
Real-time early detection of weed plants in pulse crop field using drone with IoT

Revanasiddappa, B.^{1*}, Arvind, C. S.^{1,2} and Swamy, S.³

¹Department of Computer Science, Dr. Ambedkar Institute of Technology, Bengaluru, Karnataka, India; ²SBIC A-star Singapore; ³Department of Computer Science, Sir M.Visvesvaraya Institute of Technology, Bengaluru, Karnataka, India.

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

Abstract The real-time detection of parthenium weed plants in the pulse crop field was carried out using low altitude flying drone. A fully convolutional semantic segmentation model was proved to accurately perform object segmentation with higher time complexity. In this research, the LinkNet model with Resnet-34 was used for real-time detection of weed plants using a video feed from low altitude flying drones. Experimental results is proven that LinkNet-34 can detect overlapping and irregular shape weed objects at 0.86 mean pixel accuracy of 0.598 mean IoU at 0.217s. The processing speed was better compared to LinkNet and U-Net models. The detected weed images were stitched together to create a weed site map. The created map is automatically uploaded to google cloud for further site analysis.

Keywords: Deep learning, LinkNet, ResNet, Site map, U-Net, Unmanned aerial vehicle, Weeds

Introduction

Pulses are one of the major food crops feeding over 80% of the Indian population (Annual Report, 2017-18), eradicating malnutrition and providing a good amount of protein and better economic profit for the farmers. However, the yield of pulses, year after year, is going down due to the weeds in farmlands impeding the growth of pulse crops. Weeds in the farmlands have decreased pulse production by competing for light, nutrients, moisture in the pulse-field (Tshewang *et al.*, 2016). In traditional farming, de-weeding is done by spraying pesticide/herbicides without distinguishing pulse crops with weeds. This methodology not only results in a waste of pesticide/herbicides, but it also causes environmental and health hazards for humans (Wiles, 2009). Smart site-specific weed management that reduces pesticide consumption by 50% and environmental pollution which increases crop yield resulting in economic

* **Corresponding Author:** Revanasiddappa, B.; **Email:** sidduaitse@gmail.com, revanasiddappa@dr-ait.org.

profits (Jensen *et al.*, 2012). To obtain these benefits, automatic identification of weeds and their positions is necessary for site-specific spraying.

The automatic spraying of pesticides requires the generation of a weed cover map. However, in past, the researchers have used binary segmentation methods with projection transformation to calculate the weed area (Tellaeche *et al.*, 2011). With the advancement of an unmanned aerial vehicle to protect plants, weed detection is done by capturing data from different height using the drone and projection transformation is performed for weed detection in maize and sunflower fields (Pe íez-Ortiz *et al.*, 2016; Lo íez-Granados *et al.*, 2016; Borra-Serrano *et al.*, 2015). However, it remains difficult to obtain accurate information on the weed area by flying the drone at that height. To overcome this drawback, low-altitude flying nano drones with a camera fitted can be used to get accurate information on weed areas on a small scale and to distinguish weed and crop.

Deep learning, a branch of machine learning is being widely applied in various areas of research and has become a powerful method for image classification (Krizhevsky *et al.*, 2012; Szegedy *et al.*, 2014) and object detection (Erhan *et al.*, 2013). Typically, object detection like fast-rcnn (Girshick *et al.*, 2014), faster-rcnn (Burlina *et al.*, 2016), and YOLO (Redmon *et al.*, 2016) perform well when the object bounding boxes tightly surrounds the region of interest. In this research, weeds or pulse plants do not have definite boundaries and might overlay with each other. Morphological diversity of weed growth will cause challenges for using object detection with a tight bounding box.

The main aim of semantic segmentation was to obtain the class results of each pixel at the corresponding position. The patch-level method which uses features of image patches to train the classifiers (Wei *et al.*, 2015) is time-consuming and the performance of an algorithm is affected by the limited number of patches. To overcome this drawback a pixel-wise fully convolutional network (FCN) was developed which was able to obtain the position of every pixel and features of points are taken to train the classifiers (Shelhamer *et al.*, 2017). FCN takes an RGB or single-channel image of any size as input and retains spatial information such that it can classify each pixel on the feature map. This helps in generating field maps of a weed location.

Deep learning is being widely used in many applications of agriculture. Semantic segmentation algorithm like FCN has been used for weed detection from images captured from UAV for different crops like rice, soybean, sunflower (Huang *et al.*, 2018). The demonstrated techniques use offline image data processing to generate a weed map of the captured data.

LinkNet-34, which was proposed by (Zhou *et al.*, 2018) is a deep fully convolutional neural network for segmentation that uses ResNet-34 as

backbone architecture which is pre-trained on ImageNet (Krizhevsky *et al.*, 2017) as its encoder. ResNet34 is designed to work for mid-resolution images of size 256x256 and does accurate fast semantic segmentation which can detect irregular shape objects like weeds accurately because of its symmetric encoder and decoder structure level features.

