Self-Supervised Overlapped Multiple Weed and Crop Species Leaf Segmentation under Complex Light Condition

Anand Muni Mishra
Chitkara University Institute of Engineering and Technology, Chitkara University
Prabhjot Kaur
Chitkara University Institute of Engineering and Technology, Chitkara University
Mukund Pratap Singh
Bennett University
Santar Pal Singh (spsingh78@gmail.com)
Rashtrakavi Ramdhari Singh Dinkar College of Engineering, Begusarai

Research Article

Keywords: Segmentation, U-SegNet, PSP-U-SegNet, Semantic Segmentation, Weed Species

Posted Date: July 21st, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1851240/v1

Additional Declarations: No competing interests reported.

Self-Supervised Overlapped Multiple Weed and Crop Species Leaf Segmentation under Complex light Condition

Anand Muni Mishra[1], Prabhjot Kaur[1], Mukund Pratap Singh[2], Santar Pal Singh[3*]

[1]*Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India*
[2]*School of Computer Science Engineering and Technology, Bennett University, Greater Noida, India*
[3]*Dept. of Comp. Sc. &Engg.Rashtrakavi Ramdhari Singh Dinkar College of Engineering, Begusarai, Bihar, India*
*(*Corresponding author email id: spsingh78@gmail.com)*

Abstract:

Weeds are unwanted plants that compete with target crops and absorbed the required nutrients from the soil, sunlight, air, etc. The farmers are suffering from weed identification detection due to the homogeneous morphological feature of weed and crop leaves. Computer vision is a sophisticated technique which widely used for weed and crop leaves identification and detection in the agricultural field. This work has used the three different datasets such as "Deep weed", "Crop Weed Filed Image Dataset" (CWFID), and "Leaf Segmentation Challenge" (LSC), and collected 5090 images for training the model. In this work we have used three grass species as *Setariaverticillata, Digitariasanguinalis, Echinochloa crus-Galli, and three broad leaf species as Cerastiumvulgatum L.,Chenopodiumalbums, Amaranthusretroflexus in Vignamungo crops.* The first dataset includes 2000, the second 1720 images, and 1370 images from the third dataset. We have proposed a deep learning segmentation model as PSPNet-U-SegNet for data classification and compared the data accuracy from existing segmentation models such as U-SegNet, and U-SegNet.Theresult has been shown the deep weed dataset has achieved98.97% data accuracy and 8.9m IoU. Our findings demonstrate that adding more datasets to the actual field picture collection improves network performance while requiring less manual annotation work.

Keywords:-Segmentation; U-SegNet; PSP-U-SegNet; Semantic Segmentation; Weed Species.

1. Introduction

The weed is an unwanted plant in the crop fields. Weeds compete with crops for soil, nutrients, and sunshine, which causes crops to develop slowly and become smaller. This reduces agricultural production. Therefore, these nutrients are required for the growth of crop plants but due to the presence of weeds plant, the crop growth is affected[1]. There are two important factors for crop yield loss first is weed density and mix and similar morphological properties of weed/crop plant. In the current situation, the farmers manually assess the weeds[2]. Another important factor is overlapped weed plant. For these, the weed plant identification of overlapped weed plants, detection, coverage area, and growth stages are measured in the work. However, it has a tedious task for weed identification and classification which has affected crop yield. So the

automation of these tasks is interest to researchers in recent years[3].

Weed recognition has the focus of the computer vision system. The primary problems is changing the morphological property of weed/ crop plant due to environmental conditions. So, the collection the data field images is tedious task and another is identification and classification of overlapped weed/crop leaves in computer vision. The main objective to creating effective models for identification and classification of overlapping weed /crops leaves, uneven weed patch densities, varying sizes across multiple images, and discriminating the similar morphological property of weeds and crops leaves [4].In addition, the large extent of plantings or mixed crop weeds of computational time problem in image processing methods. The recent Deep learning technique has proven to overcome the limitation of the classical image processing model[5].

Consequently, CNN has had great success classifying plant species. Crop disease detection, plantsegmentation,andweedcharacterization,amongotherapplications.However,CNNshavesomedisadvantages[6].The enormous amount of manually annotated images needed to establish a model is one of them. Annotating the needed images by hand is a time-consuming and in some cases impossible operation. In 2016, the author has proposed the deep learning-based semantic segmentation for weed identification and detection of weed and crop leaves[7].

They attempted to find any weed species but managed to separate several crop species with a pixel accuracy of 0.79. However, their findings showed that employing CNNs for the task of identifying crops and weeds has enormous promise. By accurately classifying pixels as "corn" or "weed" in a two-class classification issue, they enhanced the approach to differentiate maize from seven distinct kinds of weed species in later trials [8].This work has attained 0.94 per-pixel accuracies. This work has achieved anF1 score of 0.80on crop and an Intersection-Over-Union metric of 0.81 based on pixel base data classification of weed and crops leaves[9].The primary benefit of this work is can reliably detect weed-infested areas and compare the accuracy of existing models to that of the proposed PSP-U-SegNet model.

i. Compute overlapped weed image and density in the infected regions using vegetation pixel-wise tile classification of weed data.
ii. Use pixel and tile wise data classification with different tile size as 25x25(px), 50x50(px), and 75x75(px), for binary classification of image and achieve 9.7 IoU of data segmentation.
iii. Enhance scalability and generalizability using semantic and vegetation segmentation.

Since various plant species may only be identified by precise and nuanced taxonomic keys that may not even always be apparent in an image, segmenting multi-species overlapping weeds is a more challenging and difficult challenge than it has been previously [10].In this study, we use a combined strategy to discuss multispecies overlapping segmentation. To

eliminate the requirement for manual annotation, we first provide a unique approach to integrating synthetic and single-species datasets. Then, we suggest a novel architecture to carry out multispecies semantic segmentation effectively.

The rest of the paper is structured as follows: related work in illustrated in Section 2. The Section 3 focused on the data description and methodology is discussed in Section 4. The performance analysis of model and results are discussed in Section 5. Finally, the conclusions drawn are presented in Section 6.

2. Literature Study

For a more thorough explanation, Precision agriculture has benefited from recent developments in deep learning to overcome the drawbacks and rigidity of conventional approaches. This literature has categorized three different learning techniques first is weed identification, and the second is Deep and transfer learning.

Mishra et al. (2022) has been discussed the different types of biennial and perennials, a monocot, and broad leaves weeds species and also described biological control methods. It has also described the morphological and texture property of common perennial weeds such as "Paspalum Distichum", "Cynodon Dactylon", "Scirput Maritimus", and "Cyperus Rotundus" in Paddy crop agriculture. Furthermore, they have also described the weed control technique such as Biological, Cultural, Physical, and Chemical control methods. The authors use instance and semantic segmentation techniques for object detection, and Gray Level Co-occurrences Matrix (GLCM), Hue, Saturation, Value (HSV) are used for feature extraction. The author has applied different CNN techniques for image data classification and compares the technique based on the performance of the model. There are a few performance parameters that have been discussed by the author in terms of Precision, Recall, F1-score, Accuracy, Absolute Error (AE), and Mean Absolute Error (MAE) [11].

X Ma. et al. (2020) have discussed the Positive Enable (PE) technique for finding the exact location of plant objects and classifying the data using the Encoder-Decoder Conventional Neural Network (ED-CNN) model. The authors have used 4000 weed images and it has been collected from the "Deep weed dataset" for the experiment. There are five different families of weeds such as Cronopus Didymium, Fumaria Parvilfora, Lathyrus Aphaca, and Medicago polymorpha that have been used for data identification and detection. The author has applied a different CNN model for image data classification; they compare their proposed techniques based on the different performance matric. After comparing the different CNN models, the ED-CNN has achieved 94.03% accuracy [12].

Chechlinski et al. (2019) have suggested automated weeding called agro robotics. In this

technique, Weeds can be identified using robotics technology. The author has described the Internet of Things (IoT) and Deep Learning (DL) based technique which has automatically recognized weed identification and detection. The model has achieved 47-67% weed detection accuracy. It has been tested in four different plants in a stadium and under medium lighting conditions. The robotics system has used the custom semantic segmentation CNN using U-Net, DenseNet, and ResNet architecture. Out of this CNN architecture, ResNet pre-trained model achieved better 87% data accuracy. The author suggested that weed images can easily transfer to computer vision to another agro-robotic task [13].

Rasti et al. (2019) have discussed discriminating the weeds from the soya bean crop plant. They used pre-trained DL models such as AlexNet, SqueezeNet, GoogLeNet, ResNet-50, SqueezeNet-MOD1, and SqueezeNet-MOD2 for training the model. Furthermore, 11,600 weed images have been collected from Crop Weed Field Image Dataset (CWFID) and trained in the models. The ResNet-50 has achieved more than 92% data accuracy. AlexNet, SqueezeNet, GoogLeNet, SqueezeNet-MOD1, and SqueezeNet-MOD2 has been achieved 94%, 91%, 87%, 90%, 95% data accuracy consecutively. The author calculated the processing time of the ResNet CNN model and achieved 40.73 s, to process 11,600 images. However, the author suggested that it can also implement in biotic and abiotic leaf disease identification and detection [14].

Teimouri et al. (2018) have discussed 10 different types of weed species that grew in rabi and Kharif crops. They explained the morphological and texture property of weed leaves. Furthermore, the author described weed detection and classification techniques. There are 9649 weed/crops images collected for the standard data repository as the CWFID dataset. In this context, the author used three different classifiers such as ResNet 150, Google Net, and VGGNet-16 pre-trained CNN model for data classification. Out of these VGG-16 model has achieved the 96% of data accuracy [15].

Kropff et al. (2021) have suggested the weed identification and detection technique based on four different steps i.e. data collection, data segmentation, feature extraction, and finally data classification. Data has been collected from a multi-class Deep Weed dataset. After that data has been annotated as “CynodonDactylon”, “Convolvulus Arvensis”, “Poa Annual”, “MedicagoPolymorpha”, and “HypochaerisRadicata”. The unstructured RGB data has been resized as 256×256×3 and then implemented in the semantic segment for object detection. For the classification used SegNet, U-Net, and ResNet151 CNN model and achieved 93.05%, 93%, and 92.78% data accuracy. The author has compared the proposed model in terms of accuracy and found that the SegNet CNN model provides better accuracy. The author also discussed the computation time of image processing in the CNN model. From the experimental results, it is found that the SegNet classifier consumed less time i.e. 0.90 ms[16].

