International Journal of Recent Advances in Multidisciplinary Topics
Volume 3, Issue 10, October 2022
https://www.ijramt.com | ISSN (Online): 2582-7839

Crop Weed Prediction and Smart Crop Yielding

S. Aiswarya[1*], A. Alisha Jumaynah[2], D. Anukeerthana[3], M. Keerthana[4]
[1,2,3,4]*Student, Department of Computer Science and Engineering, Bannari Amman Institute of Technology, Erode, India*

Abstract*: Machine learning is used to identify crop weeds in the project "Crop Weed prediction and smart crop yields." In India, agriculture is one of the most significant and old professions. The production of food must be handled with the utmost care because agriculture is the foundation of India's economy. Plants become infected by weeds like viruses, fungi, and bacteria, which results in decreased output of both quality and quantity. Farmers are leaving the agricultural industry in large numbers. Therefore, taking good care of plants is essential for the same. Image processing offers more effective approaches to find weeds on plants that are brought on by fungus, bacteria, or virus. Weed detection by merely using the eyes is ineffective. The quality of plant nutrients is also harmed by overuse. Farmers suffer a significant loss in production as a result. Therefore, it is useful to apply image processing techniques to identify and categorize weeds in agricultural applications. The agriculture industry has great potential to reduce food shortages and supply wholesome, nutritious food. Farmers face a difficult problem when trying to identify crop weeds since weed invasion causes significant crop loss and quality degradation. The disadvantage of traditional weed identification is that it requires skilled taxonomists to correctly identify weeds based on physical characteristics. In order to identify crop weeds early on and shorten the time needed to improve crop production and crop quality, classification and prediction accuracy findings are used.

***Keywords*: Crops, weeds, classification, accuracy, prediction, detection.**

1. Introduction

System Specification:
Hardware Configuration:
Hard Disk: 450 GB
Keyboard: 110 Keys
Monitor: 17-inch LCD display
Mother Board: Intel
Mouse: HP Mouse
Processor: Core 2 Duo
RAM Capacity: 2.5 GB
Speed: 3 GHZ
System bus: 32-bit
Software Configuration
Operating System: Windows 7 Home Basic
Front End: Python

A. Existing System

Periodic weed outbreaks caused by plant weeds result in widespread mortality and hunger. The 1943 epidemic of helminthosporium of rice in north-eastern India is thought to have resulted in a significant loss of food crops and the deaths of one million people. Some crop farming has been abandoned because of the catastrophic impact of plant weeds. Plant weed losses are projected to have cost Georgia (USA) $653.06 million in 2007. (Jean, 2009). Because the preventive measures we take to protect our crops are not even one-tenth as extensive as those in the USA, the estimate for India is higher than for the USA.

1) Drawbacks

The primary strategy used in practice for finding and identifying plant weeds is the naked eye observation of experts. However, huge farms may not be able to afford the excessively expensive expert monitoring that is required on a constant basis. Additionally, in some poor nations, farmers may need to travel great distances to see experts, which makes doing so expensive and time-consuming. Farmers may also be unaware of non-native weeds.

In order to augment the current delivery channels offered by the department, it is planned to make pertinent information and services available to the farming community and the private sector through the use of information and communication technology. In order to address all of an Indian farmer's informational demands on agriculture, Farmers' Portal is an effort in that direction. A farmer will be able to get all pertinent information on specific topics pertaining to their village, block, district, or state once they are on the Farmers' Portal. The farmer phones this service centre through the farmers' portal and gets all of his questions answered. People at the service either look for pertinent information on their own or have received extensive training on how to assist farmers. This has the drawback that nobody at the help centre can see the precise issue the farmer is having. They sometimes identify weeds incorrectly because they are unable to fully comprehend the seriousness of the problem. The crop could be destroyed as a result of this.