In this paper, a robust real-time fast semantic segmentation with automatic weed map generation of detected weeds using nano drone with high-resolution camera can fly at low altitude and transmit real image feed to the ground system to do real-time processing of the captured data is proposed.

The main objectives of this study were to propose a semantic segmentation pipeline which can work in real-time to detect weeds from the drone video feed and create a weed map that can be uploaded to the cloud, and to compare the performance of semantic segmentation models, namely U-Net, LinkNet and LinkNet-34 for real-time weed detection.

Materials and methods

The perception system encompasses the steps is shown in Figure 1. The pipeline for weed detection included four different stages. (i) The capture of weed plant data was used an unmanned aerial vehicle UAV. (ii) Data were prepared by manual annotation. (iii) The weed plant detection model was trained by hyperparameter tuning. (iv) Weed localization (v) and weed map creation were uploaded to the cloud. The detector was run on the prepared images to generate the candidate weed regions. The generated candidate region of Interest (ROIs) from the second step was localized the weed regions which fed into weed map creation, and the generated weed map is uploaded to the cloud for further analysis.

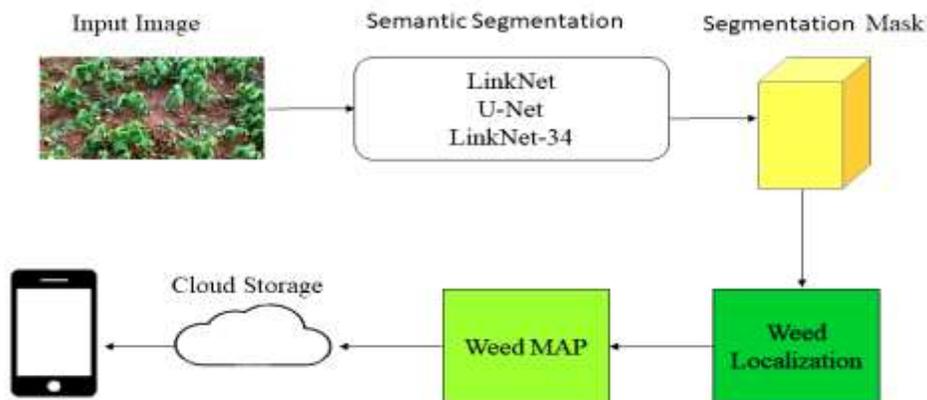


Figure 1. The perception of weed detection system

Data acquisition

Raw *Parthenium hysterophorus* is, a native American tropics weed (Manpreet *et al.*, 2014) which infects farmland and pastures often causing disastrous loss of yield as reflected in common name of famine weed. This weed in contact with human can cause dermatitis and respiratory malfunction (Munesh *et al.*, 2010). Thus, early detection of parthenium in the pulse field is necessary.

The weed images in the pulse field was captured using Tello drone which has 5-megapixel camera of [1280x720] resolution image with low altitude stable flight (Figure 2). Data were captured on September 21, 2019, which was approximately 35 days from the sowing of the horse gram pulse. The horse gram pulse field is located in mysuru, India ($12^{\circ}.23'$ N, $76^{\circ}.64'$ E). The drone had manual flight, so that it can be flown at low altitude with stop and go method to capture the weed. Around 9.82 minutes of flight data were captured which resulted in 294 frames at 30 frames per second.



Figure 2. Weed images in pulse-field

Data preparation

The acquired image is of 1280x720 resolution which is large in size that might overload the GPU memory. Therefore, each image is resized to 512x512 resolution and ground truth (GT) is manually annotated using an open-source labeling tool called labelme (Arvind *et al.*, 2019). In this experiment, two categories are considered as weed and background which included healthy crops and soil as shown in Figure 3. The annotated ninety percent of samples

were randomly selected in which eighty percent is considered for training, 10% for validation, and the remaining 10% for testing dataset. Validation of the training model is directly tested from the drone itself.

The horse gram pulse crop field data were captured with parthenium weeds from low altitude flight drones. Figure 4 shows The input image of horse gram pulse-field captured from low altitude drone, the white region showing manually annotated weed plant and weed plant region outline is shown in yellow color showing the plant is of irregular shape and size with overlapping features that are labeled as a, b and c respectively (Figure 4).

