Classification of paddy crop and weeds using semantic segmentation

Radhika Kamath, Mamatha Balachandra, Amodini Vardhan and Ujjwal Maheshwari

Cogent Engineering (2022), 9: 2018791
Classification of paddy crop and weeds using semantic segmentation

Radhika Kamath¹, Mamatha Balachandra²*, Amodini Vardhan² and Ujjwal Maheshwari²

Abstract: Weeds are unwanted plants in a farm field and have harmful effects on the crops. Sometimes rigorous weeds bring down the crop yield significantly, causing huge losses to farmers. A prevalent method of controlling weeds is the use of chemical herbicides. These herbicides are known to cause harmful effects on our environment. One of the ways to control the ill effects of herbicides is to follow the Site-Specific Weed Management (SSWM). Site-specific weed management is to use the right herbicide for the right amount on agricultural land. This paper investigates a semantic segmentation approach to classify two types of weeds in paddy fields, namely sedges and broadleaved weeds. Three semantic segmentation models such as SegNet, Pyramid Scene Parsing Network (PSPNet), and UNet were used in the
segmentation of paddy crop and two types of weeds. Promising results with an accuracy over 90% has been obtained. We believe that this can be used to recommend suitable herbicide to farmers, thus contributing to site-specific weed management and sustainable agriculture.

**Subjects:** Computer Engineering; Computer Science (General); Information & Communication Technology (ICT); Computing & IT Security

**Keywords:** computer vision; deep learning; semantic segmentation; precision agriculture; weed management

1. Introduction

Weeds can be defined as undesirable plants growing with crops. Generally, they are referred to as plants out-of-space. They compete with crops for soil nutrients and water. Competition occurs when there is more plant demand for soil moisture, sunlight, and carbon dioxide, and demand exceeds supply. The possible outcome of this competition is the development of a characteristic crop-weed association. That is, crops and weeds grow and mature under mutual suppression. More often, it is the crop that is suppressed in the absence of an effective weed management strategy. This competition usually does not result in the death of the rice crop but definitely leads to reduced yield. Many of us know about weeds, but the problem caused by weeds and the importance of weed management is not really recognized by the general population and sometimes even by agriculturists. One of the main reasons for agriculturists to ignore weeds is the high cost involving manual weeding and expensive herbicides. Other than competing with rice plants for soil nutrients, water, and space, they also act as an alternate host to various pests, which in turn attack crops and destroy them.

Paddy (*Oryza sativa* L.) is one of the most important crops of India and of many parts of the world like China, Japan, Thailand, Sri Lanka etc. Even though paddy is grown in most parts of India, its yield and output are comparatively low. The main reason for this is pests like weeds and improper weed management. Weeds are the most severe and main biological constraints to crop production in India. Weeds are the main reason for heavy yield losses and sometimes are responsible for complete crop failure (Parameswari & Srinivas, 2017). It is reported that annually India is incurring a loss of INR 1050 million because of weeds in paddy fields (Gharde et al., 2018).

Weeds in paddy fields can be broadly categorized into three types. They are broadleaved weeds, grasses, and sedges (Subudhi et al., 2015; Sureshkumar et al., 2016; Naidu, 2012) as shown in Figure 1a, 1b, and 1c, respectively.

The standard way of handling weeds in India is hand weeding, mechanical weeding, and herbicides. Hand weeding involves pulling the weed plants with their roots or using tools like sickle, hoe, or spade. Hand weeding is a time-consuming and labor-intensive job. Moreover, in the coastal Karnataka region, there is an acute shortage of agricultural laborers. Therefore, because of the reduced availability of laborers and costly labor, hand weeding is not preferred in these regions. Mechanical weeding is carried out using a machine called a rotary weeder. This is not effective in paddy fields because of the complex and difficult field conditions for paddy. In addition, it seems impractical for the directly seeded paddy fields to use mechanical weeders because there are no crop rows in direct-seeded rice or paddy fields.

Herbicides are chemicals that are capable of killing some plants (weeds) without significantly affecting other plants (crops). Herbicides have many ill effects on the environment. Most of the farmers in India do not have basic education, and they do not have any knowledge about site-specific treatment in agriculture. Therefore, most of the time, any type of weed is treated with a broadcast application of herbicides. This method of excessive use of herbicides has resulted in the contamination of the groundwater and herbicide-resistant weeds. In addition, it has been shown that the farmer’s expenditure can be reduced by 40% if the suitable herbicide is used in the right
amount at the right time (Nistorup Jørgensen et al., 2007). Moreover, most automated weed recognition techniques are only able to discriminate between crop and weed. When controlling weeds with herbicides, it is important to know the species of the weeds so that the right herbicide is applied. Spraying a specific type of herbicide via SSWM can help in reducing herbicide usage largely, and thus there will be less environmental pollution and increased profits to the farmers. Therefore, it would be of great help to the farmers if there were automation of detection and identification of weeds in paddy fields in precision agriculture. This research work shows the feasibility of an approach for detecting and identifying two types of weeds and paddy crop itself using semantic segmentation.