B. Proposed System

In the suggested approach, the farmer initially provides the photographs. Through the Python application created specifically for the farmer's use, the photographs are received from the farmer. The farmer selects the right image of the leaf or stem, preferably, from the Choose File option, and then uploads it. The farmer obtains an ID after uploading an image, which he must use later to check the manures for the problematic weed. The JUPYTER NOTEBOOK processes the

*Corresponding author: aiswarya.cs19@bitsathy.ac.in

image that the farmer uploads, and image-processing algorithms are then applied to the obtained photographs to extract important features that are required for additional analysis. The photos are then classified using a variety of analysis approaches in accordance with the particular issue at hand. The JUPYTER NOTEBOOK detects the marijuana kind and displays it, along with the affected area and the herb's severity. The database is updated with the manures for the discovered weed and the recommended dosage for the plant. The farmer must select the View Message button on the app in order to view the details. The farmer can examine the uploaded details by entering the ID that was previously displayed to him.

1) Features

- The automatic detection of plant weeds is a crucial research area since it may be useful for keeping an eye on vast fields of crops.
- As a result, weeds can be identified automatically based on symptoms that show up on plant leaves.
- This makes machine vision possible, which will allow for image-based process control, robot navigation, and autonomous inspection.
- Visual identification, in contrast, requires a lot of physical labour, is less precise, and can only be done in a narrow area.

C. Modules

The following modules are used for detecting weed in crop:

- Image Preprocessing
- Segmentation
- Feature Extraction
- Classification

1) Image Preprocessing

The acquired image undergoes preprocessing. The RGB image is first converted to L*a*b* colour space as part of the preprocessing. The chromaticity layers a* and b*, as well as the luminance layer L*, make up the L*a*b* colour space. The layers a* and b* include all of the colour data. To convert an RGB-colored image to L*a*b* space, colour form must be created. The format is eventually applied to the obtained image using the makecform() function.

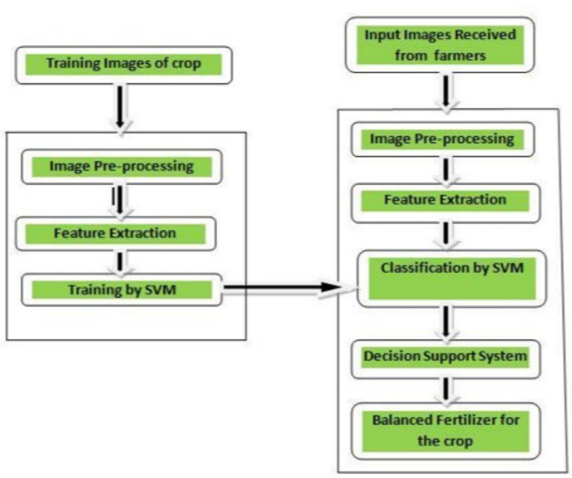

Fig. 1.

2) Segmentation

While there are several segmentation algorithms in use, k-means clustering is one of the better techniques for weed detection. K-means clustering is a prominent vector quantization technique for cluster analysis in data mining. It was originally developed for signal processing. The goal of k-means clustering is to divide n observations into k clusters, each of which is made up of the observations that belong to the cluster that has the closest mean to it and acts as the prototype of the cluster. As a result, the data space is divided into Voronoi cells. While the expectation-maximization method generally finds clusters with similar spatial size, k-means clustering tends to find clusters of varied shapes.

The k-nearest neighbour classifier, a well-known machine learning technique for classification that is frequently confused with k-means due to the k in the name, and the algorithm share some similarities. To categorise new data into existing clusters, one can use the 1-nearest neighbour classifier to the cluster centres discovered by k-means.

Applying k-means clustering, categorise the colours in the a*b* colour space. We make three clusters because the image contains three colours. Euclidean Distance Metric is used to measure distance. Make a list of every pixel in that image using the K-means results.

K-means clustering attempts to divide a set of observations (x_1, x_2,..., x_n), where each observation is a d-dimensional real vector, into k (n) setsS = S_1, S_2,..., S_k in order to reduce the within-cluster sum of squares (WCSS) (sum of distance functions of each point in the cluster to the K center). In other words, it seeks to discover:

$$\underset{\mathbf{S}}{\arg\min} \sum_{i=1}^{k} \sum_{\mathbf{x}\in S_i} \left\| \mathbf{x} - \boldsymbol{\mu}_i \right\|^2$$

Where μ_i is the mean of points in S_i.