3. Dataset description

In this study, the functional dataset was trained using instance and semantic segmentation. Three distinct datasets—"Deep-weed," "CWFID," and "LSC"are used in this study to annotate images. To expedite the manual annotation of real image datasets (dataset i)[14]. we first present certain changes. Additionally, we present various techniques for creating datasets without the need for manual annotation: a) A technique for creating artificial datasets based on a single plant image (dataset ii))[17]and a technique for creating actual field datasets made up of numerous plant images of a single weed species (dataset iii)[18].The complete discussion of the dataset is given in the next subsection.

i. Dataset A: *Deep weed data set*

to create an appropriate image collection for training and validation. Due to its potential to enhance agricultural output, research into robotic weed management has expanded recently. Deep learning is the best method for identifying different weed species in the challenging grasslands habitat because of its unmatched accomplishments. The first sizable, publicmulticlass image collection of weed species from Australian grasslands is provided by this study, enabling the development of reliable classification techniques to enable effective robotic weed treatment.This work has collected 1720 weed broad leaf weed species such as *Cerastiumvulgatum L.,Chenopodiumalbums,andAmaranthusretroflexus*[19].

ii. Dataset B:*CWFID dataset*

This dataset has standard weed and crop image data repository and there are 2000 grass are collected for training the model. Furthermore, Setariaverticillat and Digitariasanguinalis have collected 1200 and 800 weed images from grass weed species which are available online (http://github.com/cwfid)[20]. For each image from the dataset, we present a ground truth vegetation segmentation mask and manual annotation of the plant category (crop vs. weed).

iii. Dataset C: *LSC dataset*

This work has collected 1370 "*Echinochloa crus-galli"*grassweed images from the LSC dataset. We developed the Leaf Segmentation Challenge (LSC) to show the current status of leaf segmentation technology and the challenges of segmenting all leaves in a plant picture.Thecountablity of the dataset is described in figure 1.

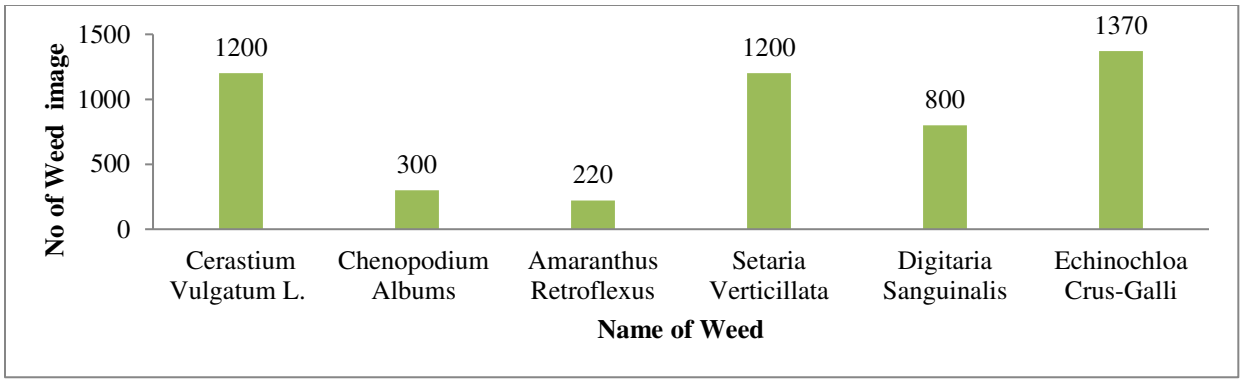


Figure1. Dataset description

4. Methodology

The suggested method determines the weed-infested areas, weed leaf count, weed growth, and the related weed density to treat the farmland under cultivation in a targeted manner. These four processes—pre-processing, segmentation, feature extraction, and classification—were used in this study to train the model.

4.1. Data Enhancement and pre-processing of the image

This work has collected the 5090 weed images from different sources. This dataset has been pre-processed and segmented the particular object from the image. These weed images are pre-processed using the Contrast-limited Adaptive Histogram Equalization (CLAHE) technique, which improves the image quality[21].Data pre-processing, data segmentation, feature extraction, and data classification have all been assigned to the data flow. Furthermore, data segmentation has use sematic, vegetation, and background segmentation[22].

This algorithm is a method of computer image processing that boosts contrast in pictures. The adaptive method differs from typical histogram equalization in that it computes several histograms, each one corresponding to a distinct region of the image, and then uses them to distribute the brightness values of the image. After that, each tile's transformation function is calculated. The pixels in the tile center are a good fit for the transformation functions [25]. All other pixels are given interpolated values and up to four transformation functions based on the center pixels of the tiles that are closest to them. The bulk of the image's pixels (shown in shaded blue) are interpolated bilinear; those near the edge (shown in shaded green) are interpolated linearly, and those near the corners (shown in shaded red) are converted using the corner tile's transformation function. To ensure that the output is continuous as the pixel gets closer to a tile center, the interpolation coefficients represent the locations of pixels between the nearest tile center pixels. The complete flow of data is given in figure2.

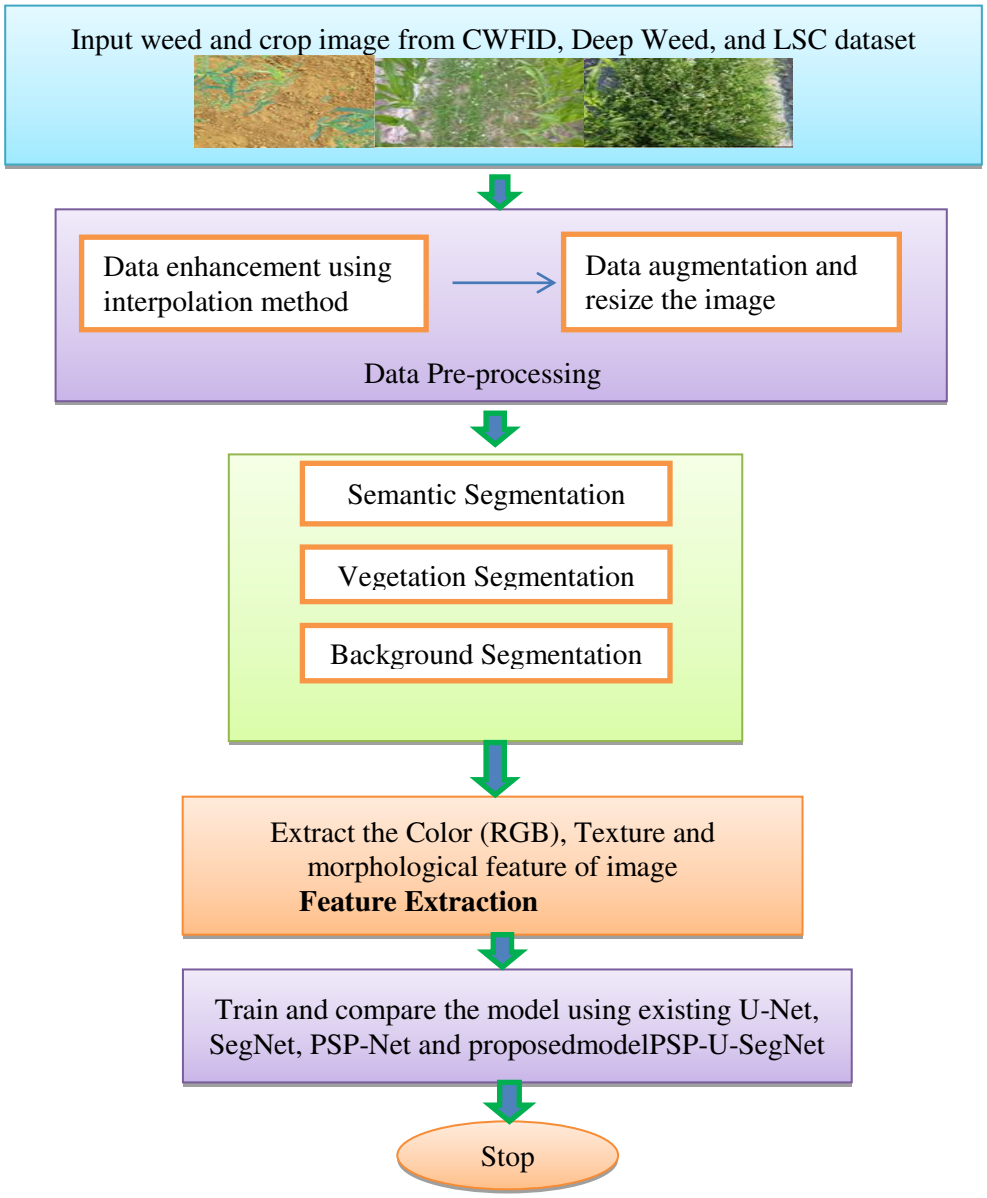


Figure 2. Complete data flow of model.

This work has enhanced the quality of the image using hologram equation 1

$$q_n = \frac{\text{No of the pixel of the image with intensity n}}{\text{total no of a pixel of an image}} \quad \text{(eq. 1)}$$

The function "f" is represented as n_r by n_c matrix of integer pixel which has like 0 to m-1. This intensity value is up to 0 to 255. Let $'q'$ are normalized hologram of $'g'$. This hologram equation can define as equation 2.

$$h_{ij,j} = floor((m-1)\sum_{n=0}^{g_{i,j}} p_n)) \quad \text{(eq. 2)}$$

The floor function rounds down/up to the nearest integer to the transform equation 3.

$$T(k) = floor((m-1)\sum_{n=0}^{k} p_n)) \qquad \text{(eq. 3)}$$

This equation has been imported from equation 4.

$$z = T(y) = (m-1)\int_{0}^{y} p_y(y)dy \quad \text{(eq. 4)}$$

Where p_y is the probability density function (pdf) of, 'T' is the distribution function of y.

Assume T is invertible and differentiable, y multiplied $(L-1)$ by which has defined as in equation 5.

