IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 6.252

Volume 2, Issue 1, November 2022

Weeds Detection Using Semi-Supervised Learning For Precision Agriculture

B. Leela Prasad[1], P. Krishna[2], M. Harsha Siril[3], Sk. E. Abhinaya[4], P. Kalpana[5], N. Haritha[6]

GMR Institute of Technology, Rajam, Andhra Pradesh, India

21345A0509@gmrit.edu.in[1], 21345A0510@gmrit.edu.in[2]

Abstract: *Farmers around the world are facing challenges of growing more food for the increasing world population. Especially facing problem in identifying weeds in the crops. Many ways to kill weeds by using traditional techniques but these techniques have disadvantages such as time consuming, manpower, spraying herbicides that can harm the real crops. By using Deep learning-based semi-supervised approach can detect weeds. An Autonomous robots using a Convolutional Neural Network (CNN) based unsupervised segmentation was used to capture the images of weeds and taken as dataset. A fine-tuned CNN(ResNet50) identifies weed infected region. With image-processing technique, the super pixels are extracted using the Simple Linear Iterative Clustering (SLIC) algorithm. A pre- trained ResNet50 used as a feature extractor that acts on classifiers such as SVM, Gaussian Naive Bayes, Neural Network, and Random Forest. Also trained U-Net model on the training split of the dataset. Major datasets like Crop Weed Field Image Dataset (CWFID) and Sugar Beets dataset are used in this approach. A Pre-trained ResNet50 eliminates the need for designing hand-crafted features for weeds detection. Once the weed-infested regions have been identified, the weed density can be computed with maximum accuracy. So, Weeds can be identified and eliminated from the main crop..*

Keywords: Weed detection, Image Processing, Deep learning, Semi-Supervised learning, Convolutional Neural Network (CNN), U-Net

I. INTRODUCTION

Agricultural sector plays an important role in contributing to the Indian economy. Many countries are working on agricultural development that results in sustaining the agricultural production but weed is the only crop threat that shrinks the production rates. Traditional techniques can to use for weeds detection but these techniques have disadvantages such as time consuming, manpower, spraying herbicides that can harm the real crops. On other hand Artificial Intelligence and machine learning are present trending and growing technologies that can also use for smart agriculture and increases productivity. As our country population increases day by day there is a need to increase food productivity and this can be achieved by using these technologies only.

An Autonomous robots using a Convolutional Neural Network (CNN) based unsupervised segmentation was used to capture the images of weeds. These robots rely on systems, including machine vision, to identify and localize weed plants. A typical image processing-based weed detection approach consists of four stages: pre-processing, segmentation, feature extraction, and classification. And by using this technique the super pixels are extracted using the Simple Linear Iterative Clustering (SLIC) algorithm. Input image (IRGB) is first resized to 500 × 500 sq. pixels using the bicubic interpolation method implemented by OpenCV library. Each pixel in the image has to be clustered into one of the two classes - background or vegetation. For this purpose, we use the CNN based approach proposed for unsupervised segmentation. And also trained U-Net model on the training split of the datasets. U-Net has been shown to be an effective supervised learning approach for pixel-wise segmentation in different use cases. A variety of machine learning algorithms such as Support Vector Machines, Random Forest Classifiers, Gaussian naive Bayes, and multilayer perceptron networks have commonly been used for classification and performances can be compared.

Instead of extracting features based on the biological morphology, physical appearance of the crop/weed, a pre-trained CNN is utilized. The use of pre-trained CNN as a feature extractor eliminates the need for designing hand crafted features. ResNet50 is utilized to extract features. It is also important to note that network is trained using the masked tile images instead of complete RGB images this allows the network to make predictions accurately. Once the weed-infested regions have been identified, the weed density can be computed from the vegetation coverage in each individual region. And Mean intersection-over-union (MIOU) is a popular metric to evaluate pixel-wise segmentation network. The accuracy of weed distribution and density estimation is evaluated. This approach can robustly identify weed infected regions, Compute weed density.

II. LITERATURE SURVEY

[1] Shorewala, S., Ashfaque, A., Sidharth, R., & Verma, U. (2021). Weed density and distribution estimation for precision agriculture using semi-supervised learning. *IEEE access*, *9*, 27971-27986.