1.1. Semantic segmentation in weed detection and identification

Deep learning has been used in many fields of agriculture and has been developed into a powerful method for image classification (Kamilaris & Prenafeta-Boldú, 2018). Deep Learning is a subset of Machine Learning. It extends classical machine learning by adding more complexity or depth and uses deep neural networks or multiple layers of neurons to extract features progressively from raw input. By allowing various transformation functions to transform the data, the data can be represented more hierarchically while at the same time maintaining abstraction of the data (Schmidhuber, 2015; Bengio et al., 1995). Furthermore, by mapping the consistency of features through a feed-forward process and optimizing the regularized loss function from the backpropagation process, deep learning networks are expected to recognize the patterns hidden in the dataset and automatically create feature maps to extract features from the dataset. The model then uses the recognized knowledge from pattern learning to perform various tasks like classification, segmentation and so on.

The goal of semantic segmentation is to perform pixel-wise classification. Each pixel is assigned a class from a pre-defined set of classes. Fully convolutional networks (FCNs) are generally used for semantic segmentation. These FCNs are made up of algorithms, which learn features automatically and build forward and reverse processes in an end-to-end manner. These often include encoder-decoder scheme. In recent years, FCNs have been used in various fields such as remote sensing, medical image analyses, and other computer vision-based applications. It is also a popular technique used in weed detection and identification. Many research studies have used deep learning-based semantic segmentation for weed detection. Segmentation models like SegNet, and UNet are the most commonly used models for automatic weed detection and identification. Contributions of this paper are as follows:

- This paper demonstrates the detection and identification of two types of weeds in paddy crop using deep learning-based semantic segmentation.

- Result analysis and comparison of different segmentation models.

2. Related studies

Precise weed control has become important to reduce the ill effects of herbicides. Precise weed control is implemented through SSWM. SSWM contributes to enhancing crop production, however the development of robust weed detection algorithms that are accurate and consistent in variable crop conditions is still a major challenge. This section summarizes the important previous studies focused on crop and weed discrimination, detection, and classification using conventional machine learning algorithms and deep learning-based techniques.

2.1. Conventional machine learning techniques for weed recognition

In the early days of computer vision-based weed detection, many researchers used conventional machine learning techniques in combination with image features to accomplish the task of weed recognition. These techniques can work with smaller data sizes and are not computationally intensive. However, these techniques cannot be used for in-field real-time weed detection and identification tasks.

In (Ahmed et al., 2012), the weed and crop were classified using a Support Vector Machine (SVM) classifier. In this approach, a multi-class SVM was designed for classification to achieve the
Figure 1. Broad categories of weeds found in paddy fields.

classification of five types of weeds. The authors have used a one-against the rest method for SVM multi-class classification. In a one-versus rest approach, a binary classifier is designed to classify weeds, where one sample is identified from the rest of the samples. Feature selection was used to select the best performing nine features using the stepwise feature selection method, which is based on the forward-feature selection and backward-feature elimination method. In the stepwise feature
In (Gao et al., 2018), the discrimination between maize crop and three types of weeds was done using the Random Forest (RF) classifier. A total of 185 features based on reflectance indices and vegetative indices were extracted. Important features were selected using two methods, Principal Component Analysis (PCA), and RF classifier. While calculating hyperparameters for RF, three different combinations of features were used. They were (i) features extracted using PCA, (ii) features selected using RF, (iii) with all features. The RF with PCA extracted features performed very poorly when compared to the features selected using RF and total features. The RF classifier with features selected using RF had a classification precision of 95% for maize crop and for three types of weeds, C. arvensis, Rumex, and C. arvense, are 95.9%, 70.3%, and 65.9%, respectively. In identifying weeds such as Rumex and C. arvense, all the three models performed on average with a precision around 68%. However, in identifying the crop maize and other weed C.arvensis, all the models performed very well with the classification precision around 93%. The authors also report that the band near the red edge frequently appears in the 30 features selected by RF and thus has very high discriminative power.