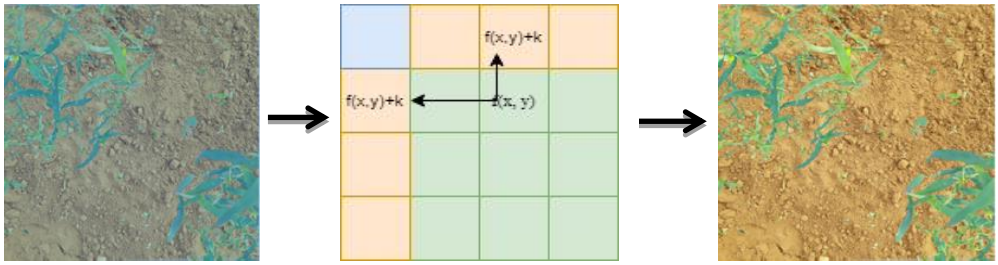
$$p_z(z) = \frac{1}{L-1} = 1 \quad \text{(eq. 5)}$$

This equation defines as a high-density pixel.

$$f(x,y) = T(f(x,y) + k) \quad \text{(eq. 6)}$$

Where $f(x,y)$ is the coordination of x and y axis value, k is the constant value it will be $0\ to\ 255$.

The approximation of weed and crop image pX(x) are illustrated transformation in equations 1 and 2. Although the histograms produced by the discrete version won't be completely flat, they will be flattened, which will improve the contrast of the image. The picture improvements took an average of 15 minutes. Enhance the quality of the weed image technique is given in figure 3.



Blur weed image (Chena podium Album L.)

Using hologram transform enhance image

Enhanced weed image

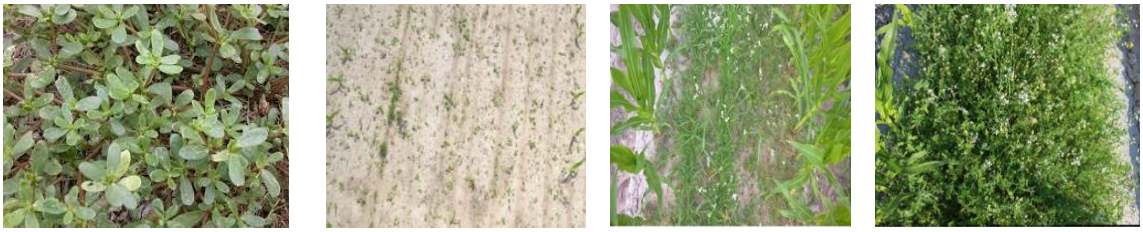
Figure 3. Enhance weed image

4.2. Overlapping plant leaves and density estimation of weeds

Generally, most of the different varieties of the plant are germinated in the field. This study useda"*Vignamungo"* plant field image with seven different classes of weed images. All these classes have overlapped weed images. The sample of some overlapped weed plants is given in figure 4. Most weed leaves are overlapped which has decreased the performance of the classifier. The tiles classification is a sophisticated technique for identifying weed and crop plants. This work uses $25x25, 50x50, 75x75,$ and $100x100$ sizes of tile for calculating the overlapped weed image. The weed density is calculated based on weed-infested regions. The weed-infested region is identified by tile classification which can be calculated by vegetation coverage in each region. In this work, the weed density has been calculated as Weed Cluster Rete (WCR) [23] defines in equation 7.

$$\text{WCR/Weed density} = \frac{\text{Weed plant coverage in tile}}{\text{total area has covered in region}} \qquad \text{(eq.7)}$$

This density estimate will help in selecting suitable areas for weeding along with herbicides in the field.Some overlapped image is given in figure 4.



Cerastiumvulgatum L

Chenopodiumalbums

Amaranths retroflexus

Setariaverticillata

Figure 4. Overlapped weed plant leaves

4.3. Weed/ crop image data Segmentation

Enhanced 5090 images are put as input for the pipeline. For the segmentation, images are grouped into three clusters first is semantic segmentation for homogeneous weed object second is background segmentation for discrimination of object and third is vegetation segmentation for foreground segmentation of object The semantic segmentation creates the homogeneous target object in the same pixel intensity of objects. For the discrimination of objects, there is two other segmentation such as vegetation and background segmentation[24]. This segmentation technique creates the vegetation mask and mask object it may be weed leaf or crop leaves. The complete process has been done using tile classification. The tile has been generated in the Region of Mask (RoI). The complete segmentation has over using vegetation, semantic, and background segmentation techniques. The detailed descriptions are given in the next sub-section.

4.3.1. Vegetation segmentation of the object

After pre-processing of an image, image segmentation is the next specific task for discriminating the weed and crop plant from field image data. The vegetation segmentation is the foreground of the specific object. These objects can discriminate the overlapped weed image and location estimation of the object. When the picture mask is applied, the only pixels that appear in the vegetation are those that are not zero. following binary image segmentation, a particular plant or weed displayed in different colors of the image and individual plants should be segmented [25].This particular task is challenging because weed and crop plants grow together. Sometimes weeds and crops leaves are overlapped occur. The vegetation segmentation can also include information such as the growth stage of weed/plant, leaf count; stem position, biomass amount, and others. Furthermore, it can also calculate the plant coverage ratio in the field; interspacing of plants, and count of plant cam measure in the field. Some weed vegetation segmentation is given in figure 5.

4.3.2. Background segmentation of objects

The vegetation segmentation is foreground segmentation that can discriminate the specific object. Our system's initial stage is foreground-background segmentation, which takes into the difference between the actual picture and a background model. Foreground refers to areas where the observed picture and the backdrop model differ considerably. The background image has a different frequency of pixel it may be a high or low-density pixel. A collection of photos of the empty working space is usually used to create the backdrop model. Because the same model is used for consecutive photos, background removal only works for static backgrounds. It has a high-density pixel object[26].The background segmentation includes high and low-density pixels of the complete object.

4.3.3. Semantic segmentation of objects

Semantic segmentation is the process of assigning a label to each pixel in an image. This contrasts with classification, which gives the entire image a single label. Semantic segmentation treats many objects belonging to the same class as a single entity. These techniques create the weed or crop object inhomogeneous color object which has helped to identify the weed or crop object. There are some weed and crop objects given in figure 6.

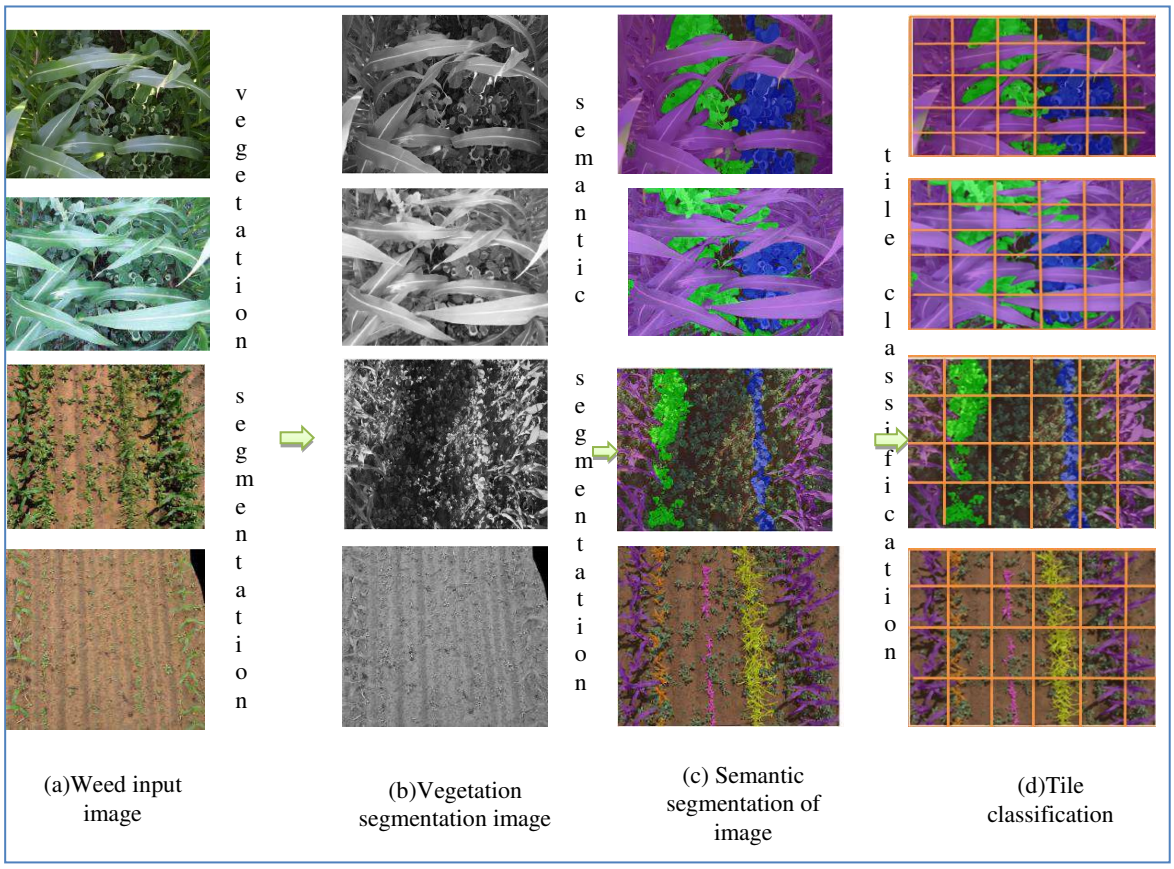


Figure 6. Some weed and crop plant segmentation

Figure 6 includes some different categories of images such as vegetation segmented and semantic segmented images. The vegetation segmentation image includes foreground high density of the object and these objects have the same density pixel using semantic segmentation[27]. The object has been identified using tile classification. The tile includes high-density pixels. After that pixels are put on a future vector for feature extraction.