In this paper weeds can be identified using deep learning-based semi-supervised techniques. Foreground vegetation pixels containing crops and weeds are first identified using a CNN. Subsequently, the weed infected regions are identified using a fine-tuned CNN and algorithms like SVM, Random Forest Classifiers, Gaussian naive Bayes used for classification.One of the limitations of this work is the iterative nature of generating vegetation masks. By using CNN based techniques weeds can be accurately identified and using SLIC, SVM, Random Forest Classifiers, Gaussian naive Bayes algorithms for classification weeds can be classified perfectly from main crop.Weed infested regions are identified with a maximum recall of 0.99. Weed density in these regions is estimated with an accuracy of 82.13%.

[2] Jin, X., Che, J., & Chen, Y. (2021). Weed identification using deep learning and image processing in vegetable plantation. *IEEE Access*, *9*, 10940-10950.

To identify weeds in vegetable plantation using deep learning and image processing. A Center Net model was trained to detect vegetables. To extract weeds from the background, a color index was determined and evaluated through Genetic algorithms (GAs) according to Bayesian classification error. For evaluation Genetic algorithms (GAs) according to Bayesian classification error can be used to exactly separate weeds and vegetable plants. The trained CenterNet achieved a precision of 95.6%, a recall of 95.0% and a F1 score of 0.953.

[3] Sodjinou, S. G., Mohammadi, V., Mahama, A. T. S., & Gouton, P. (2021). A deep semantic segmentation-based algorithm to segment crops and weeds in agronomic color images. *Information Processing in Agriculture*.

This work proposed a segmentation method. Agronomic images of two different databases were used for the segmentation algorithms. Using the thresholding technique, everything except plants was removed. Afterward, semantic segmentation was applied using U-net followed by the segmentation of crops and weeds using the K-means subtractive algorithm. Based on the confusion matrix, the true-positive and true-negative values were 0.995 2 and 0.898 representing the rate of crops and weeds, respectively.An advantage of using K-means is to be easy to implement.The best precision of segmentation was achieved using GAC (geometric active contour) method amounted to 94.56%.The proposed algorithm performed very well (i.e., accuracy of 99.19%) in segmentation having an efficient segmentation of crops and weeds.

[4] Selvi, C. T., Subramanian, R. S., & Ramachandran, R. (2021, March). Weed Detection in Agricultural fields using Deep Learning Process. In *2021 7th International Conference (ICACCS)* (Vol. 1, pp. 1470-1473). IEEE.

A fast-growing area of research today is artificial intelligence, specifically deep learning. Object recognition, making use of computer vision, is one of its numerous applications. This work suggests a deep learning with image processing-based framework to classify, various crops and weeds. This experiment is tested over 10,000 images and takes 48 hours to yield 94.74% accuracy. CNET has gained 92.08% accuracy.

[5]. Jogi, Y., Rao, P. N., & Shetty, S. (2020, November). CNN based Synchronal recognition of Weeds in Farm Crops. In *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)* (pp. 1373-1378). IEEE.

Weed is an unwanted plant that is found in the field. They can do some harm to the main crop. Traditional methods are supported but they have lot of disadvantages. A new system has been proposed to perform the real time identification of weeds using a deep learning method. The image is being captured by webcam and is processed by raspberry pi. using this method with maximum accuracy.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Impact Factor: 6.252
Volume 2, Issue 1, November 2022

III. METHODOLOGY

3.1 Convolution Neural Network

A convolutional neural network (CNN) is a type of artificial neural network used in image processing and processing that is specifically designed to process pixel data. CNNs uses deep learning to perform both generative and descriptive tasks, often using machine vison that includes image and video recognition, along with recommender systems and natural language processing. A CNN has convolutional layer, pooling layer, and a fully connected layer. The convolution layer is the core building block of the CNN. It carries the main portion of the network's computational load. The pooling layer replaces the output of the network at certain locations by deriving a summary statistic of the nearby outputs. The fully connected layer helps to map the representation between the input and the output

3.2 Simple Linear Iterative Clustering (SLIC)

SLIC algorithm generates super pixels by clustering pixels based on their colour similarity and proximity in the image plane. This is done in the five-dimensional space. We need to normalize the spatial distances in order to use the Euclidean distance in this 5D space because the maximum possible distance between two colours in the CIELAB space is limited whereas the spatial distance in the xy plane depends on the image size. Therefore, In order to cluster pixels in this 5D space, a new distance measure that considers super pixels size was introduced.