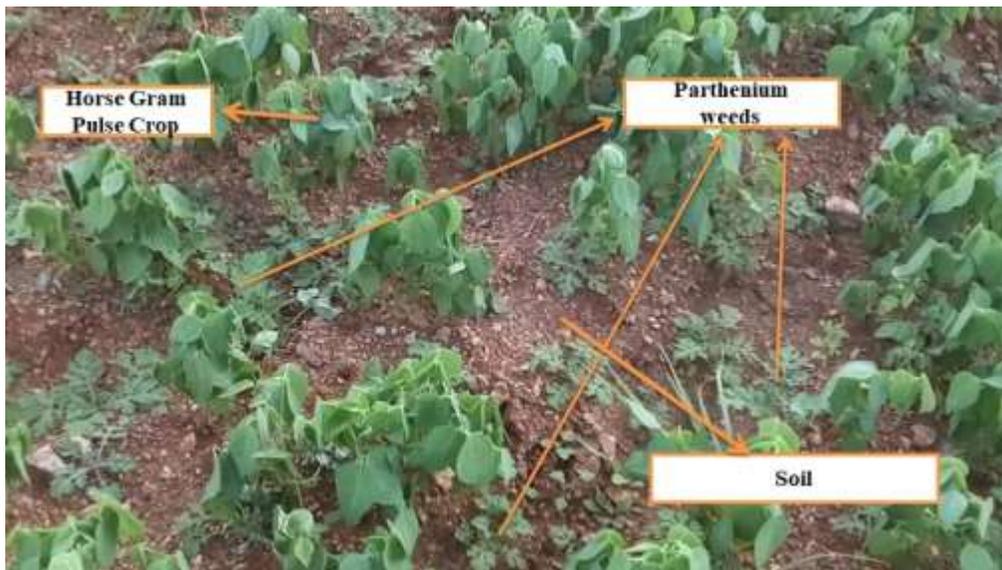


Figure 3. Horse gram Pulse crop field data

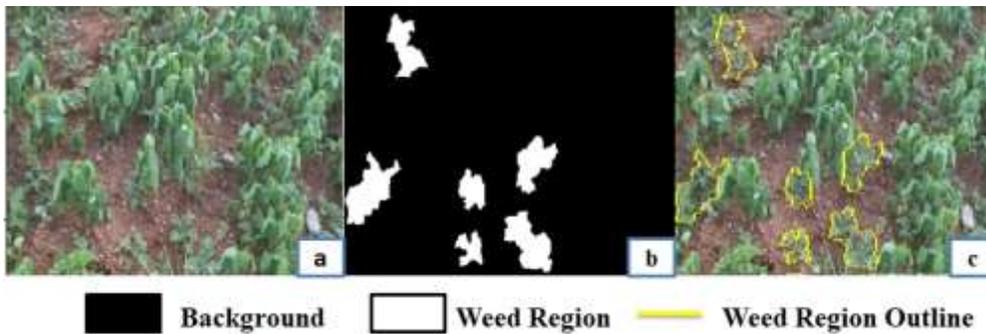


Figure 4. Input image of Horse gram pulse-field

Drone communication

Nano Tello drone is connected to a local system running the ubuntu operating system using wireless communication. The communication to a drone is via Tello API. The binding of drone uses local IP address with local video port. The video streaming uses a h264 video encoder (Richardson *et al.*, 2003) with a sender and receiver video socket. The video streaming is done at 30 frames per second using multithreading concepts (Cetinkaya *et al.*, 2019). Flight control during data acquisition is via manual flight but testing is fully autonomous using Tello API commands. Command communication uses the multithreading principle. The coverage range of wireless communication is up to 100 meters in radius and 50 feet in altitude using a WiFi extender. Figure 5 shows the secured video streaming and command communication between drones to a local system using wireless communication.



Figure 5. Secured video streaming and command communication

Weed detection training architecture

In the weed detection problem, the original size of the image captured and the annotated masks are 1280x720. For algorithmic standard the image captured and the mask is resized to 512x512. The image contains two classes of an object mainly (i) pulse crop (ii) weed. Captured weed and crops are overlaying with the same color but with different leaf shapes. Considering these features, semantic segmentation algorithm like U-net, LinkNet, and, LinkNet-34 is designed to process 512x512 image resolution as input, so that it can preserve image features.

The U-Net semantic segmentation model was developed by Ronneberger (Ronneberger *et al.*, 2015) that performed well on biomedical image data. U-Net model contains encoder and decoder structure, where encoder would shrink the data, and decoder expanding the features which are encoded. The lost information during encoding is copied during decoding, so that lost edge information is replenished. This model can accurately predict the edge pixel values, which is very important in correctly segmenting the overlay weed plant region. The U-Net model which is used in this research work is shown in Table 1. It showed the parameters used for the training U-Net model.