4.3.4. Tile classification of the object

Further, more input weed images data has been taken from the "Deep Weed", "CWFID", and LSC"datasets and acquired black gram field images. The concept of inputting any single weed image (H_{RGB}) has represented the image. The object has been identified by vegetation mask (H_{veg}) which has been generated by H_{RGB} and applied for H_{masked}. The H_{RGB} has achieved Region of Concern(RoC)which is denoted H_{masked} as an object. Furthermore, the masked image (H_{masked}) is distributed as small tiles (H_{tile}), often the patches are square tiles. It may be 25x25(px), 50x50(px), and 100x100(px).The term tile (H_{tile}) denotes the morphological characteristic of weeds taken from and in possession of the vegetation pixels at any given time in the image (H_{tile}). Additionally, the resulting scores are used to categorize plants as either weeds or crops. A binary classifier is used to categorize these plants (crop/weed). Utilizing the vegetation segmentation approach for classification, Weed and crop density performance measurements have been completed [28]. There are a few abbreviations used in the algorithm (OWID) given in **Table 1**.

Table 1. Abbreviations used in the algorithm 1 (OWID)

Notation	Meaning
H_{RGB}	Input colored weed image of weed
H_{veg}	Vegetation mask
H_{masked}	Fragmented image
H_{tile}	Square tiles
VCRE	Vegetation Cluster Rate Estimation
S_h	The complete height of Weed
O_{wd}, F_V	Segmented weed object, Feature Vector
$M_p(indices)$	Max pool indices, exponent variable of softmax.
CE, and BCE	Cross-Entropy and Binary Cross-Entropy

The steps of the proposed Overlapped weed/crop image Data (OWID) algorithm are given in Algorithm 1and Algorithm 2 and figure 7**.**

Algorithm 1: Estimation of Overlapped weed/Crop Image Data (**OWID**)

Input: Take H_{RGB} of the field acquired from dataset.

Output: OverlappedWeed object (O_{wd}) and density $(W_{density})$

- Given (H_{RGB}), Generate the vegetation mask (H_{veg}) using CNN-based segmentation;
- overlay H_{RGB} With H_{veg} to get H_{masked};
- divide the image H_{masked} into smaller regions H_{tile} (square tiles);
  - ❖ for (H_{tile} in H_{masked}) do

Classify H_{tile} into crop, weed, or background of image;

Classify high-density pixel and put on Feature Vector-1 (FV_1)

- Set threshold value as 2700 pixels.
  - ❖ $If\ (pixel\ value > 2700)$

Print →over-segmentation

else

- Compare each pixel from the object and segment the object as (O_{wd})
- Return→ O_{wd}
- estimate $W_{density}$ where $W_{density} = \frac{Weed\ plant\ coverage\ in\ region(O_{wd})}{Total\ land\ area}$
- Return→ $W_{density}$
- end

Applying segmentation based on CNN, create the vegetation mask (H_{veg}) from the picture (H_{RGB}), which is taken from a common data store. This segmentation is overlaid H_{RGB} with H_{veg} to get H_{masked}. It has divided the image H_{masked} into smaller regions H_{tile} (square tiles). Furthermore, Classify H_{tile} into crop, weed, or background of the image. The high-density pixel and put on Feature Vector$(FV_1)$with threshold value as 2700 pixel and check over segmentation. Further, the segment the object as (O_{wd}) use for $W_{density}$ calculation.

Algorithm 2.Execute Overlap Weed Data (EOWD)

Input: Overlapped Weed object (O_{wd}) and density $(W_{density})$

Output:MAE, IoU, MIoU

The input O_{wd} object is divided into 80:20 ratios of data.

- The pre-processed image on modify max-pool layer of PSP-U-SegNetand update max pool layer Max pool, int l =256; Update last 3 max-pool layer of

- Max Pool -Layer PSP-U-SegNet(3PSPU-SegNet)

Input $C64(3\times3\times256)$ and remove one skip connection

Print→ $C64(3\times3\times255)$

- Input O_{wd} and remove one layer skip connection and print: $C64(3\times3\times254)$
- Input $C64(3\times3\times254)$ and remove one layer skip connection $C64(3\times3\times256)$
- If (accuracy of PSP-U-SegNet> accuracy of U-Net, SegNet and U-SegNet)

  Print→ highest accuracy, IoU, MIoU.

  elseif implement U-Net

  elseifSegNet

  else U-SegNet

print→ highest accuracy, IoU, MIoU.

- Stop

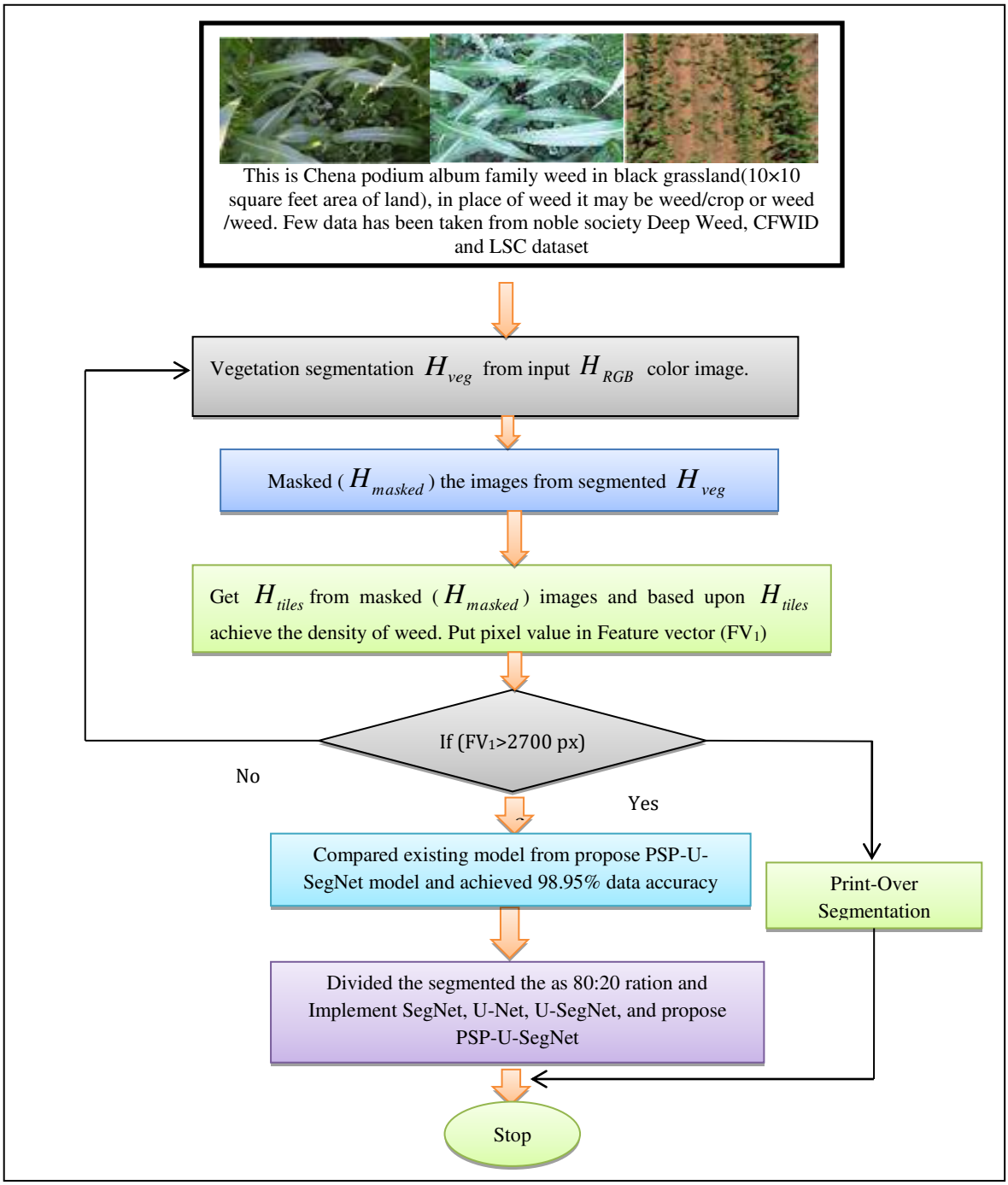


Figure7. Flow chart of the proposed model

4.4. Data classification using Proposed Model

This work has trained three existing CNN model such as U-Net, SegNet, U-SegNet and proposed model PSP-U-SegNet.The learning rates are slow in U-Net, SegNet, and U-SegNet CNN having a slow

learning rate due to deeper intermediated layers. The proposed PSP-U-SegNet model has been ignored over the deeper intermediate layer. We solve this problem by offering a global prior representation that is both effective and efficient which is discussed in the next subsection.

4.4.1. PSP-U-SegNet(Pyramid Scene Parsing Network U-SegNet)

The proposed PSP–U-SegNet (PSP-U-SegNet) model has included the functionality of PSP-Net, U-Net, and SegNet models. It has a total of 83 Convo layers, which has 25 convo layers of PSP-Net, 16 convos up and 16 convos down sampling layers from U-Net, andtheremaining 26convo layers from the SegNet CNN model are included. The proposed model has included input, convolutional, softmax, Up-sampling, and Max pool layer. Further, replacing the three max pool layers out of 5 layers from up-sampling layers of the pyramid.Finally, softmax layers will generate a final result of image classification. This work has 83Conv, 5 max pools, and 5 Up-sampling layers applied in the hybrid SegNet CNN model[29]. After pre-processed image, $(w \times h \times 3)$ has input for the proposed CNN model. The morphological feature map of weeds of an image has achieved by the proposed model. The certain scale of image feature map has been reduced using Max–pooling layer and up-sampling process. The final result has been shown after processing the soft-max layer in pixel-wise data representation of each class and creating the pyramid. The proposed model has shown a "U" shape [19]. Initially, U-Net was invented for biological-image segmentation but it has also achieved good results in other industries.

There are two main reasons for use of this U-Net and SegNetCNN model. Firstly, it can extract exhaustive features from local information through convolution layers. Secondly, it will provide the best accuracy for the limited number of samples. The classical U-Net and SegNetmodel has large consumption of calculation resources and speed was slow, therefore the proposed model has simplified these factors. This work has very similar to the SegNet CNN model for image segmentation using the skip connection method. The skip connection method has been lost using up-sampling of the bottom layer in theSegNet CNN model. The classical SegNet model is the skeleton of the proposed model. There is more time consumption of pooling in the basic SegNet model. Therefore, it's mandatory to reduce the number of pooling layers at first. This work has been performed by the Skip Connection Technique (SCT) in the SegNet CNN model. This technique arranged spatial information at the same level after using the up-sampling bottom layer. The Batch Normalization (BN)was added in the final stage in the convolutional layer to guarantee data stability[30].