3.3 U-Net Model

UNet is a convolutional neural network architecture that expanded with few changes in the CNN architecture. It was invented to deal with biomedical images where the target is not only to classify whether there is an infection or not but also to identify the area of infection. It is an architecture for semantic segmentation. It consists of a contracting path and an expansive path. The contracting path follows the typical architecture of a convolutional network. It consists of the repeated application of two 3x3 convolutions, each followed by a rectified linear unit and a 2x2 max pooling operation with stride 2 for down sampling. At each down sampling step we double the number of feature channels.

3.4 Vegetation Segmentation :

First the input image is resized to 500 × 500 sq. Each pixel has to be clustered into two classes background or vegetation using CNN based approach. This iterative approach is solved in two steps label prediction and learning network parameters. And maintains constraints for predicting the cluster or class to which each pixel might belongs to. The main constraints are feature similarity, spatial continuity and final constraint is placed on the number of unique clusters into which the image is segmented. Super pixels can be defined as a group or cluster of pixels that exhibit common characteristics such as pixel intensity and proximity. The network extracts the super pixels using the Simple Linear Iterative Clustering (SLIC) algorithm which operates in a five-dimensional space.

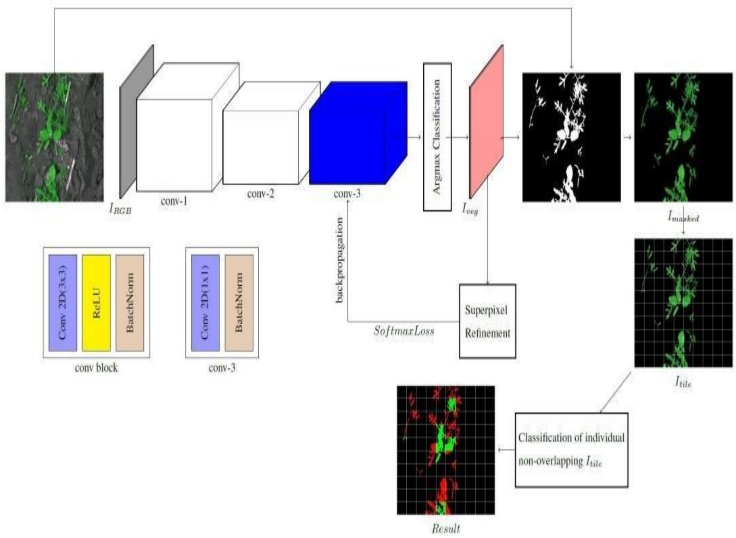

Figure : Overview of the proposed approach

A. Tile Classification

Once the vegetation mask is generated, the input image IRGB is overlaid with resulting in the masked image Imasked. This masked image contains only the RGB pixels for the vegetation (crops and weeds), which ensures that classification is performed based on vegetation features alone. Imasked is divided into smaller tiles. Also, the outline color represents the classification for the region: Blue

- Vegetation, Orange - Background/Soil. A variety of machine learning algorithms such as Support Vector Machines, Random Forest Classifier, Gaussian naive Bayes, and multilayer perceptron networks have commonly been used for classification.
- Feature Vector Based Classification.
- Image-based Classification.
- ResNet50 is utilized.

B. Weed Density Estimation

Once the weed-infested regions have been identified, the weed density can be computed from the vegetation coverage in each individual region. The cluster rate was denoted by CR, is used to quantify, or model the weed density.

Dataset:

In this approach Crop/Weed Field Image dataset [65] and the Sugar Beets dataset are used.

CROP/WEED FIELD IMAGE DATASET (CWFID): It contains images acquired by an autonomous field robot BoniRob from a carrot farm, which are split into train and test set in the ratio of 2:1 SUGAR BEETS DATASET: It contains field images acquired by the same autonomous robot BoniRob from a sugar beet farm for over three months. While the entire dataset is quite extensive and includes data from multiple sensors, only a subset of the Sugar Beets dataset. It is further split into the train and test set with a ratio of 7:3.