In (Bakhshipour & Jafari, 2018), classification of sugar beet crop and four types of weeds was done using an SVM and Artificial Neural Network (ANN) classifiers using shape features. The performance of the SVM and ANN is evaluated and compared. A feed-forward multi-layer perceptron ANN is developed with two hidden layers and a Levenberg-Marquardt (LM) backpropagation learning algorithm. Tangent Sigmoid and logarithm sigmoid functions were used as transfer functions. For dimension reduction, PCA was used to reduce 31 original features into four components. The best network model was selected using MSE (mean squared error) and coefficient of determination (R2). The SVM with Radial Basis Function (RBF) kernel was used. Features reduced using PCA were also used in SVM. The 5-cross validation was used to find the optimal values for the penalty parameter (C), also known as regularization parameter, which indicates model’s tolerance
to the misclassification, and $\Gamma$ (gamma). The gamma parameter is the inverse of the standard deviation of the RBF kernel (Gaussian function), which is used as a similarity measure between two points. The correct identification of the weeds by ANN and SVM was 92% and 93% while for sugar beet crop was 93% and 96%. Even though the results obtained by both classifiers are almost the same, the authors conclude that SVM performs better with shape-based discrimination.

In (Ashok Kumar & Prema, 2016), the crop and weed discrimination was done using Relevance Vector Machine (RVM), SVM, and RF classifiers. Tamura texture features were extracted using the curvelet transform. The results showed that the RVM classifiers outperformed the other two classifiers with an accuracy of 99%. The authors also report that the testing time of the RVM classifier is less than that of the SVM classifier.

In (Santiago et al., 2019), the classification of sugar cane crop and different types of weeds found in sugar cane fields is done using the SVM classifier with the help of visual bag-of-features approach. In visual bag-of-words, the image is represented as a histogram representative of local features.

### 2.2. Deep learning-based image classification for weed detection and identification

In (Ma et al., 2019), the discrimination of rice seedlings and weed was done using the deep FCN, SegNet (Badrinarayanan et al., 2017). SegNet is based on encoder and decoder (codec) structure which has a lower computational cost and higher precision. There were three classes, namely soil, rice, and weeds. The structure of the SegNet is as shown in Figure 2. SegNet achieved a pixel-wise classification accuracy of 91% for soil, 94% for rice, and 94% for weeds. The classical semantic models such as UNet achieved an accuracy of 97% for soil background, 46% for rice, and 69% for weeds and FCN model had an accuracy of 83% for soil background, 92.1% for rice, and 92.9% for weeds (Xiaomeng et al., 2018), (Long et al., 2015).

In (Andrea et al., 2017), the maize crop and weed discrimination was done using a convolutional neural network (CNN). Four types of CNN models, namely LeNet, AlexNet, cNet, and sNet, were used, and their performance was analyzed. In addition, these models were also executed on different hardware set-up like normal CPU, Raspberry Pi 3 Model B, and with Nvidia Graphics Card. The cNet model outperformed all other models with 16 filters with an accuracy of 97%. The model, which ran on Nvidia Graphics Card, was supported CUDA platform executed 18 times faster than a normal CPU and 170 times faster than the model executed on the Raspberry Pi 3 Model B. Figure 3 shows the cNet architecture.

In (Liu & Bruch, 2020), weed detection in romaine lettuce crops was done using a deep learning model. Around 3000 images of lettuce crops and weeds were collected from an organic farm. The green vegetation was retained using Otsu’s thresholding (Otsu, 1979). The vegetation was labeled as crop and weeds semi-automatically using MATLAB. First, around 500 images were manually labeled. Then, a CNN model with YOLO-v2 (Redmon & Farhadi, 2017) was trained with these images to label the rest of the images. All the images labeled by the model were again manually inspected and corrected through adjustment. Then, the labeled images were used as training data for CNN models such as ResNet-50, ResNet-101, MobileNet, InceptionResnet V2, SqueezeNet, VGG16, VGG19 as the feature extraction layers for the YOLO-v2 model to identify lettuces. The...
Authors report that the CNN model with VGG16 outperformed all other models with a mean average precision of 93% with new data (20% test data, the data the model had never seen).

In (Inkyu et al., 2017), the classification of sugar beet and weeds has been done using a CNN. The images collected were converted to SegNet format and annotated. For model training, the maximum number of iterations was 640 epochs, the learning rate was set to 0.001, the batch size was 6, weight decay rate was 0.005. With the test data, an average accuracy of 80% was obtained by this model.
In (Milioto et al., 2018), the classification of sugar beet crops and weeds was done using CNN. The input is represented in 14 different formats based on indices such as Excess Green Index (EGI), Excess Red Index (ERI), etc., as shown in Figure 4. The input is resized, and channel-wise contrast normalization is done.

