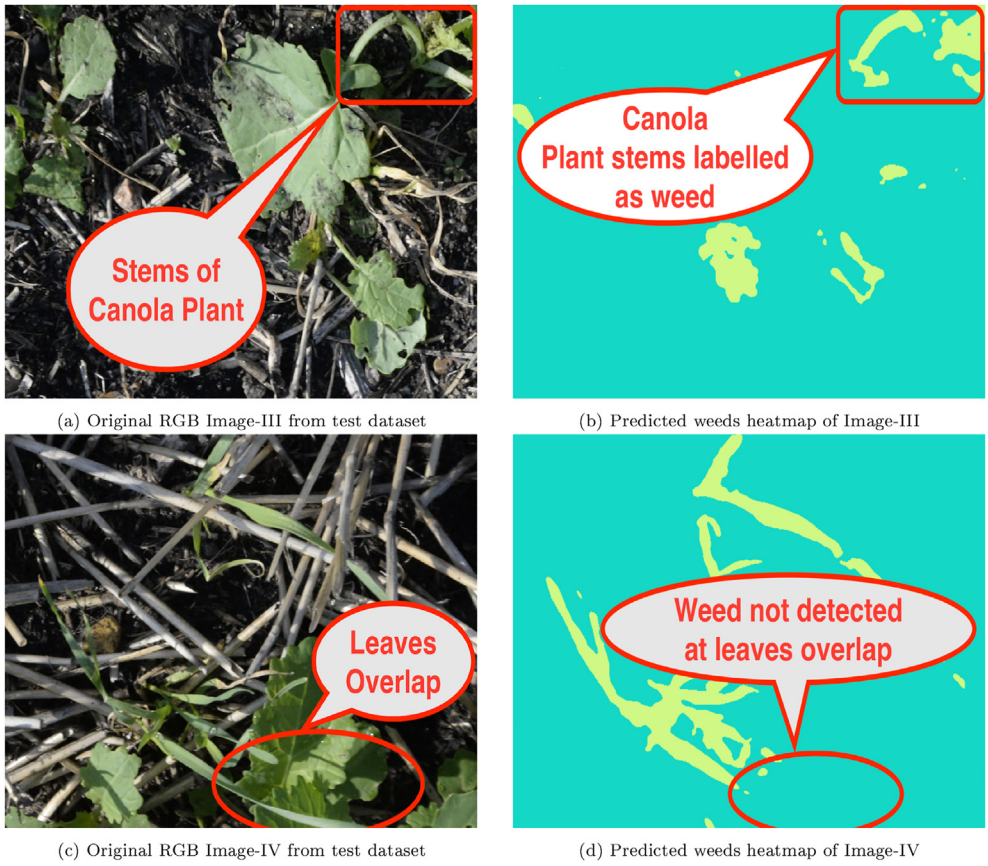
(d) Predicted weeds heatmap of Image-II

Fig. 10 – Original RGB test images and predicted heatmaps.

architectures, SegNet performs slightly better than UNET. Accuracy of the trained models is very high, which can be misleading as background and crop pixels are majority classes. In scenarios where class imbalance exists, majority class classifier accuracy is minimum criterion. Our trained model exceeds majority class classifier accuracy of 98.23%. This problem is addressed by IOU metric which provides criterion for each of the classes. Ideally, IOU value should be one. In our case, IOU of weed class is critical in accurate weed mapping and density estimation. ResNet-50 based SegNet model has highest IOU for weeds. Similarly, MIOU and FWIOU are also highest for ResNet-50 based models. ResNet-50 based SegNet have shown best performance on all metrics except F1 score. ResNet50 based UNET has highest value for F1 score.