IV. RESULTS AND DISCUSSION

4.1 Performance Metrics

The main performance metrics are mIoU, Precision, Recall, F1-Score, Accuracy and for weed density estimation performance metrics used are Cluster Rate, Absolute Error, Mean Accuracy, MAE, RMSE.

mIoU: Mean intersection-over-union (mIoU) is a popular metric to evaluate pixel-wise segmentation networks.

$$mIoU = \frac{\sum_i x_{ii}}{C\left(\sum_i \sum_j x_{ij} + \sum_j x_{ji} - x_{ii}\right)}$$

Precision: Precision refers to the number of true positives divided by the total number of positive predictions.

Recall: The recall is calculated as the ratio between the numbers of Positive samples correctly classified as Positive to the total number of Positive samples.

F1-Score: The F1-score combines the precision and recall of a classifier into a single metric by taking their harmonic mean. **Accuracy:**Accuracy tells you how many times the ML model was correct overall.

Cluster Rate: The cluster rate was denoted by CR, is used to quantify or model the weed density.

$$CR = \frac{\text{Weed plant coverage in the region (in pixels)}}{\text{Total land area of the region (in pixels)}}$$

Mean: The weed density for each tile is measured by the cluster rate. The estimated cluster rate (CRest) is compared against cluster rate in ground truth pixel-wise annotations (CRgt). The following three metrics are computed to quantify the error in weed density estimation:

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

IJARSCT

Impact Factor: 6.252

Volume 2, Issue 1, November 2022

$$\text{Absolute Error} = |\, CR_{gt} - CR_{est} \,|$$

$$\text{Mean Accuracy} = 1 - \sum_i \frac{\text{Absolute Error}/CR_{gt}}{N}$$

$$\text{MAE} = \sum_i \frac{\text{Absolute Error}}{\text{N}}$$

$$RMSE = \sqrt{\frac{\sum_i \text{Absolute Error}^2}{N}}$$

Mean intersection-over-union (mIoU) computed for both. It is observed that the unsupervised network produced a slightly higher score for the CWFID and significantly outperformed U-Net.

Model	Dataset	mIoU
UnSupervised Segmentation	CWFID	0.928
UnSupervised Segmentation	Sugar Beet Dataset	0.82
UNet	CWFID	0.913
UNet	Sugar Beet Dataset	0.76

Table-1: Quantitative evaluation for vegetation segmentation

The improvement in classifier performance due to weighted training using different approaches. This result substantiates previous finding by demonstrating that sampling techniques (random sampling and SMOTE) helps in improving the classifier performance for an unbalanced dataset. The performance is measured using the computed precision and recall for the weed class on the test set. While the precision and recall values improve relatively due to sampling techniques that address class imbalance, the absolute values remain below the acceptable threshold.

Random forest classifier achieved a recall of 1.0 but an extremely poor precision all tiles were predicted as weed infested.

To validate the choice of tile size, results from regions of different sizes were compared and retrained the classification models by both increasing and decreasing the side length (100 and 25 pixels, respectively) and reports precision and recall values for all the machine learning classifiers.

Once the weed-infested regions are identified, the cluster rate for each tile can be computed from the segmented vegetation pixels. A comparison of the estimated cluster rate for the weed infested regions with the ground truth values. The results show that the weed density can be estimated reasonably across both the datasets. The results reported for vegetation segmentation and weed distribution show that the proposed approach results in a mean absolute error of 5% on CWFID and 1% on Sugar Beets datasets. This indicates that the proposed method can handle the error arising due to above listed sources reasonably. Moreover, the RMSE of less than 8% on two datasets of two different crop/weed species demonstrates that the proposed approach is scalable and can be adapted to any crop/weed species. Once the weed plant distribution and density have been estimated, it is possible to decide about which regions should be selectively treated with agrochemicals.

Dataset	Mean Accuracy	MAE	RMSE
CWFID	75.24	5.02	7.85
Sugar Beets	82.13	1.62	3.06

Table- 2: Accuracy, MAE and RMSE of weed density estimation.