The CNN model with a rectified linear unit (ReLu) was used with three different image datasets captured at different places. The experiment was also carried out by inputting only RGB representations. The CNN model with input with the extra representations performed well with an accuracy of 95%. In addition, the authors report that inputting additional information results in speeding up the training process, and generalization capability of the model. Thus, it increases the overall performance of the model.

In (Jialin et al., 2019), weed detection in perennial ryegrass was done using deep learning techniques. Two datasets were constructed, one dataset with images containing a single weed species and another dataset with images consisting of multiple weed species. These two datasets were used to train deep learning models such as AlexNet, GoogLeNet, and VGGNet. The first dataset was a balanced dataset while the other was an unbalanced dataset. For the single species dataset AlexNet and VGGNet performed better when compared to GoogLeNet and also these models performed well with other datasets consisting of multiple weed species.

In (Hamza Asad & Bais, 2019), weed detection in the canola field has been done using maximum likelihood classification and deep learning techniques. As the first step, the soil background and vegetation are segmented using maximum likelihood classification. The pixels were then segmented as either belonging to a crop class or a weed class. Finally, deep learning techniques such as SegNet, UNet, and encoder architectures like VGG16 and ResNet-50 were used, and their
performance is compared. It was found that SegNet based on ResNet-50 performed well when compared to other architectures.

In (Fawakherji et al., 2019), the discrimination of crop and weeds was done using deep learning techniques using context-independent pixel-wise segmentation. The authors claim that this method is particularly useful for those data where object-annotated data is not available or is very less. The model used is UNet based on a modified VGG-16 encoder followed by a binary pixel-wise classification layer.

Chechlinski et al., (2019) discuss the implementation of crop and weed discrimination using a deep learning model and its deployment in the low-cost mobile Single-Board Computer (SBC)
Raspberry Pi 3 Model B+. The ground truth labels used were generated semi-automatically with the UNet model. The developed deep learning model involves the concepts of DenseNet, ResNet, and MobileNets. The model's deployment on the Raspberry Pi 3 Model B+ gave satisfactory results with detecting weed area up to 67% and misclassification of crop area by up to 0.9%. The crop and weed area segmentation was done from videos at more than 10 frames per second.

3. Methodology

3.1. Dataset preparation
Two digital cameras (Canon Power shot SD3500 IS and Sony DSLR) were used to acquire images in paddy fields around Manipal town, Karnataka State, India. The cameras were fixed at 2 ft., 3 ft., and 4 ft. above the ground (top-view) and only natural lighting conditions were used. All pictures
were taken in the year 2016, 2017, and 2018. The images consist of different growth stages of paddy, from one leaf stage to flowering stage. The images were stored in RGB color space with a resolution of 3456 × 2592 in JPG format. The MATLAB R2017b was used to pre-process the images before annotation. The images were resized to 1296 × 966 when working with MATLAB R2017b. The images acquired from the field presented various challenges before annotation due to complex background and complex plant objects. Since the paddy fields require standing water, we encountered problems such as reflections, shadows of plants, and other objects near the fields. So extracting only green plants becomes a challenge. To overcome this problem, we segmented the foreground object (plants) from the background using the YCbCr model as shown in Figure 5. Figures 6a and Figures 7a show sample images from the dataset, and Figures 8 and 12 show their corresponding images after soil and water background removal.

Paddy field images created by (Ma et al., 2019) were also considered in this research. The soil background of these image Figure 8b shows the image after removing water and soil background.

3.2. Dataset description

Paddy-weed images taken from the above said two different sources were combined. The dataset consists of four classes: paddy, two types of weeds—sedges and broadleaved weed, and background. While preprocessing two types of images were generated. The first set of images has only one type of class (plant) in the picture, that is, either paddy, or sedges or broadleaved weed. The second set of images consists of multiple classes in the same picture, for example, paddy with broadleaved weed or paddy with sedges. These images were divided into two types of datasets:

Dataset-1: It contains images of only weeds. The images in this dataset had only one kind of weed along with the background. This dataset is used to display how the model can differentiate between different kinds of weed. There are 875 images in the training set and 130 in the testing dataset.

Dataset-2: It consists of images of paddy with weeds. There are four classes in this dataset—paddy, broadleaved weed, sedges, and background. There are images with more than one class such as paddy with broadleaved weed or paddy with sedges. This dataset also has isolated images from the first dataset. This dataset shows a close resemblance to an actual paddy field, which generally has more than one type of weed in a field along with paddy crop. There are 1470 images in the training set and 220 images in the testing set.