Table 1. Typical U-Net architecture with 5 encoder and 5 decoder layers for weed detection

Encoder	Layer type	Feature Size	Number Of Features	Decoder	Layer Type	Feature Size	Number Of Features
E1	Conv1_1	3x3	16	D5	Upsample	2x2	128
					Concat	[D5,E4]	
	Conv1_2	3x3	16		Conv6_1	3x3	128
	Max Pool	2x2	8		Conv6_2	3x3	128
E2	Conv2_1	3x3	32	D4	UpSample	2x2	64
					Concat	[D4,E4]	
	Conv2_2	3x3	32		Conv7_1	3x3	64
	Max Pool	2x2	16		Conv7_2	3x3	64
E3	Conv3_1	3x3	64	D3	UpSample	3x3	32
					Concat	[D3,E2]	
	Conv3_2	3x3	64		Conv8_1	3x3	32
	Max Pool	2x2	32		Conv8_2	3x3	32
E4	Conv4_1	3x3	128	D2	UpSample	3x3	16
					Concat	[D2,E1]	
	Conv4_2	3x3	128		Conv9_1	3x3	16
	Max Pool	64	32		Conv9_2	3x3	16
E5	Conv5_1	3x3	256	D1	Conv10_1	1x1	Sigmoid

Table 2. LinkNet deep learning model

Encoder	Block	Feature	Decoder	Block	Feature
enc1	Encoder Block	(64,64)	dec4	Decoder Block	(256,512)
				Shortcut	[enc3,dec4]
enc2	Encoder Block	(64,128)	dec3	Decoder Block	(128,256)
				Shortcut	[enc2,dec3]
enc3	Encoder Block	(128,256)	dec2	Decoder Block	(64,128)
				Shortcut	[enc1,dec2]
enc4	Encoder Block	(256,512)	dec1	Decoder Block	(64,64)
				Shortcut	[dec1,input]
Encoder Block		Decoder Block		Shortcut	
conv[(3x3),(n,n)] conv[(3x3),(n,n)] conv[(3x3),(n,n)] conv[(3x3),(n,n)/2]		Conv[(1x1),(m/4,n)] Full conv[(3x3),(m/4,m/4),^2] Conv[(1x1),(m,n/4)]		conv(1,1) L2 regularizer	

LinkNet is light weighted semantic segmentation architecture which is lighting fast as stated by author Abhishek (Abhishek and Eugenio, 2017). With 11.5 million parameters, the network has a series of encoder and decoder blocks which downsample up to image and up-sample it with few convolutional layers. The structure of the network used in this research work is shown in Table 2 with a few parameters. Real-time segmentation is achieved to require for real-time weed plant detection through live video fed from the drone camera.

LinkNet model parameter tuning does not improve model performance. The accuracy without compromising on time complexity was improved. LinkNet is integrated with ResNet-34 (He *et al.*, 2015) a pre-trained model such that encoded features of ResNet help to improve the decoder image features. The LinkNet-34 network structure is shown in Table 3 that is improved the segmentation accuracy by removing false positive like soil and also perform real-time segmentation of weed plant from a live video feed from the drone camera. Different training hyperparameters for training U-Net, LinkNet, and LinkNet-34 are mentioned in Table 4.

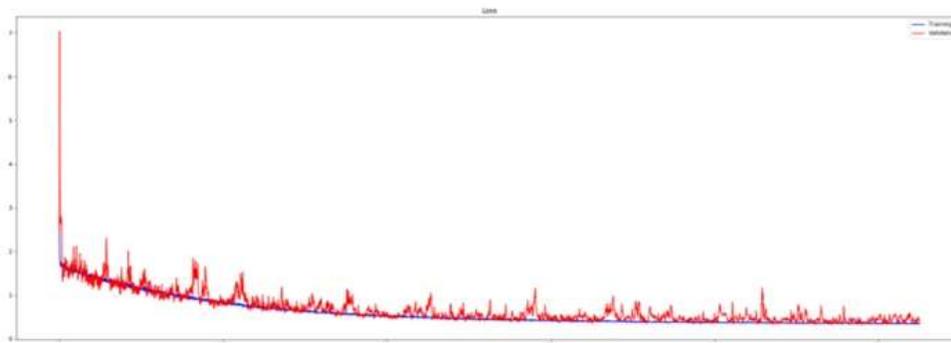
The training loss as represented in the red line getting reduced after each epoch is shown in Figure 6. The red line of the LinkNet-34 loss value is reduced after each training epoch indicating that the algorithm can detect the region of interest.

Table 3. Proposed LinkNet 34 deep learning model architecture for weed plant detection

Encoder	Layer Type	Decoder	Decoder	Layer Type
Enc 1	(7x7,64)	Dec 1	Decoder Block	256
Enc 2	(3x3,64)	Dec 2	Decoder Block	128
Enc 3	(3x3,128)	Dec 3	Decoder Block	64
Enc 4	(3x3,256)	Dec 4	Decoder Block	64
Enc 5	(3x3,512) Resnet-34	Dec 4	Transpose Conv	3x3 ²
Decoder Block		Dec 5	Conv 5_1 Conv 5_2	3x3 3x3
<ul style="list-style-type: none"> • conv 1x1 • Batch Norm • Transpose conv • Batch Norm • Conv 1x1 • Batch Norm 				

Table 4. U-net, LinkNet and LinkNet-34 training parameters for weed detection

Training Parameters	U-Net	LinkNet	LinkNet-34
Training Images	297	297	297
Number of Weed ROI	1467	1467	1467
PreTrained Model	No	No	ResNet-34
Image Dimensions	[512x512]	[512x512]	[512x512]
Learning Momentum	0.9	0.9	0.9
Learning Rate	1e-5	1e-6	1e-6
Number of Classes	2	2	2
Number of Epochs	40	350	450
Steps per Epoch	100	100	100

**Figure 6.** Training loss

Dice co-efficient is shown in equation 1 to determine the accuracy of foreground pixels.

$$\text{Dice Coefficient} = \frac{2 \cdot \text{TP}}{\text{TP} + \text{FP} + \text{TP} + \text{FN}} \quad (1)$$

Where TP is the number of True predicted weed pixels, FP is the number of False predicted weed pixels, FN is the number of false predicted plant regions as weed pixels.