This paper has proposed a PSP-U-SegNet model with a skip connection method with a unified parameter (kernel Size 3, 3) of the convolution layer. This work has used kernel size, padding, and activation functions. There are 3, 3 is kernel size, for padding used 0 in the outer ring of image. The ReLu activation function used 0, Conv 64, ConV128 mask, and finally, the kernel size of the outer layer has (1, 1). Furthermore, the sigmoid function handles the binary (0~1) image segmentation problem. The complete steps of proposed model are described in figure 7.

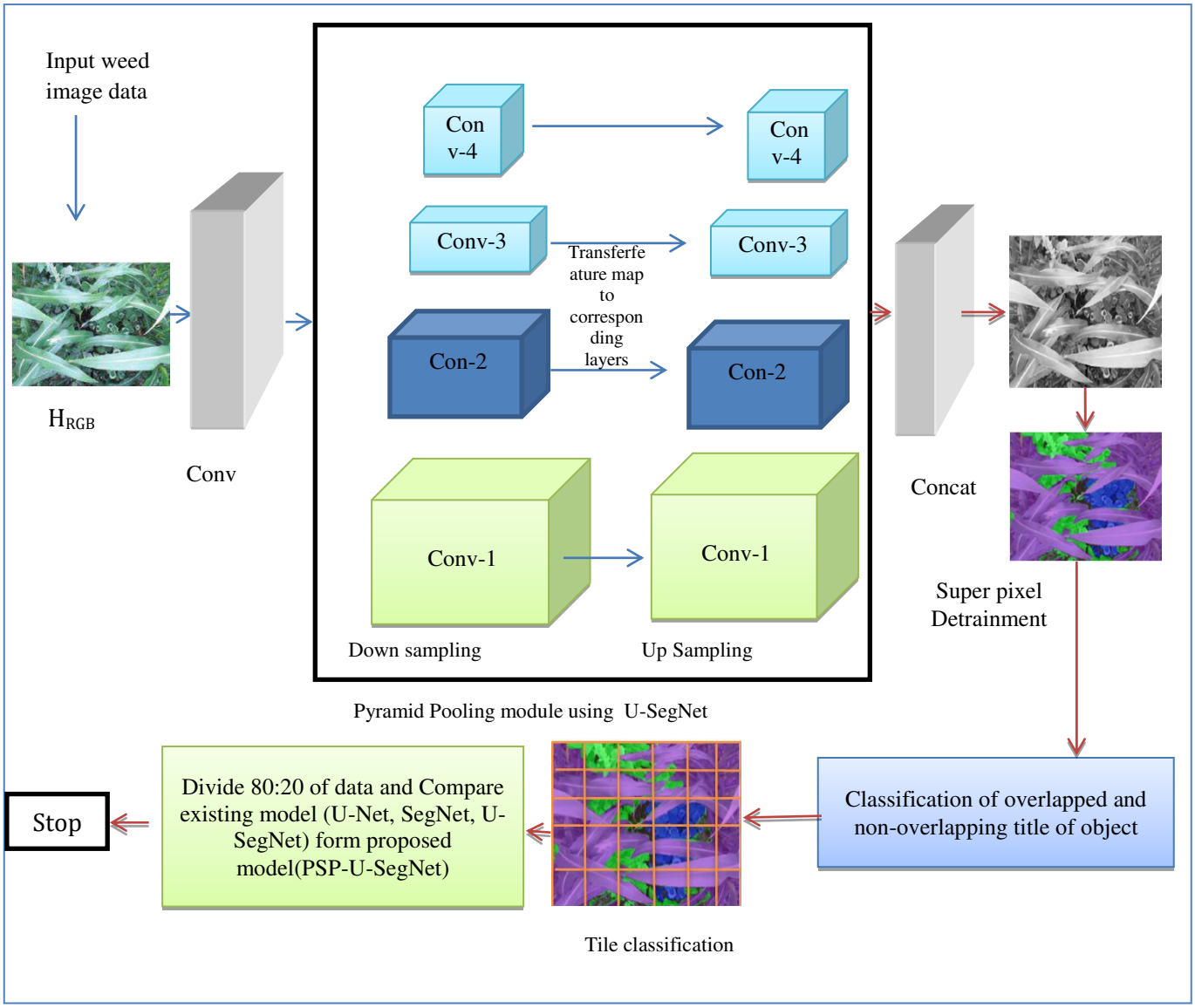


Figure 7. Proposed PSP U-SegNet CNN model

This work has used cascade representation of object detection. Moreover, it will demonstrate the successful non-trivial semantic segmentation object detection. In this work, the proposed model has changed the three max-pool layers out of 5 layers in a new framework of semantic segmentation. The max-pool information is proceeding before forward to the next stage and finally third is before executing the semantic segmentation of the weed object to explore contextual information. Overall changes are improving the flow and accurately achieving the object of the image. The detailed description is given below in figure 8.

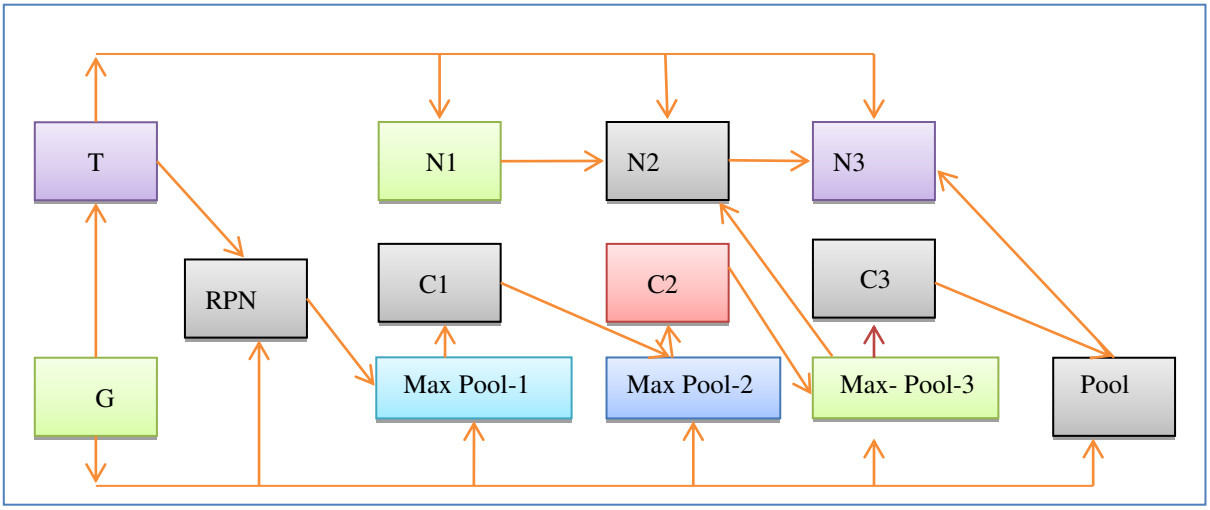


Figure 8.Modified Max pool network architecture of PSP-U-SegNet

Here 5 different max pool networks which are closely related to Region Proposal Network (RPN) and CNN feature (G, T). The RPN can parallel predict the object in semantic and vegetation segmented objects. Here C1, C2, and C3 are predicted masks of objects, andN1, N2, and N3 are bounding boxes. The bounding box and predicted mask has been shown below in Equations 7 and 8.

$$y_t^{box} = q(y, s_{t-1}), s_t = C_t(y_t^{box}) \quad \text{(eq.7)}$$

$$y_t^{mask} = q(y, s_{t-1}), N_t = N_t(y_t^{mask}) \quad \text{(eq.8)}$$

Where y is the backbone feature of CNN feature, y_t^{box} and y_t^{mask} is donated as bounding box and predicted mask feature. C_t, N_t is the box and mask head and t is a stage, and s_t, y_t is predicted box and mask head.

4.4.2 Interleave Execution of weed image

Processing of weed image object as two branches of bounding boxes in parallel execution in training stage (eq.1) and both two branches are not directly interacted within a stage. So it is mandatory to improve the architecture at N_{t-1} the head. The interleaved execution and mask information flow is expressed as equation 9 and 10.

$$y^{box}{}_t = \eta(y_1 y_{t-1}), s_t = C_t(y_t{}^{box}) \quad \text{(eq.9)}$$

$$y_t{}^{mask} = \eta(y, s_{t-1}), N_t = N_t(y_t^{mask}, N_{t-1}) \quad \text{(eq.10)}$$

Where N_{t-1} the intermediate object is a feature and $t-1$ is a stage of mask representation.

4.1.1. Object detection flow of Weed image data

Weed object detected using region of interest (RoI) future and it has been implemented before de-convolutional of data with the spatial size is 14×14. In stage's' forwarded all the mask head

with the use of RoIs and finally computed the object. Here 'F' is a function thatcombinesthe feature of the current stage and here $N_t(F(y_t^{mask}, N_{t-1})$ is a feature transformation function with four 3×3 convolutional layers. The objection detection has been done through the back propagation technique in equation 11.

$$y_t^{mask} = q(y, s_{t-1}), N_t = N_t(F(y_t^{mask}, N_{t-1})$$
$$F(y_t^{mask}, N_{t-1}) = y_t^{mask} + h_t(N_{t-1})$$
$$N_1 = N_1(y_t^{mask})$$
$$N_2 = N_2(F(y_t^{mask}, N_1)$$
$$\vdots$$
$$N_{t-1} = N_t(F(y_t^{mask}, N_{t-2})) \quad \text{(eq.11)}$$

This work has been directly combined with Mask R-CNN and Cascade R-CNN which is denoted as a Hybrid cascade Mask R-CNN.