Dataset-1 consists of images of one type of weed, either broadleaved weed or sedges. The intention of creating this dataset is to check the ability of the models to distinguish between weed types. Dataset-2 contains both images with only one weed type and also images with paddy along with two types of weeds.
Table 1. Results of Dataset-1

| Network | Frequency weighted IoU | Mean IoU | Class Wise IoU |
|---------|-------------------------|----------|----------------|
|         | Background | Broadleaved weed | Sedges |
| PSPNet  | 0.936112   | 0.723028 | 0.983647 | 0.693135 | 0.492301 |
| SegNet  | 0.823767   | 0.601746 | 0.942751 | 0.471455 | 0.391032 |
| UNet    | 0.742380   | 0.388569 | 0.821453 | 0.281921 | 0.262331 |

3.3. Models description

Semantic segmentation models are built upon a standard CNN network. Image segmentation models implemented in this research use the concept of encoder-decoder structure. The encoder downsamples the spatial resolution of images to develop a lower resolution feature mapping, which is used for classification into output labels. The decoder then upsamples the feature representation into a full resolution image again. For encoding classification, networks such as ResNet-50, AlexNet, VGG-16 (Kaiming et al., 2016; Krizhevsky et al., 2012; Simonyan & Zisserman, 2014) are used, and the decoder network is built on top of the encoder. In this research, three semantic segmentation models have been used—PSPNet, UNet, and SegNet.

In PSPNet, a segmentation architecture ResNet-50 base model was implemented. In PSPNet, a feature map is generated from the base network. This feature map is then downsampled to different scales. Convolution is implemented on these pooled feature maps and then feature maps are upsampled and concatenated together. A final convolution layer is applied to produce final segmented outputs. PSPNet is generally used for datasets in which objects are present in different sizes. In our dataset, as crop images are captured from different heights, objects are of different sizes.

The UNet architecture uses an encoder-decoder framework in which encoder-decoder layers are symmetrical to each other. UNet also uses skip connections. Skip connections are extra connections, which link upsampling layers with earlier downsampling layers. Skip connection helps in the reconstruction of segmentation boundaries after downsampling and hence produces a more precise output image. ResNet-50 base model was used in this segmentation architecture.

The SegNet model also uses an encoder network that is topologically identical to the 13 convolutional layers in VGG16 network. Each encoder layer has a corresponding decoder layer followed by a multi-class softmax classifier that produces class probabilities for each pixel. SegNet does not have any skip connections.

3.4. Implementation

Annotated images for the semantic segmentation models were generated using playment.io tool. Each RGB original image has a corresponding annotated image with the same name in png format. Each pixel of the annotated image was labeled with a class id: Background—0, Broadleaved weed—1, Sedges—2, Paddy—3. Thus, the entire annotated image has color between 0 and 3.

The Google Colab facility was used for the implementation of this research. Keras, Tensorflow, and image-segmentation-keras APIs were used for the implementation of different models. In the models, transfer learning was used to get better results. Weights obtained from ADE-20 K dataset (Zhou et al., 2017) were used as initial weights in the models. ADE-20 K dataset weights are trained on 150 different classes. Three semantic segmentation models—UNet, PSPNet, and SegNet, were used in this research. These models were implemented with different base models—ResNet-50 and VGG-16. In this research, PSPNet and UNet gave better results with ResNet-50 encoder and SegNet with VGG-16.
| Network   | Frequency weighted IoU | Mean IoU | Background | Broadleaved weed | Sedges | Paddy |
|-----------|-------------------------|----------|------------|------------------|--------|-------|
| PSPNet    | 0.89548                 | 0.624393 | 0.942821   | 0.581821         | 0.41156| 0.51137|
| SegNet    | 0.78291                 | 0.513497 | 0.91018    | 0.491223         | 0.31156| 0.381025|
| UNet      | 0.71934                 | 0.318804 | 0.81221    | 0.261874         | 0.242678| 0.358453|

Table 2. Results of Dataset-2

Kamath et al., Cogent Engineering (2022), 9:2018791
https://doi.org/10.1080/23311916.2021.2018791
Adaptive Momentum Optimization (Adam) optimizer was used in all the models. In addition, hyperparameter tuning was implemented for parameters such as batch size, learning rate and epochs. An initial learning rate of 0.001 was used along with a learning rate scheduler. Models were trained with different batch sizes (16,32,64) with SegNet and PSPNet giving better results with 32 batch size and UNet with 16. Initial epochs were set to 100 with an Early Stopping patience of 15 epochs.