The increase in dice co-efficient with an increase in training epoch showing that the network can be predicted foreground weed pixels by penalizing wrong labels better (Figure 7).

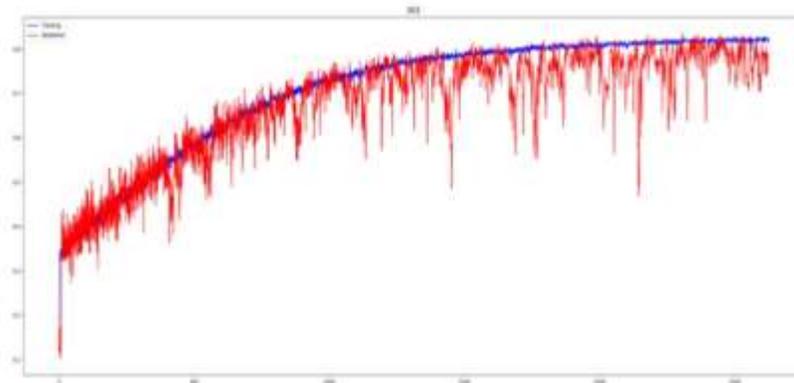


Figure 7. Dice coefficient score

Inference weed detection

Real-time weed detection is achieved with an autonomous flight. The drone which transmits [1280x720] resolution image is taken as input data for the trained model as it can handle variable size input data. During inference, the following steps are followed as per algorithm 1 for real-time weed plant segmentation.

Secure Communication(): Drone and local system are paired together to a wireless hotspot for secured wireless communication.

Command(): Autonomous flight of Tello Drone is achieved using Tello drone API commands which helps to control the drone from a local system. The command is sent according to fly a mission.

Image Data Transfer(): Tello drone uses UDP protocol for sending of captured images to the local system via wireless.

Weed Segmentation(): Trained semantic segmentation model like U-Net, LinkNet, and, LinkNet-34 is used for segmentation.

Algorithm 1: Real-Time Weed Plant Segmentation: Input is RGB Image $I_{(x, y)}$ from Tello Drone camera and output is Weed Plant segmentation region. if Secure Communication(), while True, if Image Data Transfer(), Command = Take off, Command = Up 6 feet, Command = Move Forward/Backward, Weed Segmentation() = $I_{(x, y)}$, Weed Plant Segmentation ROI = $I_{(x, y)}$ Weed Segmentation, end if, end while and end if.

Field map generation

For further analysis, after each mission getting a complete picture of site-specific weed region inside of the pulse grain field is necessary. To generate a

complete picture after each flight weed detected region, images are stitched together to generate a panoramic view using image stitching methodology. Panoramic weed map is generated using scale-invariant features of the detected image using SIFT (Lindeberg, 2012) algorithm and RANSAC (Schnabel *et al.*, 2007) is used to remove outlier. Homography is computed to match feature points and warp perspective image generating which provides site-specific weed region (Agarwal *et al.*, 2005). The flight generated weed field maps is used SIFT and homographic transformation algorithm, where the multiple weed detected images are stitched to generate the site-specific weed map (Figure 8).

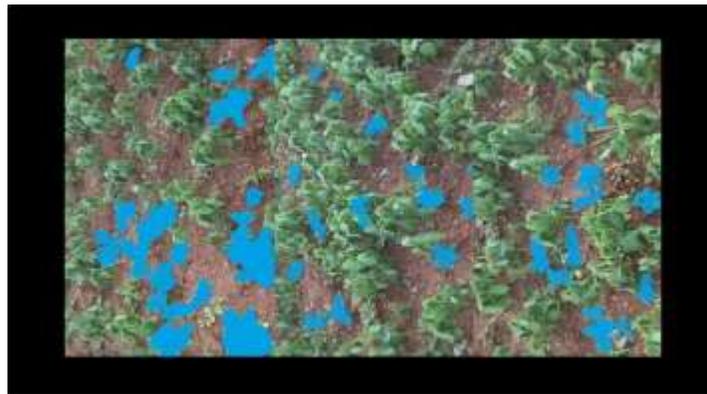


Figure 8. Generated Weed Plant field map

Field map upload to cloud

The Google Drive API allows the uploading of a weed map file when it is created or updated (Dinatha *et al.*, 2016). The upload of weed map uses HTTPS file upload steps as shown in algorithm 2.