4.4.3 *Learning the weed object using the proposed model*

This work presented the PSP-U-SegNet for semantic segmentation of weed/crop pictures.Figure 8 shows the different boxes and masks which has to interact with different branches. This work use RoI Align such as 7×7 and 14×14 feature map. Each stage is predicted by the box head and the entire mask head has predicted as the pixel-wise mask. The loss function takes the form of multi-task learning given in equations 12, 13, 14 and 15

$$M = \sum_{t=1}^{T} \beta(M_{cbox}^{t} + M_{mask}^{t}) + \lambda M_{seg} \quad \text{(eq.12)}$$

$$M_{cbox}^{t}(d_i, s_t, \hat{d}_t, \hat{s}_t) = M_{cls}(d_t, \hat{d}_t) + M_{reg}(s_t, \hat{s}_t) \quad \text{(eq. 13)}$$

$$M_{mask}^{t}(n_t, \hat{n}_t) = BCE(n_t, \hat{n}_t) \quad \text{(eq. 14)}$$

$$M_{seg} = CE(t, \hat{t}) \quad \text{(eq. 15)}$$

Here M_{cbox}^{t} cover the loss of the bounding box which has been predicted as the stage of t, and it has to combine as M_{cls} and M_{reg} which is defined as weed classification and bounding box regression. M_{mask}^{t} is denoted as a prediction mask in any stage of t which is called the binary cross-entropy (BCE). M_{seg} is used to balance the phases and tasks of segmentation. It is designated as semantic segmentation loss in the concept of cross-entropy.
This work is used by default $\beta = [1, 0.4, 0.24]$, $\lambda = 1$ and $t = 2$. The furthers section are discussed as evaluations of model

5. Training and Evolution of model

This work has taken 5090 data from various datasets such as Deep Weed, CWFID, and LSC datasets distributed on 80:20 ratios of data.The complete distribution of the dataset is given in table 2.

Table 2. Dataset distribution

Name of Dataset	Total	Training image	Testing Image
Deep Weed	5090	1600	120
CWFID		1800	200
LSC		1200	170

5.1.Performance of Vegetation segmentation of the model

The description of vegetation segmentation using the existing and proposed model is given in the next subsection.

i. Qualitative evaluation of model:

The result of vegetation segmentation using a few input images from three different datasets. Itcan be observed that PSP-U-SegNet outperformed the other model. After the discrimination of objects using vegetation segmentation, semantic segmentation prepares the homogeneous color object model with the same pixel intensity of the object. In observation, the “Deep weed” dataset has provided the finer object detection. The background and vegetation segmentation can finer detail of the vegetation of the object. It is also interesting the U-Net can identify tiny groupings of vegetation objects. Further, classify as asingle pixel of the object.It is because it prioritizes the spatial continuity of vegetation clusters, whereas U-Net tends to focus on a pixel's immediate surroundings. The CWFID dataset, which has weak contrast when compared to the LSC, showed a considerably stronger trend. The “Deep Weed” has a more prominent dataset using the PSP-U-SegNet classifier. The detailed vegetation in segmentation is shown in figure 9.

ii. Feature vector-based tile classification and effect of tile

As previously mentioned, the vegetation segmentation $H_{veg}$is used to detect the areas of vegetation in the pictures that contain crops and weeds. The output is a masked picture created by overlaying the input image H_{RGB} with H_{veg}. Then, non-overlapping tiles (sub-images) and Titles are separated from this masked picture. A pre-trained U-Net classifier is then used to retrieve the characteristics from each Title. Table 4 shows how well various classifiers perform when identifying terms such as "weed" or "crop" using these attributes. Take note of the enhancement in classifier performance brought on by weighted training utilizing various methods. By showing how sampling strategies (random sampling) to aid in enhancing the classifier performance for an imbalanced dataset, this study supports prior findings.The accuracy

and recall computed for the weed class on the test set are used to gauge performance. While sampling methods that account for class imbalance result in a relative improvement in the accuracy and recall values, the absolute values still fall below the acceptable cutoff. As shown in Table 8, the suggested model PSP-U-SegNet classifier obtained an accuracy of 96.98% and a recall of 97.98 percent. Every tile was expected to be covered in weeds. This highlights how these classifiers are unable to reliably distinguish between feature vectors produced by the suggested pipeline that correspond to agricultural and weed plants. Two observations served as the basis for the intuitive choice of tile size (a square with a side of 50 pixels) as primarily to either weed or agricultural plants rather than both, and (2) it prevented the creation of zones where virtually all of the pixels belonged to a cluster of vegetation. Due to how similar crop and weed plants would seem, there would not be sufficient descriptive information for the classifier to differentiate between them.

Nevertheless, the outcomes from regions of varied sizes were examined to justify the choice of tile size. In this study, we used both side length increases and decreases (100(px) and 25(px), respectively) to retrain the classification models.Classifiers trained using tiles of side lengths 25(px) and 50(px) perform better than those trained on 100100 tiles on average, taking into account both accuracy and recall values. Further, table 5 has shown the computation time by passing the tile processing. For patch sizes with sides of 50 and 100 pixels, computation time is comparable, while side lengths of 25 pixels resulted in a considerable increase in computation time. The explanation was that patches with sides longer than 25 pixels had a significantly higher percentage of tiles with vegetation pixel density than 10% compared to the previous two.

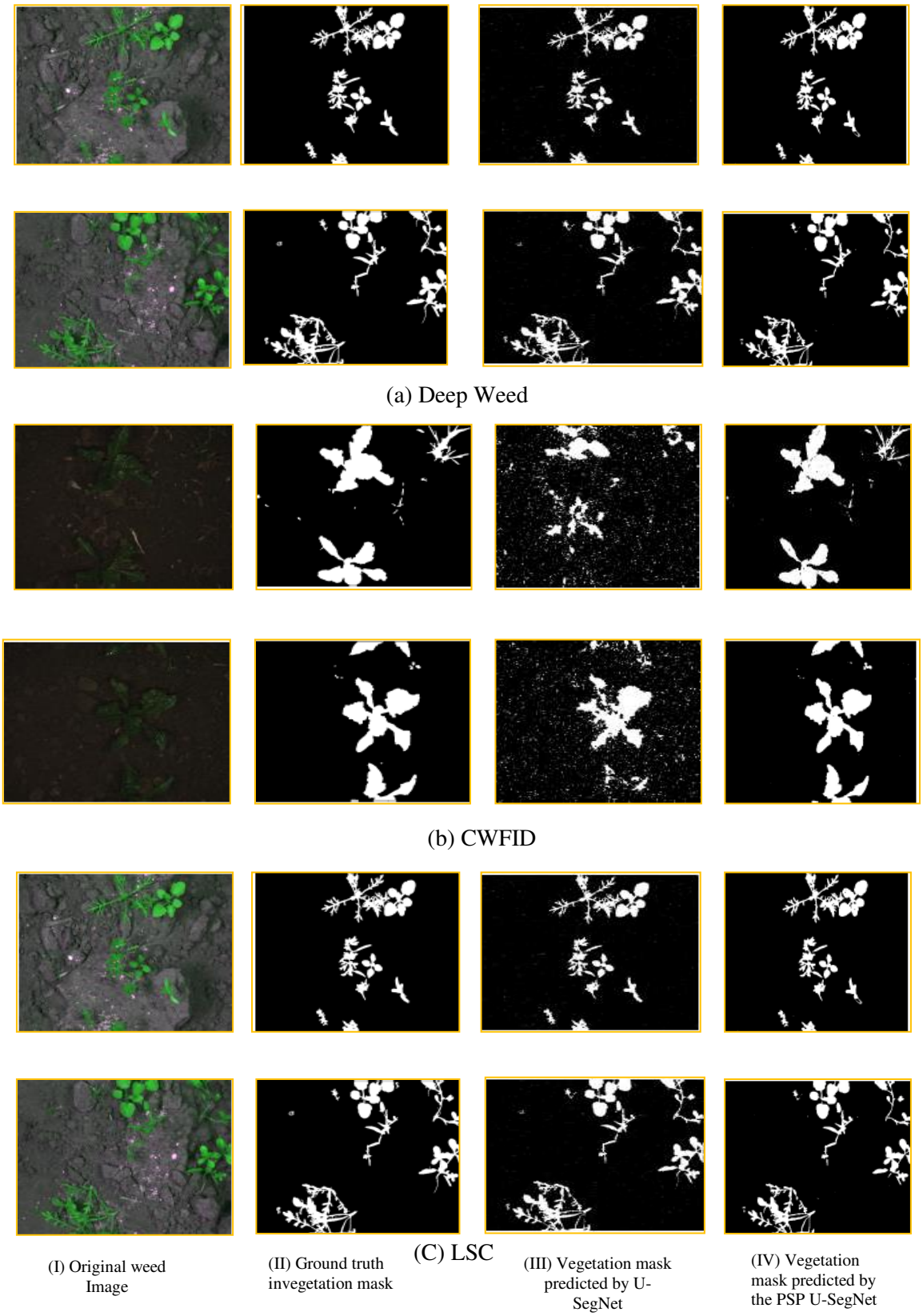


Figure 9. Vegetation mask of weed image(left to right): (I) Original weed Image, (II) Ground truth

vegetation mask, (III) Vegetation mask predicted by U-SegNet (IV) Vegetation mask

Table 3. Quantitativeevolution of data

Model Name	Dataset	mIoU
U-Net	Deep Weed	0.911
	CWFID	0.872
	LSC	0.902
SegNet	Deep Weed	0.922
	CWFID	0.772
	LSC	0.913
U-SegNet	Deep Weed	0.934
	CWFID	0.741
	LSC	0.921
PSP-U-SegNet	Deep Weed	0.961
	CWFID	0.802
	LSC	0.943

The proposed model PSP-U-SegNethas provided 0.961% of discrimination from vegetation and background segmentation of objects from an image. The other existing classifier as U-SegNet provides 0.92%, and SegNet has 0.91% for the LSC dataset.

iii. Comparison of pixel-wise dense predictions

The patch-wise predictions may be utilized to provide accurate pixel-wise weed/crop segmentation, even though it is not the suggested method's main goal. Therefore, we compare the anticipated ground coverage's accuracy using the F1 score measure (Equation 8). End-to-end segmentation networks were suggested by the authors of [22] and [23] for predicting dense crop/weed maps on the Deep weed, CWFID, and LSC datasets. The maximum-minimum value for the class of weeds is (0.41, 0.43) precision value in Deep weed dataset in $50x50$ tile classification. The CWFID and LSC weed class has 0.39 and 0.75 and 0.42 and 0.34 precision values respectively. In observation, the CWFID dataset has more accurate than the other dataset. Another parameter as F1-score has been reported as a maximum of 0.28, 0.36, and 0.28 in $50x50(px)$for Deep weed, CWFID, and LSC datasets respectively. Our method falls short in terms of pixel-level precision in comparison (the maximum F1 value for weed class is 0.36) in the CWFID dataset. The complete pixel data segmentation is given in table 4.