3.4.1. Results
Intersection over Union (IoU) also known as the Jaccard Index, is the area of overlap between the ground truth and the predicted segmentation divided by the area of union between ground truth and predicted segmentation. The IoU is calculated individually for all classes and then averaged to calculate the mean IoU. Frequency weighted IoU is calculated for images where there is class imbalance. Since the proportion of background in our research is far more than classes such as broadleaved weed and sedges, class imbalance persists here. Thus, the best metric to compare models is frequency-based IoU because it calculates a weighted mean IoU based on the frequency of each class.

Pixel accuracy is another metric of evaluation for semantic segmentation models. It is calculated as the percentage of pixels in the image classified correctly. However, since pixel accuracy can provide misleading results in case of class imbalance, this metric was not used in this research. Hence, frequency weighted IoU metrics were used to compare segmentation models. The best results were obtained with PSPNet and hence the images shown below contain results from PSPNet. The PSPNet model is able to identify both the classes and the shape of the plant to some extent.

Table 1 shows the classification result of Dataset-1 and Table 2 shows the classification result of Dataset-2. From Table 1 and Table 2, we conclude that PSPNet performs better than SegNet and UNet models. The pooling in PSPNet is applied on the whole image and on the increasingly smaller regions. This allows the network to capture information not only at the global level but also at different object levels around the pixel of interest. Therefore, it learns more context relations. Thus, it gives good performance when there are objects at multiple scales. In (Ma et al., 2019), FCN was used to classify crop and weed with over 90% accuracy. No weed classification was done. The classification of paddy crop and weeds has been done by (Kamath et al., 2019), (Kamath et al., 2020), (Kamath et al., 2018) using conventional machine learning techniques like SVM, Random Forest classifiers, and multiple-classifier systems and obtained a decent accuracy around 80% to 90%. The drawback of classic machine learning techniques is that they cannot be used in real-time in-field weed segmentation problems. Semantic segmentation is a preferable method for real-time in-field weed segmentation problems.

Figures 9–14 show the results in images. In the images, sub-image at the left is the original image, sub-image at the center is the predicted output image, and the sub-image at the right is the ground truth image. Broadleaved weeds and paddy were identified well, but sedges weeds were not so well identified. This can be attributed to the similarities between paddy and sedges. Another reason could be the lesser number of samples for sedges weed type.

* In Dataset-2, paddy was represented in yellow color. Figure 15 shows identification of paddy and broadleaved weed.

4. Conclusion and future work
This paper investigated semantic segmentation models such as PSPNet, UNet, and SegNet for the segmentation of paddy crop and two types of weeds in paddy fields, namely, broadleaved weed and sedges type weed. The paddy field image dataset was taken from two different sources wherein the images were acquired under natural conditions. The soil and water background was removed to retain only green vegetation. The dataset was divided into two types, Dataset-1 and Dataset-2. The Dataset-1 consisted of only weed-type plants and
Dataset-2 consisted of images of weed along with paddy crop. The annotation of images were done using playment.io tool and each pixel is assigned to one of the classes among four classes, namely, Background-0, Broadleaved weed-1, Sedges-2, and Paddy-3. The results of this investigation were compared with classical machine learning techniques and segmentation models such as FCN. PSPNet performed better when compared to UNet and SegNet. The result is satisfactory, with the mean IoU around 70% to 80% and frequency weighted IoU is around 80% to 90%. This can be further improved by increasing the dataset size. This work can be used to recommend suitable herbicides to the farmers. We believe that this work encourages the development of in-field weeding machine/robots or in-field herbicide sprayer for paddy fields. Overall results indicate that deep learning-based semantic segmentation of paddy crop and weeds can be used for sustainable site-specific weed management and thus towards safe food production.

Future work is to expand the dataset and use various semantic segmentation models to analyze its performance.

Funding
The authors have no funding to report.

Author details
Radhika Kamath¹
E-mail: radhika.kamath@manipal.edu
Mamatha Balachandra²
E-mail: mamtha.bc@manipal.edu
ORCID ID: http://orcid.org/0000-0003-2201-8730
Amodini Vardhan³
E-mail: amodinivardhan@gmail.com
Ujjwal Maheshwari²
E-mail: ujjwalmaheshwari@gmail.com
¹ Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India.
² Institute of Computer Science, Tartu Ulikool, ESTONIA.

Disclosure statement
No potential conflict of interest was reported by the author(s).

Author Statement
Our group’s key research interest includes using computer vision-based techniques for weed detection and identification. It involves the creation of a digital image dataset of crop and weeds. We also work on publicly available crop and weed image datasets. We also work on building prototypes for automatic weed detection and identification using IoT boards. The aim of our research is to explore new technologies in agriculture to develop a sustainable and ecofriendly decision-making model that will help farmers to make the right decision at the right time.