Algorithm 2. Weed Map Cloud Upload: Input is RGB $M_{(x, y)}$ generated weed map and output is uploaded message of Generated $M_{(x, y)}$ to cloud. A *POST* request to the upload URL is created. Query parameter *uploadType* is multi-part, $M_{(x,y)}$ Weed Map is added to the *request body*. HTTP Head with Content-Type set to MIME is added. Then, the request for Map Upload to specific URL is send and successfully uploaded message.

Results

Weed region segmentation

Mean Intersection over Union (MIoU) is used to measure the accuracy which determined the union of ground truth region with predicted weed region.

In addition to MIoU measure, the speed of the deep learning algorithm is also measured as it is key for real-time weed detection using drone feed. MIoU is calculated as per equation 2. The weed detection system is developed in python using open-source deep learning Tensorflow 1.12 (Abadi *et al.*, 2016) and OpenCV 4.1 (Gollapudi, 2019) library under Ubuntu 18.04 operating system with 32 GB RAM and NVIDIA 1060 6GB GPU.

$$MIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{P_{ii}}{\sum_{j=0}^k P_{ij} + \sum_{j=0}^k P_{ji} - P_{ii}} \quad (2)$$

Result showed K which represents weed region and crop with soil as background and k in equation 2 was set to 2. In this experiment, i represents ground truth class; j represents the predicted class; P_{ii} represents the number of true positive, which represent a number of pixels of ground truth class and predicted class as same; P_{ij} represents the number of false positives, the misclassified number of pixels and P_{ji} represents the number of false negatives, namely the number of pixels that were falsely classified. Weed detection accuracy at different MIoU threshold levels for three different algorithms with time complexity is shown in Table 45. The pixel classification in terms of the F1 score is evaluated. The Proposed LinkNet-34 is outperformed when compared to the existing U-Net and LinkNet model. The F1 score accuracy rate for weed and background pixel-wise classification for three models was 0.943, 0.897, and 0.836, respectively.

Table 5. Evaluation of pixel classification using f1Score

Approach	Type of Pixel	F1 Score	
		Weed (Parthenium)	BackGround
U-Net	Weed (Parthenium)	0.8367	0.1633
	Pulse Grain with Soil as Background	0.1084	0.8916
LinkNet	Weed (Parthenium)	0.8973	0.1027
	Pulse Grain with Soil as Background	0.0368	0.9632
LinkNet-34	Weed (Parthenium)	0.943	0.057
	Pulse Grain with Soil as Background	0.0257	0.9743

The experimental results of weed detection in the pulse grain field using three different semantic segmentation models is shown in Table 6. The performance metrics such as mean pixel (MP) and mean intersection of union

(MIoU) is used for the estimation of correct pixel prediction. The proposed LinkNet-34 method is outperformed the regular U-Net and LinkNet model with an average accuracy value of 86.7% with a runtime accuracy of 0.217 seconds. The LinkNet-34 is shown a more shortcut connection with Resnet as a backbone model to predict the weed plant pixels accurately compared to U-Net and LinkNet model.

Table 6. Experimental results of weed detection in pulse grain field using three different semantic segmentation models

Approach	Training Images	Number of Weed Samples	Validation Images	Mean Pixel Accuracy of Validation Data	Mean IoU of Validation Data	Testing Data	Mean Pixel Accuracy of Testing Data	Mean IoU of Testing Data	Speed
U-Net	297	1467	30	0.712	0.549	967	0.704	0.527	0.312s
LinkNet				0.824	0.573		0.812	0.553	0.176s
LinkNet-34				0.867	0.598		0.843	0.581	0.217s

The parthenium weed segmentation result of three different models where Input image, ground truth, U-Net segmentation resulted in false-positive of detecting soil as weed region (Figure 9). LinkNet resulted better than U-Net and LinkNet-34 with accurately segmented weed region better than LinkNet result is indicated as a, b, c, d, and e regions.