Table 4. Pixel-wise data segmentation

Dataset	classifier	$25x25(px)$			$50x50(px)$			$75x75(px)$		
		Precision	Recall	F1-Score	Precision	Recall	F1-Score	Precision	Recall	F1-Score
Deep Weed	U-Net	0.25	0.13	0.15	0.41	0.33	0.02	0.36	0.49	0.55
	SegNet	0.11	0.03	0.04	0.43	0.02	0.04	0.34	0.33	0.35
	U-SegNet	0.34	0.33	0.35	0.22	0.07	0.11	0.34	0.33	0.35
	PSP-U-SegNet	0.31	0.16	0.21	0.43	0.07	0.28	0.34	0.43	0.38
CWFID	U-Net	0.31	0.13	0.18	0.39	0.22	0.28	0.34	0.43	0.38
	SegNet	0.16	1	0.28	0	0	-	0.35	0.24	0.29
	U-SegNet	0.32	0.3	0.31	0.3	0.52	0.38	0.42	0.21	0.28
	PSP-U-SegNet	0.29	0.07	0.11	0.75	0.07	0.36	0.36	0.16	0.22
LSC	U-Net	0.35	0.36	0.36	0.42	0.21	0.28	0.37	0.14	0.2
	SegNet	0.32	0.5	39	0.27	0.42	0.33	0.31	0.29	0.3
	U-SegNet	0.32	0.16	0.21	0.28	0.09	0.14	0	0	-
	PSP-U-SegNet	0.44	0.13	0.2	0.34	0.39	0.36	0.29	0.67	0.41

However, a method to choose particular regions must be added to the segmentation networks to selectively treat specified parts. There will inevitably be an overlap of weed and crop pixels for the majority of the tiles if it is separated into sections like square tiles. The dominating label for such tiles will be used to determine how to handle a certain area. As a result, the selective treatment is unaffected by correctly recognized pixels that are in the minority for a specific tile. The computation time of tile processing is given in table 5.

Table 5. Computation time of tile processing

Tile size length	75x75 (px)	50x50 (px)	(25x25) px
computations time	0.94	0.9	6.54

We contend that the suggested method places more emphasis on accurately identifying the treatment regions than it does on correctly identifying such pixels. Additionally the enormous data needs of the suggested technique are far lower than among an end-to-end segmentation network, which enhances generalisation and scalability. The suggested method may also be applied to any crop-weed combination because it does not need the creation of custom features. The value loss via cross-entropy is displayed in table 6.

Table 6. Value loss using Cross-Entropy

Dataset	Binary Cross Entropy Loss						Weight cross Entropy Loss					
	PR0	RE0	F1 S0	PR1	RE1	F1 S1	PR0	RE0	F1 S0	PR1	RE1	F1 S1
Deep Weed	0.5	0.6	0.79	0.2	0.7	0.44	0.8	0.5	0.74	0.31	0.91	0.46
CWFID	0.7	0.6	0.80	0.3	0.8	0.44	0.9	0.6	0.76	0.32	0.92	0.47
LSC	**0.9**	**0.8**	**0.82**	**0.4**	**0.9**	**0.46**	**1**	**0.7**	**0.77**	**0.34**	**0.93**	**0.49**

The soft-max layer of the proposed PSP-U-SegNet model has checked the cross-entropy and weight-cross entropy loss of images. This work used three datasets Deep Weed, CWFID, and LSC, out of these the Deep weed has 0.5, precision and in the case of weight cross-entropy, the minimum precision has 0.8.The CWFID and LSC have maximum value loss.

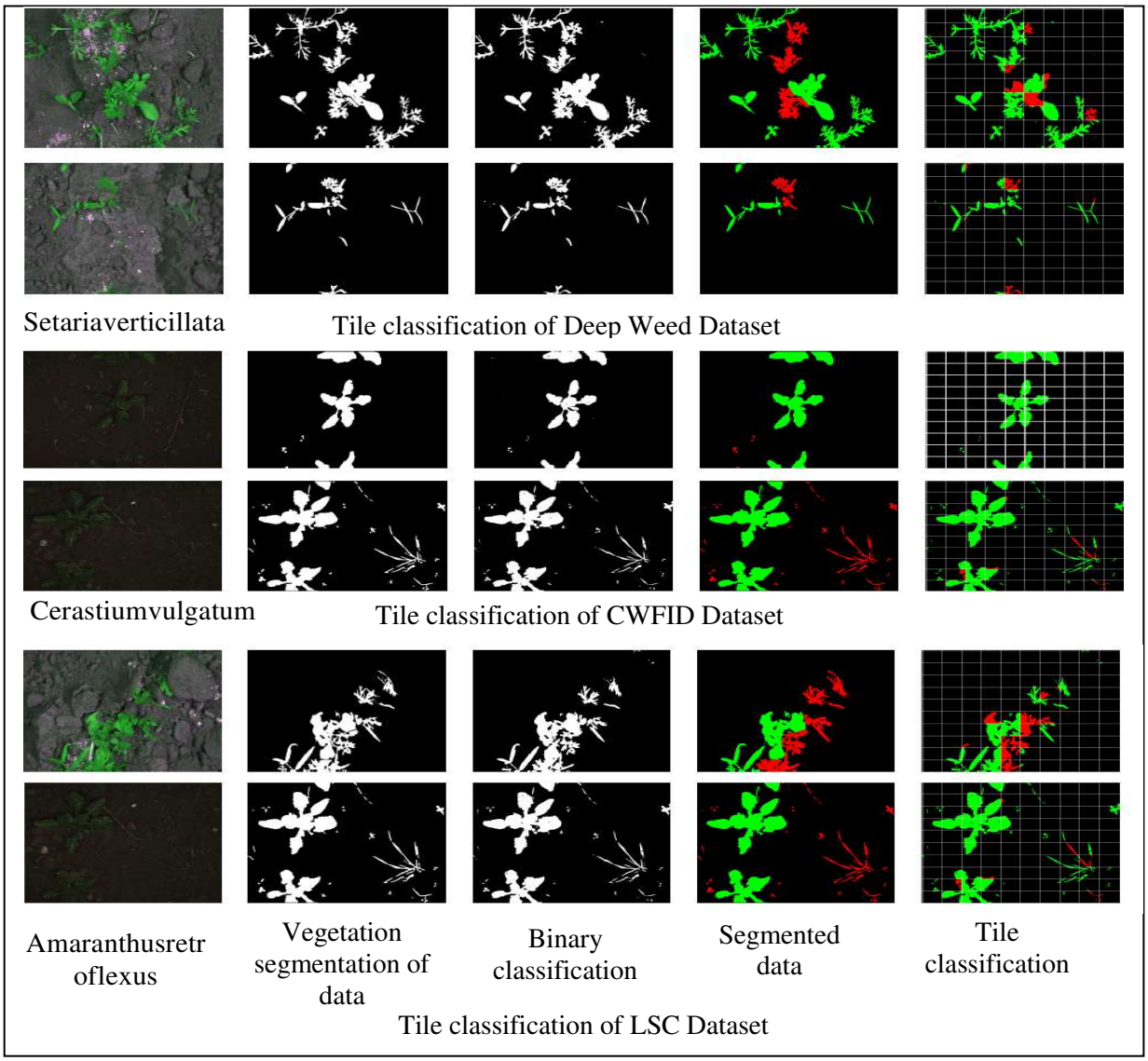


Figure 10. Vegetation mask of data

This work has estimated the weed object based on tile classification of "*AmaranthusRetroflexus*" weed data. After analyzing the vegetation segmentation and binary classification,the data has binary classified as a gray-scale image. For keen observation of the object, it has been segmented using tile classification. The classified object may be overlapped, and the partial or full object detected. The detected object is estimated by error rate which are given in table 7.

Table 7 The error rate estimation of image data

Dataset	Mean Accuracy	MAE	RMSE
Deep Weed	**82.13**	**1.62**	**2.06**
CWFID	75.24	5.02	7.85
LSC	82.13	1.62	3.06

Table 7 has summarized the error rate based upon MA, Mean Absolute Error (MAE), and Root Mean Square Error (RMSE) of Deep weed, CWFID, and LSC datasets. After observation, the Deep weed dataset has less great MA which has 82.13 and 1.62, and 2.06 has an error rate of MAE and RMSE. The performance of the model has given in table 8.

Table 8. Performance of Existing model to the proposed model

Dataset	Classifier	Precision	Recall	F1-Score	Accuracy
Deep Weed	U-Net	86.97	87.45	87.97	89.93
	SegNet	87.65	86.54	88.90	90.98
	U-SegNet	89.76	90.09	90.56	91.90
	PSP-U-SegNet	96.98	97.98	97.88	98.96
CWFID	U-Net	84.97	85.45	85.97	90.90
	SegNet	88.65	87.54	87.90	93.87
	U-SegNet	91.76	89.99	90.56	95.97
	PSP-U-SegNet	96.98	95.98	95.88	97.99
LSC	U-Net	88.97	87.45	85.97	84.23
	SegNet	89.65	88.54	89.90	85.45
	U-SegNet	89.79	90.09	90.56	93.67
	PSP-U-SegNet	90.98	92.98	91.88	93.98

In the Deep weed,a dataset the proposed model has achieved 96.98%, 97.98%, and 98.96% Precision, Recall, and data Accuracy respectively using the proposed model PSP-U-SegNet. The existing model U-Net classifier has achieved 89.93%, 90.90%, and 84.23 % data accuracy. Another existing CNN model (SegNet and U-SegNet) has achieved 90.98%, 93.87%, and 85.45% data accuracy which has less accurate wih proposed model.