In the paper “Classification of Paddy crop and weeds using semantic segmentation.” We discussed on different uses of different deep learning models to classify paddy crop and weeds. This research on further extension can be used to identify the weed area and depending on the density and type of the weed, a suitable herbicide in the required amount can be recommended. Thus, resulting in the reduced usage of herbicide and reducing the cost burden of the farmers.

Citation information
Cite this article as: Classification of paddy crop and weeds using semantic segmentation, Radhika Kamath, Mamatha Balachandra, Amodini Vardhan & Ujjwal Maheshwari, Cogent Engineering (2022), 9: 2018791.

References
Ahmed, F., Al-Mamun, H. A., Hussain Bari, A. S. M., Hussain, E., & Kwan, P. (2012). Classification of crops and weeds from digital images: A support vector machine approach. Crop Protection, 40, 98–104. https://doi.org/10.1016/j.cpprotection.2012.04.024
Andrea, C.-C., Barreno Mauricio Daniel, B., & Barriuonuevo José Misaol, J. Precise weed and maize classification through convolutional neuronal networks. In 2017 IEEE Second Ecuador Technical Chapters Meeting (ETCM), 1–6. IEEE, 2017.
Ashok Kumar, D., & Prema, P. (2016). A novel approach for weed classification using curvelet transform and Tamura texture feature (ccttf) with rvm classification. International Journal of Applied Engineering Research, 11(3), 1841–1848 http://www.scopus.com/inward/record.url?eid=2-s2.0-84991310024&partnerID=MN8TOARS.
Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(12), 2483–2495. https://doi.org/10.1109/TPAMI.2016.2644615
Bakhshipour, A., & Jafari, A. (2018). Evaluation of support vector machine and artificial neural networks in weed detection using shape features. Computers and Electronics in Agriculture, 145, 0 153–160. https://doi.org/10.1016/j.compag.2017.12.032
Bengio, Y., LeCun, Y., Noh, C., & Burges, C. (1995). Lerec: A nn/hmm hybrid for on-line handwriting recognition. Neural Computation, 7(6), 0 1289–1303. https://doi.org/10.1162/neco.1995.7.6.1289
Chechilinski, Ł., Siemiatkowska, B., & Majewski, M. (2019). A system for weeds and crops identification—reaching over 10 fps on raspberry pi with the usage of mobilenets, densenet and custom modifications. Sensors, 19(17), 3787. https://doi.org/10.3390/s19173787
Fawakherji, M., Youssef, A., Bloisi, D., Pretto, A., & Nardi, D. Crop and weeds classification for precision agriculture using context-independent pixel-wise segmentation. In 2019 Third IEEE International Conference on Robotic Computing (IRC) Naples, Italy, 146–152. IEEE, 2019.
Gao, J., Nuytens, D., Lootens, P., Yong, H., & Pieters, J. G. (2010). Recognising weeds in a maize crop using a random forest machine-learning algorithm and near-infrared snapshot mosaic hyperspectral imagery. Biosystems Engineering, 170, 0 39–50. https://doi.org/10.1016/jbiosystemseng.2018.03.006
Gharde, Y., Singh, P. K., Dubey, R. P., & Gupta, P. K. (2018). Assessment of yield and economic losses in agriculture
due to weeds in India. *Crop Protection*, 107, 0 12–18. https://doi.org/10.1016/j.croppro.2018.01.007

Hamza Asad, M., & Bois, A. (2019). Weed detection in canola fields using maximum likelihood classification and deep convolutional neural network. *Information Processing in Agriculture* 7 (4) doi:10.1016/j-ipa.2019.12.002

Inkyu, S., Chen, Z., Popovic, M., Khanna, R., Liebisch, F., Nieto, J., & Siegwart, R. (2017). Weednet: Dense semantic weed classification using multispectral images and mav for smart farming. *IEEE Robotics and Automation Letters*, 3(1), 0 588–595 doi:10.1109/LRA.2017.2774979

Jinlin, Y., Schumann, A. W., Cao, Z., Sharpe, S. M., & Boyd, N. S. (2019). Weed detection in perennial ryegrass with deep learning convolutional neural network. *Frontiers in Plant Science*, 10 1422–1429 . https://doi.org/10.3389/fpls.2019.01422

Kaiming, H., Zhang, X., Ren, S., & Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition Las Vegas, NV, USA, 770–778, 2016.*