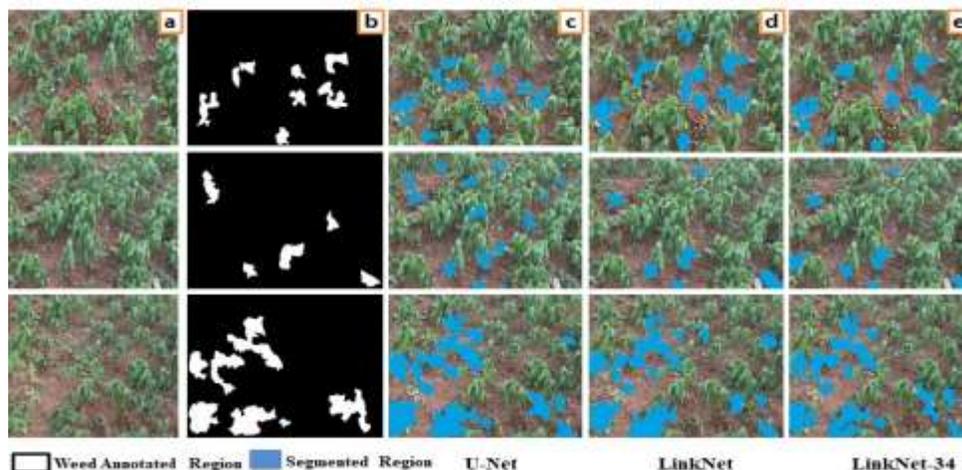


Figure 9. Parthenium weed segmentation

Discussion

The results of the present study revealed that early detection of parthenium weed plants in the pulse crop field can be detected using low

altitude flying drone in real-time. According to previous research work on weed detection during the seedling stage in paddy fields using a fully convolutional network was stated by Ma, *et al.* (2019) using high-resolution image data. In our study, a real-time weed map is created for site is based weed management using deep learning-based semantic segmentation algorithm output. A comparative study was conducted between U-Net, LinkNet, and, LinkNet-34 deep learning models for real-time weed segmentation. In which experimental results showed that LinkNet-34 model can segment parthenium weed plants, which were overlapping, and in-between pulse crop with high accuracy in real-time as shown Table 6. Huang *et al.* (2018) created weed mapping using unmanned aerial vehicle imagery using FCN deep learning architecture with an accuracy of 93.4%, not in real-time as FCN was very slow compared to LinkNet. The proposed automatic weed map generation was the time efficient and is conducted after each flight by stitching weed plant detected images using feature-based homography method as shown in Figure 8. The generated weed map is uploaded to a cloud drive for end-user field analysis as explained in algorithm 2. Even though, the proposed weed, plant segmentation pipeline was working in real-time using deep learning from low altitude drone video data. It failed to cover the complete field as the flight was only 13 minutes for the experimental drone and feature-based image stitching fails to create a good weed map, if there were 10 frames difference from continuous frame sequences. To overcome flight time issues of Tello drone either multiple flights needed to be flown or used an advanced version of low altitude drone-like Marvic Mini which is better flight time and stable video capture. Feature-based stitching can be replaced with learning-based image registration method which can generate better weed map for analysis.

Acknowledgements

This work was fully conducted and supported by Dr. Ambedkar Institute of Technology, Bengaluru. Author 2 is presently working in affiliation 2 during the submission of this manuscript and no research activity was conducted at affiliation 2.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D.G., Steiner, B., Tucker, P.A., Vasudevan, V., Warden, P., Wicke, M., Yu, Y. and Zhang, X. (2016). TensorFlow: A system for large-scale machine learning. Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16), 265-283.

- Abhishek, C. and Eugenio, C. (2017). LinkNet: Exploiting Encoder Representations for Efficient Semantic Segmentation. IEEE Visual Communications and Image Processing (VCIP) DOI:10.1109/VCIP.2017.8305148.
- Annual Report (2017-18). Directorate of Pulses Development, Ministry of Agriculture & Farmers Welfare Government of India. Retired from <http://dpd.gov.in/Annual%20Report%202017-18.pdf>.
- Agarwal, A., Jawahar, C. V. and Narayanan, P. J. (2005). A Survey of Planar Homography Estimation Techniques. Retired from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.102.321&rep=rep1&type=pdf>.
- Borra-Serrano, I., Peña J. M., Torres-Sánchez, J., Mesas-Carrascosa, F. J. and López-Granados, F. (2015). Spatial Quality Evaluation of Resampled Unmanned Aerial Vehicle-Imagery for Weed Mapping. *Sensors (Basel)*, 15:19688-19708.
- Burlina, P, editor MRCNN: A stateful Fast R-CNN (2016). 23rd International Conference on Pattern Recognition (ICPR), 12:4-8.
- Cetinkaya, O., Sandal, S., Bostancı, E., Güzel, M.S., Osmanoglu, M. and Kanwal, N. (2019). A Fuzzy Rule-Based Visual Human Tracking System for Drones. 2019 4th International Conference on Computer Science and Engineering (UBMK) 1-6.
- Arvind, C. S., Prajwal, R., Bhat, P. N., Sreedevi, A. and Prabhudeva, K. N. (2019). Fish Detection and Tracking in Pisciculture Environment using Deep Instance Segmentation. TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON), Kochi, India, pp. 778-783. DOI: 10.1109/TENCON.2019.8929613.
- Dinatha, R. C., Sukarsa, I. M. and Cahyawan, A. (2016). Data Exchange Service using Google Drive API. *International Journal of Computer Applications*, 154:12-16.
- Erhan, D., Szegedy, C., Toshev, A. and Anguelov, D. (2013). Scalable Object Detection Using Deep Neural Networks. IEEE Conference on Computer Vision and Pattern Recognition arXiv:1312.2249v1.
- Girshick, R., Donahue, J., Darrell, T. and Malik, J. (2014). Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. IEEE Conference on Computer Vision and Pattern Recognition arXiv:1311.2524v5.
- Gollapudi, S. (2019). Learn Computer Vision Using OpenCV with Deep Learning CNNs and RNNs. 151 p.
- He, K., Zhang, X., Ren, S. and Sun, J. (2015). Deep Residual Learning for Image Recognition. (2016). IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770-778.
- Huang, H., Deng, J., Lan, Y., Yang, A., Deng, X. and Wen, S. (2018). Accurate Weed Mapping and Prescription Map Generation Based on Fully Convolutional Networks Using UAV Imagery. *Sensors* 18.
- Jensen, H. G., Jacobsen, L. B., Pedersen, S. M. and Tavella, E. (2012). Socioeconomic impact of widespread adoption of precision farming and controlled traffic systems in Denmark. *Precision Agriculture*, 13:661-77. <https://doi.org/10.1007/s11119-012-9276-3>.
- Krizhevsky, A., Sutskever, I. E. and Hinton, G. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *International Conference on Neural Information Processing Systems*, 25:1097-105.
- Krizhevsky, A., Sutskever, I. and Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Commun. ACM*, 60:84-90.
- Lindeberg, T. (2012). Scale Invariant Feature Transform. *Scholarpedia*, 7:10491.
- Lopez-Granados, F., Torres-Sanchez, J., De Castro, A. I., Serrano-Perez, A., Mesas-Carrascosa, F. J. and Pena, J. M. (2016). Object-based early monitoring of a grass weed in a grass