These results are consistent with other similar works. For example, Xu Ma et al. show that SegNet is performing better than UNET in detecting weeds in rice fields [47]. In comparison to their work, the proposed methodology has shown an improvement of 24% and 16% in MIOU and FWIOU respectively. Higher values of MIOU and FWIOU could be attributed to the high-class imbalance. The lower value of weed class IOU is due to the complex nature of the images. The dataset is collected under actual field conditions. Most of the existing weed detection and semantic segmentation studies on agriculture images either involve synthetic data or they do data collection in a controlled environment which can not completely model actual field conditions like varied background, illumination effects, crop-weed distribution and orientations in images. Similarly, approaches that require acquisition of high-resolution images of crop and weed leaves are not scalable. These approaches are impractical for large farm holdings in the Canadian Prairies. Additionally, acquiring images through UAVs is not flexible due to high wind conditions in the Prairies.

The trained semantic models zero out crop pixels along with background pixels. Crop pixels are not separately classified because the main objective of the study is to estimate weed density (w_d) for variable rate herbicide application.



(a) Original RGB Image-III from test dataset

(b) Predicted weeds heatmap of Image-III

(c) Original RGB Image-IV from test dataset

(d) Predicted weeds heatmap of Image-IV

Fig. 11 – Original RGB test images and predicted heatmaps are highlighting the points where model fails to distinguish between weeds and canola plants.

However, crop density (c_d) can be estimated by subtracting weed density from background segmented vegetation density (v_d) given by following equation:

$$c_d = v_d - w_d \tag{10}$$

Fig. 10 shows original images and heatmaps of weeds as predicted by SegNet_ResNet-50. We can see canola and two types of weeds in Fig. 10a and c. The model combines the two types of weeds into one class and crop and background into other class. It is a binary classification of pixels. The model learns shape of crop plants and clubs it with background, and labels remaining vegetation in the image as weeds. Higher IOUs, MIOU and FWIOU are attributed to a bigger dataset for model training. UNET and SegNet work good on small dataset, but higher training instances increase model's ability to generalise well. However, there is one disadvantage of this methodology as it requires crop-specific model training which is manageable in Canadian Prairies because of small crop mix.

Fig. 11 shows example original test images and their predicted weed heatmaps for the points where model fails to distinguish between weeds and canola plants. One of the most common points where model confuses is stem part of canola plants as highlighted in Fig. 11a. Model classifies canola stem pixels as weeds confusing it with grass weed or wheat plants. Though model is good at mapping weeds in canola lines but it sometimes confuses in images where weeds leaves overlap canola leaves. In Fig. 11c wheat leaf is overlapping canola leaf. Model fails to map wheat leaf as shown in Fig. 11d.

5. Conclusion

Accurate mapping of weeds is a pre-requisite for weed density estimations and variable rate herbicide prescription. Semantic segmentation using deep learning is a promising technique for this purpose. Bottleneck for employing semantic segmentation is unavailability of labelled agriculture images at pixel level, which has been addressed in this paper. After acquiring high-resolution RGB images from canola field, background is segmented as a first labelling step, and then minority class pixel are manually labelled in background-segmented image. The trained model on this dataset maps only weeds and combines crop pixels along with background pixels. Our methodology has better results when we compare it to some of the most recent works involving semantic segmentation and weed detection. It has shown MIOU value of 0.8288 and FWIOU value of 0.9869 for ResNet-50 based SegNet model. Also, the paper makes comparison between UNET and SegNet meta-architectures for ResNet-50 and VGG16 base models. It is found that SegNet has better performance than UNET on this dataset. For base model feature extractors, ResNet-50 has superior performance than VGG16. In future work, soil properties will be included in the study for investi-

gating a relationship between weed densities and soil characteristics with the purpose to facilitate variable herbicide prescription for different soil zones.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We would like to thank Mr. Kevin Shearer of SAIRS Ltd. (https://www.sairs.ca/) for providing us with high-resolution images of canola at two different growth stages. Special thanks to Mr. Brett Bauereiss of JB Farms who gave access to his land and canola fields for this project. We also acknowledge the guidance and feedback of Dr. Joseph Piwovar in the field of remote sensing to conduct this study.