Kamath, R., Balachandra, M., & Prabhu, S. (2019). Raspberry pi as visual sensor nodes in precision agriculture: A study. *IEEE Access*, 7, 0 45110-45122. https://doi.org/10.1109/ACCESS.2019.2908864

Kamath, R., Balachandra, M., & Prabhu, S. (2020). Paddy crop and weed discrimination: A multiple classifier system approach. *International Journal of Agronomy*, 2020, 1–16. https://doi.org/10.1155/2020/6474536

Kamath, R., Balachandra, M., & Prabhu, S. (2018). Paddy crop and weed classification using color features for computer vision based precision agriculture. *International Journal of Engineering and Technology (UAE)*, 70(4), 0 2909–2916 doi:10.14419/ijet.v7i4.21511

Kamilaris, A., & Prenafeta-Boldu, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147, 0 70–90. https://doi.org/10.1016/j.compag.2018.02.016

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25, 0 1097–1105 https://dl.acm.org/doi/10.5555/2999134.2999257

Liu, B., & Bruch, R. (2020). Weed detection for selective spraying: A review. *Current Robotics Reports* 1 (1), 1–8 doi:10.1007/s43154-020-00011-w.

Long, J., Shelhamer, E., & Darrell, T. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition Boston, MA, USA, 3431–3440, 2015.*

Mo, X., Deng, X., Long, Q., Jiang, Y., Hongwei, L., Wang, Y., & Xing, X. (2019). Fully convolutional network for rice seedling and weed image segmentation at the seedling stage in paddy fields. *PloS One*, 14(4) 1–13 doi:10.1371/journal.pone.0215676.

Milioto, A., Lottes, P., & Stachniss, C. Real-time semantic segmentation of crop and weed for precision agriculture robots leveraging background knowledge in cnns. In *2018 IEEE International Conference on Robotics and Automation (ICRA)* Brisbane, Australia, 2229–2235. *IEEE, 2018.*

Naidu, V. S. G. R. Hand book on weed identification, 2012.

Nistrup Jørgensen, L., Noe, E., Longvad, A.-M., Jensen, J. E., Erik Ørum, J., & Rydahl, P. (2007). Decision support systems: Barriers and farmers’ need for support. *EPPO Bulletin*, 37(2), 0 374–377. https://doi.org/10.1111/j.1365-2338.2007.01145.x

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 90(1), 0 62–66. https://doi.org/10.1109/TSMC.1979.410076

Parameswari, Y. S., & Srinivas, A. (2017). Weed management in rice—a review. *International Journal of Applied and Pure Science and Agriculture*, 3 (1) (https://iapso.com/published-papers/volume-3issue-1/weed-management-in-rice-a-review.pdf).

Redmon, J., & Farhadi, A. Yolo9000: Better, faster, stronger. In *Proceedings of the IEEE conference on computer vision and pattern recognition Honolulu, USA, 7263–7271, 2017.*

Santiago, W. E., Leite, N. J., Teuel, B. J., Karkee, M., Azania, C. A. M. et al. (2019). Evaluation of bag-of-features (bof) technique for weed management in sugarcane production. *Australian Journal of Crop Science*, 130(11), 0 1819. https://doi.org/10.21475/ajcs.19.13.11.p1819

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61, 1 85–117 doi:10.1016/j.neunet.2014.09.003.

Simonyan, K., & Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.

Subudhi, H. N., Panda, S. P., Behera, P. K., & Patnaik, C. (2015). A check list of weeds in rice fields of coastal Orissa, India. *Journal of Agricultural Science*, 70(6), 0 207 doi:10.5539/jas.v7n6p207.

Sureshkumar, R., Ashoka Reddy, Y., & Ravichandran, S. (2016). Effect of weeds and their management in transplanted rice—a review. *International Journal of Research in Applied, Natural and Social Sciences*, 40 (11), 0 165–180 https://www.researchgate.net/publication/310774176_EFFECT_OF_WEEDS_AND_THEIR_MANAGEMENT_IN_TRANPLANTED_RICE_-_A_REVIEW.

Xiaomeng, L., Chen, H., Xiaojuan, Q., Dou, Q., Chi-Wing, F., & Heng, P.-A. (2018). H-denseunet: Hybrid densely connected unet for liver and tumor segmentation from ct volumes. *IEEE Transactions on Medical Imaging*, 37(012), 0 2663–2674 https://doi.org/10.1109/TMI.2018.2845918.

Zhou, B., Zhao, H., Puig, X., Fidler, S., Barriuso, A., & Torralba, A. Scene parsing through ade20k dataset. In *Proceedings of the IEEE conference on computer vision and pattern recognition Honolulu, HI, USA, 633–641, 2017.*