- crop using high-resolution UAV imagery. *Agronomy for Sustainable Development*, 36:67. [10.1007/s13593-016-0405-7](https://doi.org/10.1007/s13593-016-0405-7).
- Manpreet, K., Aggarwal, N. K., Kumar, V. and Dhiman, R. (2014). Effects and Management of *Parthenium hysterophorus*: A Weed of Global Significance. *International Scholarly Research Notices*, vol. 2014, Article ID 368647, 12 p. <https://doi.org/10.1155/2014/368647>.
- Ma, X., Deng, X., Qi, L., Jiang, Y., Li, H., Wang, Y. and Xing, X. (2019). Fully convolutional network for rice seedling and weed image segmentation at the seedling stage in paddy fields. *PLoS ONE*, 14.
- Munesh, K., Kumar, S. and Sheikh, M. A. (2010). Effect of *Parthenium hysterophorus* ash on growth and biomass of *Phaseolus mungo*, 9:145-148.
- Perez-Ortiz, M., Pena, J. M., Gutierrez, P. A., Torres-Sánchez, J., Hervas-Martinez, C. and Lopez-Granado, F. (2016). Selecting patterns and features for between- and within-crop-row weed mapping using UAV-imagery. *Expert Systems with Applications*, 47:85-94. <https://doi.org/10.1016/j.eswa.2015.10.043>.
- Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016). You only look once: Unified, Real-Time Object Detection. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) arXiv:1506.02640v5.5.
- Richardson, I. E. (2003). H.264 and MPEG-4 Video Compression: Video Coding for Next-Generation Multimedia, pp.137.
- Ronneberger, O., Fischer, P. and Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention* arXiv:1505.04597v1.
- Schnabel, R., Wahl, R. and Klein, R. (2007). Efficient RANSAC for point-cloud shape detection. *Comput. Graph. Forum*, 26:214-226.
- Shelhamer, E., Long, J. and Darrell, T. (2017). Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39:640-51. <https://doi.org/10.1109/TPAMI.2016.2572683> PMID: 27244717.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. and Rabinovich, A. (2014). Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) arXiv:1409.4842v1.9.
- Tellaèche, A., Pajares, G., Burgos-Artizzu, X. P. and Ribeiro, A. (2011). A computer vision approach for weeds identification through Support Vector Machines. *Applied Soft Computing*, 11:908-15. <https://doi.org/10.1016/j.asoc.2010.01.011>.
- Tshewang, S., Sindel, B. M., Ghimiray, M. and Chauhan, B. S. (2016). Weed management challenges in rice (*Oryza sativa* L.) for food security in Bhutan: A review. *Crop Protection*, 90:117-24.
- Wei, S., Xinggang, W., Yan, W., Xiang, B. and Zhang, Z. (2015). DeepContour: A deep convolutional feature learned by positive-sharing loss for contour detection. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 7–12 June 2015.
- Wiles, L. J. (2009). Beyond patch spraying: site-specific weed management with several herbicides. *Precision Agriculture*, 10:277-90.
- Zhou, L., Zhang, C. and Wu, M. (2018). D-LinkNet: LinkNet with Pretrained Encoder and Dilated Convolution for High-Resolution Satellite Imagery Road Extraction. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) DOI: 10.1109/CVPRW.2018.00034.

(Received: 7 January 2020, accepted: 30 August 2020)