REFERENCES

[1] Oerke E. Crop losses to pests. J Agric Sci 2006;144(1):31–43.
[2] Benbrook CM. Trends in glyphosate herbicide use in the united states and globally. Environ Sci Europe 2016;28(1):3.
[3] Osteen CD, Fernandez-Cornejo J, et al. Herbicide use trends: a backgrounder. Choices 2016;31(4):1–7.
[4] Myers JP, Antoniou MN, Blumberg B, Carroll L, Colborn T, Everett LG, et al. Concerns over use of glyphosate-based herbicides and risks associated with exposures: a consensus statement. Environ Health 2016;15(1):19.
[5] Bah MD, Hafiane A, Canals R. Deep learning with unsupervised data labeling for weeds detection on UAV images; 2018a. arXiv: 1805.12395.
[6] Plant R, Pettygrove G, Reinert W, et al. Precision agriculture can increase profits and limit environmental impacts. Calif Agric 2000;54(4):66–71.
[7] Hemming J, Rath T. Pa—precision agriculture: computer-vision-based weed identification under field conditions using controlled lighting. J Agric Eng Res 2001;78(3):233–43.
[8] Kamilaris A, Prenafeta-Boldú FX. Deep learning in agriculture: a survey. Comput Electron Agric 2018;147:70–90.
[9] Oerke EC, Gerhards R, Menz G, Sikora RA. Precision crop protection-the challenge and use of heterogeneity, vol. 5. Springer; 2010.
[10] Saari H, Pellikka I, Pesonen L, Tuominen S, Heikkilä J, Holmlund C, et al. Unmanned aerial vehicle (UAV) operated spectral camera system for forest and agriculture applications. In: Remote sensing for agriculture, ecosystems, and hydrology XIII, vol. 8174. International Society for Optics and Photonics; 2011. p. 81740H.
[11] Wu X, Xu W, Song Y, Cai M. A detection method of weed in wheat field on machine vision. Procedia Eng 2011;15:1998–2003.
[12] Bah MD, Hafiane A, Canals R. Deep learning with unsupervised data labeling for weed detection in line crops in UAV images. Remote Sens 2018;10(11):1690.
[13] Ng H, Ong S, Foong K, Goh P, Nowinski W. Medical image segmentation using k-means clustering and improved watershed algorithm. In: 2006 IEEE southwest symposium on image analysis and interpretation. IEEE; 2006. p. 61–5.
[14] Rusu RB. Semantic 3d object maps for everyday manipulation in human living environments. KI-Künstliche Intelligenz 2010;24(4):345–8.
[15] Valiente-Gonzalez JM, Andreu-García G, Potter P, Rodas-Jorda A. Automatic corn (zea mays) kernel inspection system using novelty detection based on principal component analysis. Biosyst Eng 2014;117:94–103.
[16] Slaughter D, Giles D, Downey D. Autonomous robotic weed control systems: a review. Comput Electron Agric 2008;61 (1):63–78.
[17] García-Santillán ID, Pajares G. On-line crop/weed discrimination through the Mahalanobis distance from images in maize fields. Biosyst Eng 2018;166:28–43.
[18] Guerrero JM, Ruz JJ, Pajares G. Crop rows and weeds detection in maize fields applying a computer vision system based on geometry. Comput Electron Agric 2017;142:461–72.
[19] Wendel A, Underwood J. Self-supervised weed detection in vegetable crops using ground based hyperspectral imaging. In: 2016 IEEE international conference on robotics and automation (ICRA). IEEE; 2016. p. 5128–35.
[20] Okamoto H, Murata T, Kataoka T, Hata SI. Plant classification for weed detection using hyperspectral imaging with wavelet analysis. Weed Biol Manage 2007;7(1):31–7.
[21] Zhao W, Du S. Spectral–spatial feature extraction for hyperspectral image classification: a dimension reduction and deep learning approach. IEEE Trans Geosci Remote Sens 2016;54(8):4544–54.
[22] De Rainville FM, Durand A, Fortin FA, Tanguy K, Maldague X, Panneton B, et al. Bayesian classification and unsupervised learning for isolating weeds in row crops. Pattern Anal Appl 2014;17(2):401–14.
[23] Pantazi XE, Moshou D, Bravo C. Active learning system for weed species recognition based on hyperspectral sensing. Biosyst Eng 2016;146:193–202.
[24] Rumpf T, Römer C, Weis M, Sökefeld M, Gerhards R, Plümer L. Sequential support vector machine classification for small-grain weed species discrimination with special regard to cirsium arvense and galium aparine. Comput Electron Agric 2012;80:89–96.
[25] Haug S, Michaels A, Biber P, Ostermann J. Plant classification system for crop/weed discrimination without segmentation. In: IEEE winter conference on applications of computer vision. IEEE; 2014. p. 1142–9.
[26] Liakos KG, Busato P, Moshou D, Pearson S, Bochtis D. Machine learning in agriculture: a review. Sensors 2018;18 (8):2674.
[27] Olsen A, Konovalov DA, Philippa B, Ridd P, Wood JC, Johns J, et al. Deepweeds: a multiclass weed species image dataset for deep learning. Sci Rep 2019;9(1):2058.
[28] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. p. 3431–40.
[29] Dyrmann M, Mortensen AK, Midtiby HS, Jørgensen RN, et al. Pixel-wise classification of weeds and crops in images by using a fully convolutional neural network. In: Proceedings of the international conference on agricultural engineering, Aarhus, Denmark; 2016. p. 26–9.
[30] Potena C, Nardi D, Pretto A. Fast and accurate crop and weed identification with summarized train sets for precision agriculture. In: International conference on intelligent autonomous systems. Springer; 2016. p. 105–21.
[31] Ahmad A, Quegan S. Analysis of maximum likelihood classification on multispectral data. Appl Math Sci 2012;6 (129):6425–36.
[32] Sisodia PS, Tiwari V, Kumar A. Analysis of supervised maximum likelihood classification for remote sensing image.