The task of grouping a set of items in a way that they are more similar (in some way or another) to each other than to those in other groups is known as cluster analysis or clustering (clusters).

The k-means algorithm can be used to divide the input data set into k parts for cluster analysis. Using the findings from the K means, label each pixel in the image. The clustering results are then constructed and stored in a blank cell array. Then, using pixel labels, create an RGB label. Another crucial factor is choosing the right cluster. The cluster that shows the most weed-affected area should be chosen. The features of the chosen cluster are extracted in the subsequent stage of feature extraction.

3) Feature extraction

The retrieved features of the chosen cluster. Since the chosen image is in RGB format, it is converted to grayscale. The Gray Level Co-occurrence Matrices are used as the following stage (GLCM). Gray level co-occurrence matrices are used to get the necessary statistics (GLCM). Contrast, Correlation, Energy, Homogeneity, Mean, Standard Deviation, Entropy, RMS Variance, Smoothness, Kurtosis, and Skewness are among the

13 features that are extracted and analysed. An array is used to store the thirteen features.

4) Classification

The idea of decision planes, which specify decision boundaries, serves as the foundation for support vector machines. A decision plane is a diagram that distinguishes between a collection of objects with various class memberships.

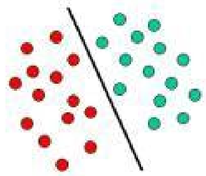

Fig. 2.

The example given above is a basic illustration of a linear classifier, or a classifier that uses a line to divide a collection of items into their respective groups (in this case, GREEN and RED). However, most classification problems are not that straightforward, and frequently more complex structures are required to achieve an optimal separation, i.e., correctly categorize new objects (test cases) based on the examples that are already available (train cases). The image that follows shows this circumstance. It is evident from the previous graphic that a curve would be necessary to fully separate the GREEN and RED items (which is more complex than a line). Adding dividing lines to hyperplane classifiers for classification problems. These tasks are best handled by Support Vector Machines. When your data contains precisely two classes, users can employ a support vector machine (SVM). An SVM sorts data into classes by locating the optimal hyperplane that divides all of the data points in one class from those in the other. The hyperplane with the biggest margin between the two classes is the optimum hyperplane for an SVM. Margin is the maximum thickness of the slab that is perpendicular to the hyperplane but has no interior data points. The data points on the slab's edge that are closest to the separating hyperplane are known as the support vectors.

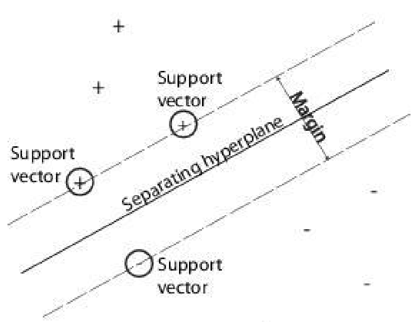

Fig. 3.

Using Support Vector Machines: As with any supervised learning model, a support vector machine is trained first, and the classifier is then cross-validated. Use the machine that has been trained to classify (predict) fresh data. Additionally, you can employ a variety of SVM kernel functions and adjust their parameters to get a sufficient level of predicted accuracy.

The support vector machines must first be trained before they can be used. Images of weed rice crop plants were used to train the SVM for weed detection. You can get the photos for Brown Spot and Rice Blast weed from this page. Additionally, the photos' characteristics were retrieved. And a dataset with image-related features contained all of the data. They are known as training images. The whole process is known as training the SVM.

The function used is: svmStructWeed = svmtrain(our,dt);

Where the dataset used by the system is one that contains feature information from images. "dt" refers to the set that contains the particular set of weeds. The photographs are then categorised as the next phase. The function is as follows:

svmclassify(svmStructWeed,feat_weed)

The array feat weed holds the characteristics of the chosen picture utilized for testing. The weed is discovered after choosing the right cluster. Additionally, a help window that displays the impacted weed is presented.