6. Conclusion

A farmland management strategy to sustainably increase productivity and revenues is known as precision agriculture. In addition to being bad for the environment, agrochemicals like weedicides are an expensive input for farming. It may be possible to drastically lower their usage by using a computer vision system to locate areas that need specific chemical treatment. In order to support precision agriculture, a PSP-U-SegNet technique to robustly predict the weed density and dispersion is provided. The suggested method only accepts color photos as input. The first step is to construct a binary vegetation mask by removing every background pixel. Precision agriculture is an approach to agricultural management that tries to gradually increase yield and revenue. In addition to being harmful to the environment, agrochemicals like weedicides are an expensive input for farming. It might be possible to drastically lower their use by using a computer vision system to locate areas that need specific chemical treatment.

A PSP-U-SegNet approach to accurately estimate weed density and dispersion is offered to enhance precision agriculture. The recommended approach only takes input from color photographs. Making a binary vegetation mask in the first stage entails erasing every backdrop pixel. A maximum recall of 0.99% is used to identify weed-infested areas, and an accuracy of 82.13% is used to assess their weed density. Reducing reliance on heavily annotated datasets is one of the main goals of our research.

As long as a platform is in place to get top views of the plants, there is no need to spend money on any additional sensors beyond a standard RGB camera. The ongoing process of creating vegetation masks is one of our work's constraints. Future research should aim to reduce the average number of iterations required by the unsupervised network to build the vegetation mask.

This would decrease how long it takes an autonomous robot to process a single color image. Another potential direction for this study is to use the two-stage detection and localization method to medical imaging for the diagnosis of diseases or lesions. Another is weed detection in mixed crop field in another challenging task for weed identification and detection. To broaden the use of precision agriculture, a similar strategy may be used for diagnosing crop diseases.

Declarations

Ethical Approval: Not applicable

Conflict of Interest: The authors declare that they have no conflict of interest

Authors' Contributions: AMM: Conceptualization, data collection and Methodology, Writing, PK: Methodology, Conceptualization and Supervision, MPS: Validation and supervision, SPS: validation. All authors have read and agreed to the published version of the manuscript.

.

Funding Information: Not applicable

Data Availability Statement: Data and source codes are available from the authors upon reasonable request

References

[1] W. Kazmi, F. J. Garcia-Ruiz, J. Nielsen, J. Rasmussen, and H. Jørgen Andersen, "Detecting creeping thistle in sugar beet fields using vegetation indices," *Comput. Electron. Agric.*, vol. 112, pp. 10–19, Mar. 2015, doi: 10.1016/j.compag.2015.01.008.

[2] Y. Lu and S. Young, "A survey of public datasets for computer vision tasks in precision agriculture," *Comput. Electron. Agric.*, vol. 178, no. May, p. 105760, 2020, doi: 10.1016/j.compag.2020.105760.

[3] R. Yasrab, J. Zhang, P. Smyth, and M. P. Pound, "Predicting Plant Growth from Time-Series Data Using Deep Learning," *Remote Sens.*, vol. 13, no. 3, p. 331, Jan. 2021, doi: 10.3390/rs13030331.

[4] R. Kamath, M. Balachandra, and S. Prabhu, "Crop and weed discrimination using laws' texture masks," *Int. J. Agric. Biol. Eng.*, vol. 13, no. 1, pp. 191–197, 2020, doi: 10.25165/j.ijabe.20201301.4920.

[5] S. M. Sharpe, A. W. Schumann, and N. S. Boyd, "Detection of Carolina Geranium (Geranium carolinianum) Growing in Competition with Strawberry Using Convolutional Neural Networks," *Weed Sci.*, vol. 67, no. 2, pp. 239–245, 2019, doi: 10.1017/wsc.2018.66.

[6] P. Lottes, J. Behley, A. Milioto, and C. Stachniss, "Fully convolutional networks with sequential information for robust crop and weed detection in precision farming," *IEEE Robot. Autom. Lett.*, vol. 3, no. 4, pp. 2870–2877, Oct. 2018, doi: 10.1109/LRA.2018.2846289.

[7] M. Qian, I. McLaughlin, W. Quo, and L. Dai, "Mismatched training data enhancement for automatic recognition of children's speech using DNN-HMM," 2017, doi: 10.1109/ISCSLP.2016.7918386.

[8] N. Colbach, A. Gardarin, and D. Moreau, "The response of weed and crop species to shading: Which parameters explain weed impacts on crop production?," *F. Crop. Res.*, vol. 238, pp. 45–55, 2019, doi: 10.1016/j.fcr.2019.04.008.

[9] J. Gao, A. P. French, M. P. Pound, Y. He, T. P. Pridmore, and J. G. Pieters, "Deep convolutional neural networks for image-based Convolvulus sepium detection in sugar beet fields," *Plant Methods*, vol. 16, no. 1, Mar. 2020, doi: 10.1186/s13007-020-00570-z.

[10] A. M. Mishra and V. Gautam, "Weed Species Identification in Different Crops using Precision Weed Management: A Review." Accessed: Apr. 17, 2021. [Online]. Available: https://niti.gov.in/national-strategy-.

[11] A. Muni Mishra *et al.*, "A Deep Learning-Based Novel Approach for Weed Growth Estimation," *Intell. Autom. Soft Comput.*, vol. 31, no. 2, pp. 1157–1173, 2022, doi: 10.32604/iasc.2022.020174.

[12] X. Ma *et al.*, "Fully convolutional network for rice seedling and weed image segmentation at the seedling stage in paddy fields," *PLoS One*, vol. 14, no. 4, Apr. 2019, doi: 10.1371/journal.pone.0215676.

[13] Ł. Chechliński, B. Siemiątkowska, and M. Majewski, "A system for weeds and crops

identification—reaching over 10 fps on raspberry pi with the usage of mobilenets, densenet and custom modifications," *Sensors (Switzerland)*, vol. 19, no. 17, 2019, doi: 10.3390/s19173787.

[14] P. Rasti, A. Ahmad, S. Samiei, E. Belin, and D. Rousseau, "Supervised image classification by scattering transform with application toweed detection in culture crops of high density," *Remote Sens.*, vol. 11, no. 3, Feb. 2019, doi: 10.3390/rs11030249.

[15] N. Teimouri, M. Dyrmann, P. R. Nielsen, S. K. Mathiassen, G. J. Somerville, and R. N. Jørgensen, "Weed growth stage estimator using deep convolutional neural networks," *Sensors*, vol. 18, no. 5, pp. 1–13, 2018, doi: 10.3390/s18051580.

[16] M. J. Kropff, L. A. P. Lotz, S. E. Weaver, H. J. bos, J. Wallinga, and T. Migo, "A two parameter model for prediction of crop loss by weed competition from early observations of relative leaf area of the weeds," *Ann. Appl. Biol.*, vol. 126, no. 2, pp. 329–346, 1995, doi: 10.1111/j.1744-7348.1995.tb05370.x.

[17] S. Haug and J. Ostermann, "A crop/weed field image dataset for the evaluation of computer vision based precision agriculture tasks," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8928, pp. 105–116, 2015, doi: 10.1007/978-3-319-16220-1_8.

[18] C. Potena, D. Nardi, and A. Pretto, "Fast and accurate crop and weed identification with summarized train sets for precision agriculture," *Adv. Intell. Syst. Comput.*, vol. 531, pp. 105–121, 2017, doi: 10.1007/978-3-319-48036-7_9.

[19] R. Kamath, M. Balachandra, and S. Prabhu, "Raspberry Pi as Visual Sensor Nodes in Precision Agriculture: A Study," *IEEE Access*, vol. 7, pp. 45110–45122, 2019, doi: 10.1109/ACCESS.2019.2908846.

[20] J. Ahmad *et al.*, "Visual features based boosted classification of weeds for real-time selective herbicide sprayer systems," *Comput. Ind.*, vol. 98, pp. 23–33, Jun. 2018, doi: 10.1016/j.compind.2018.02.005.

[21] R. Yasrab, J. Zhang, P. Smyth, and M. P. Pound, "Predicting plant growth from time-series data using deep learning," *Remote Sens.*, vol. 13, no. 3, pp. 1–17, 2021, doi: 10.3390/rs13030331.

[22] P. Kaur and V. Gautam, "Plant Biotic Disease Identification and Classification Based on Leaf Image: A Review," doi: 10.1007/978-981-15-9712-1_51.

[23] S. K. Marwat *et al.*, "Weeds of Wheat Crop and Their Control Strategies in Dera Ismail Khan District, Khyber Pakhtun Khwa, Pakistan," *Am. J. Plant Sci.*, vol. 04, no. 01, pp. 66–76, 2013, doi: 10.4236/ajps.2013.41011.

[24] C. Potena, D. Nardi, and A. Pretto, "Fast and accurate crop and weed identification with summarized train sets for precision agriculture," in *Advances in Intelligent Systems and Computing*, 2017, vol. 531, pp. 105–121, doi: 10.1007/978-3-319-48036-7_9.

[25] S. Shorewala, A. Ashfaque, R. Sidharth, and U. Verma, "Weed Density and Distribution Estimation for Precision Agriculture Using Semi-Supervised Learning," *IEEE Access*, vol. 9, pp. 27971–27986, 2021, doi: 10.1109/ACCESS.2021.3057912.

[26] J. A. Vayssade, G. Jones, C. Gée, and J. N. Paoli, "Pixelwise instance segmentation of leaves in dense foliage," *Comput. Electron. Agric.*, vol. 195, p. 106797, Apr. 2022, doi: 10.1016/J.COMPAG.2022.106797.

[27] N. R V, A. Krishnan, V. Krishnan K, H. P, and S. A. Haritha Z A, "Southern Pea / Weed Field Image Dataset for Semantic Segmentation and Crop / Weed Classification using an Encoder-Decoder Network," *SSRN Electron. J.*, Feb. 2021, doi: 10.2139/ssrn.3781351.

[28] P. Kaur, S. Harnal, R. Tiwari, S. Upadhyay, S. Bhatia, and A. Mashat, “Recognition of Leaf Disease Using Hybrid Convolutional,” 2022.
[29] A. M. Mishra and V. Gautam, “Weed species identification in different crops using precision weed management: A review,” *CEUR Workshop Proc.*, vol. 2786, no. February, pp. 180–194, 2021.
[30] I. Sa *et al.*, “WeedNet: Dense Semantic Weed Classification Using Multispectral Images and MAV for Smart Farming,” *IEEE Robot. Autom. Lett.*, vol. 3, no. 1, pp. 588–595, Jan. 2018, doi: 10.1109/LRA.2017.2774979.