V. CONCLUSION

Precision agriculture is described as a farmland management approach to maximize productivity and profits in a sustainable manner. Agrochemicals, such as weedicides, are an expensive input for farming in addition to being detrimental to the environment. A semi- supervised approach to robustly estimate the Weed Density and Distribution Estimation for Precision Agriculture Using Semi-Supervised Learning weed density and distribution to aid precision agriculture is presented. This approach relies only on colour images as input. The first step is to generate a binary vegetation mask by removing all the background pixels. An unsupervised network is used to cluster the pixels into either background or vegetation. The second step is to overlay the mask on the input colour image and divide it into smaller regions (square tiles of side 50 pixels). These smaller regions are then classified as weed or crop. The performance of classifiers such as SVM, Gaussian Naive Bayes, Neural Network, and Random Forest which uses a pre-trained ResNet50 as a feature extractor. This approach is validated on two datasets Crop/Weed Field Image and Sugar Beets. Weed infested regions are identified with a maximum recall of 0.99 and weed density in these regions is estimated with an accuracy of 82.13%.

REFERENCES

[1]. Shorewala, S., Ashfaque, A., Sidharth, R., & Verma, U. (2021). Weed density and distribution estimation for precision agriculture using semi-supervised learning. IEEE access, 9, 27971-27986.

[2]. Jin, X., Che, J., & Chen, Y. (2021). Weed identification using deep learning and image processing in vegetable plantation. IEEE Access, 9, 10940-10950.

[3]. Sodjinou, S. G., Mohammadi, V., Mahama, A T. S., & Gouton, P. (2021). A deep semantic segmentation-based algorithm to segment crops and weeds in agronomic color images. Information Processing in Agriculture.

[4]. Selvi, C. T., Subramanian, R. S., & Ramachandran, R. (2021, March). Weed Detection in Agricultural fields using Deep Learning Process. In 2021 7th International Conference (ICACCS) (Vol. 1, pp. 1470-1473). IEEE.

[5]. Jogi, Y., Rao, P. N., & Shetty, S. (2020, November). CNN based Synchronal recognition of Weeds in Farm Crops. In 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 1373-1378). IEEE.

[6]. Das, M., & Bais, A. (2021). DeepVeg: Deep Learning Model for Segmentation of Weed, Canola, and Canola Flea Beetle Damage. IEEE Access, 9, 119367-119380.

[7]. Xu, K., Li, H., Cao, W., Zhu, Y., Chen, R., & Ni, J. (2020). Recognition of weeds in wheat fields based on the fusion of RGB images and depth images. IEEE Access, 8, 110362-110370.

[8]. Wang, A., Xu, Y., Wei, X., & Cui, B. (2020). Semantic segmentation of crop and weed using an encoder-decoder network and image enhancement method under uncontrolled outdoor illumination. IEEE Access, 8, 81724-81734.

[9]. Moazzam, S. I., Khan, U. S., Qureshi, W. S., Tiwana, M. I., Rashid, N., Alasmary, W. S., ... & Hamza, A. (2021). A Patch-Image Based Classification Approach for Detection of Weeds in Sugar Beet Crop. IEEE Access, 9, 121698-121715.

[10]. Razfar, N., True, J., Bassiouny, R., Venkatesh, V (2022). Weed detection in soybean crops using custom lightweight deep learning models. Journal of Agriculture and Food Research, 8, 100308.

[11]. Beeharry, Y., & Bassoo, V. (2020, November). Performance of ANN and AlexNet for weed detection using UAV-based images. In 2020 7th International Conference on Emerging Trends in Electrical, Electronic and Communications Engineering (ELECOM) (pp. 163-167). IEEE.

[12]. Asad, M. H., & Bais, A. (2020). Weed detection in canola fields using maximum likelihood classification and deep convolutional neural network. Information Processing in Agriculture, 7(4), 535-545.

[13]. Wu, Z., Chen, Y., Zhao, B., Kang, X., & Ding, Y. (2021). Review of weed detection methods based on computer vision. Sensors, 21(11), 3647.

[14]. Tejeda, A. I., & Castro, R. C. (2019, February). Algorithm of weed detection in crops by computational vision. In 2019 International Conference on Electronics, Communications and Computers (CONIELECOMP) (pp. 124-128). IEEE.

[15]. Islam, N., Rashid, M. M., Wibowo, S., Xu, C. Y., Morshed, A., Wasimi, S. A., ... & Rahman, S. M. (2021). Early weed detection using image processing and machine learning techniques in an Australian chilli farm. Agriculture, 11(5), 387.

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