In: International conference on recent advances and innovations in engineering (ICRAIE-2014). IEEE; 2014. p. 1–4.
[33] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: International conference on medical image computing and computer-assisted intervention. p. 234–41.
[34] Badrinarayanan V, Kendall A, Cipolla R. Segnet: a deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans Pattern Anal Mach Intell 2017;39 (12):2481–95.
[35] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition; 2014. arXiv preprint arXiv: 14091556.
[36] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. p. 770–8.
[37] Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition. p. 2818–26.
[38] Chollet F. Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition. p. 1251–8.
[39] Guo Y, Liu Y, Georgiou T, Lew MS. A review of semantic segmentation using deep neural networks. Int J Multimedia Inform Retrieval 2018;7(2):87–93.
[40] Kendall A, Badrinarayanan V, Cipolla R. Bayesian segnet: model uncertainty in deep convolutional encoder-decoder architectures for scene understanding; 2015. arXiv preprint arXiv: 151102680.
[41] Li X, Chen H, Qi X, Dou Q, Fu CW, Heng PA. H-denseunet: hybrid densely connected unet for liver and tumor segmentation from ct volumes. IEEE Trans Med Imaging 2018;37(12):2663–74.
[42] Esri. Maximum likelihood classification; 2019. http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/maximum-likelihood-classification.html.
[43] Russell TC, Torralba A, Murphy KP, Freeman WT. Labelme: a database and web-based tool for image annotation. Int J Comput Vision 2008;77:157–73.
[44] NVIDIA Corporation. GeForce GTX 1080 Ti Specifications; 2019. URL https://www.geforce.com/hardware/desktop-gpus/geforce-gtx-1080-ti/specifications.
[45] Thoma M. A survey of semantic segmentation; 2016. arXiv preprint arXiv: 160206541.
[46] Zeiler MD. Adadelta: an adaptive learning rate method; 2012. arXiv preprint arXiv: 12125701.
[47] Ma X, Deng X, Qi L, Jiang Y, Li H, Wang Y, et al. Fully convolutional network for rice seedling and weed image segmentation at the seedling stage in paddy fields. PloS One 2019;14(4):e0215676.