Rice plant classification and training are done with the SVM. The dataset contains two weed plant features. The Multiclass SVM seeks to label instances by utilizing support vector machines, where the labels are generated from a finite set of several items, and can be extended to several weeds. Reducing a single multiclass problem to a number of binary classification questions is the most popular method for doing this. These are typical techniques for such reduction:

Building binary classifiers which distinguish:

1. One label versus all the others (one-versus-all) or
2. In between each set of classes (one-versus-one).
3. A winner-takes-all technique is used to classify new examples for the one-versus-all case, in which the classifier with the highest output function assigns the class (it is important that the output functions be calibrated to produce comparable scores). To classify instances in the one-versus-one technique, a max-wins voting strategy is used. In this method, each classifier assigns an instance to one of the two classes, increases the vote total for the allocated class by one vote, and then chooses the class that has received the most votes overall. SVM for a directed acyclic graph.

Error-correcting output codes:

Instead of breaking the multiclass classification problem down into numerous binary classification problems, Crammer and Singer devised a multiclass SVM method that transforms the problem into a single optimization problem. We have developed a dataset of numerous photos that includes each of the thirteen features derived using feature extraction for the current proposed method for the detection of weeds. A label set comprising integer values for the appropriate weed kind was produced from a training batch of many photos. A specific type of marijuana is identified upon maximal feature matching, and the help dialogue displays it.

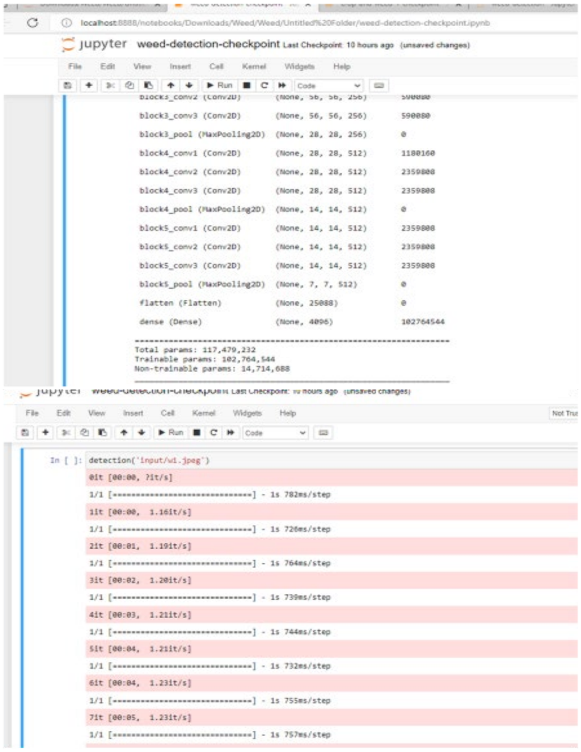

Fig. 4. Output

2. Conclusion

The methodology for classifying crops and weeds using machine learning algorithms is proposed in this paper. The technique enables us to automatically select one of two CNN models. Ensemble models are constructed using models with different data usages and objective functions.

References

[1] J. Hemming and T. Rath, "Pa—precision agriculture: computer-vision-based weed identification under field conditions using controlled lighting," Journal of agricultural engineering research, vol. 78, no. 3, pp. 233–243, 2001.

[2] X. Ma, X. Deng, L. Qi, Y. Jiang, H. Li, Y. Wang, and X. Xing, "Fully convolutional network for rice seedling and weed image segmentation at the seedling stage in paddy fields," PloS one, vol. 14, no. 4, 2019.

[3] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," IEEE transactions on pattern analysis and machine intelligence, vol. 39, no. 12, pp. 2481–2495, 2017.

[4] X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu, and P. A. Heng, "H-denseunet: hybrid densely connected unet for liver and tumor segmentation from ct volumes," IEEE transactions on medical imaging, vol. 37, no. 12, pp. 2663–2674, 2018.

[5] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 3431–3440.

[6] C. C. Andrea, B. B. M. Daniel, and J. B. J. Misael, "Precise weed and maize classification through convolutional neuronal networks," in 2017 IEEE Second Ecuador Technical Chapters Meeting (ETCM). IEEE, 2017, pp. 1–6.

[7] B. Liu and R. Bruch, "Weed detection for selective spraying: a review," Current Robotics Reports, vol. 1, no. 1, pp. 19–26, 2020.

[8] J. Redmon and A. Farhadi, "Yolo9000: better, faster, stronger," in proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 7263–7271.