3rd International Conference on
Robotics Automation and Non-Destructive Evaluation,
Chennai, India, 23 April 2022.
https://doi.org/10.13180/RANE.2022.23.04.20

MACHINE LEARNING APPROACHES TOWARDS SUGARCANE CROP DISEASE IDENTIFICATION

D. VISHNU SAKTHI[#1] , L. SHERIN BEEVI[#2]
T. VENKATA SUKESH[#3] ,V. VENU VITTAL[#4] and U. SOMA SEKHAR[#5]
[#1,#2]*Assistant Professor, Department of Computer Science and Engineering,*
R.M.D Engineering College, Tiruvallur,601206
[#1] vishnu.cse@rmd.ac.in [#2] sherinbeevi@gmail.com
www.rmd.ac.in

[# 3,#4,#5]*Student, Department of Computer Science and Engineering,*
R.M.D Engineering College, Tiruvallur,601206
[#3] sukesh992002@gmail.com, [#4] venuvittalvemuri@gmail.com,
[#5]usoma56@gmail.com

Sugarcane is an important crop which plays a vital role in overall revenue in agriculture. Almost many of countries in world are involved in sugarcane plantation and pests affect sugarcane yield or production. The use of decision support system in agriculture for crop disease identification is made possible using machine learning algorithms. We evaluate the performance of machine learning techniques for the sugarcane crop in this research. We focus on mainly four diseases that affect of yield of sugarcane crop and we train machine algorithms using the images of these crop diseases using CNN, CNN and SVM, XGBoost algorithm. The performance of these algorithms for the classification is evaluated as to how it synthesizes the features and correctly classifies the disease.

1. Introduction

Early disease detection can help you prevent losing a lot of money on your crop. If feature extraction is not performed on photos from farmland, it is difficult to anticipate images using machine learning methods such as support vector machines and BP (Backpropagation neural networks), and prediction accuracy is not adequate. Due of the complex environment in farming, classical machine learning approaches struggle to identify unique traits of all pests in farmland; this is where CNN comes in handy. Normally, there are a lot of species of pests with different or same appearance are existing in fields so that CNN plays important role in the classification of pests under their category[1]. Smart visualizing (visualizing through gadgets) is easy in comparison to manual identification, and also nowadays many are using smartphones so image recognition can play important role in pest identification. The CNN can extract the images and achieve good results. So, this can be used as a feature extraction tool. CNN has an interference process for identification and a backpropagation process for training images in the testing dataset [2].

2. Literature Review

Sugarcane is the most important crop that sees a worldwide production there are various products produced from the crop.

S. Sivasakthi et al 2019 Plant Leaf disease identification: K-Nearest Neighbor (KNN), Radial Basis Function (RBF), Probabilistic Neural Networks (PNN), Backpropagation Network (BPN) and Support Vector Machine (SVM). this process is an alternative to any training method, a simpler classifier can be used. It is a faster training method, and the hidden layer is simple to understand. It is capable of tolerating excessively loud inputs. Instances classified an excessive number of outputs that were adaptable to changing data. It is easy to implement. It is used in various types of problems. It is a simple geometric interpretation. now we will focus on the drawbacks of this work that is the speed of compilation is slow. It shows the output very late. It has very high memory space. Learning can be a slow process. It's difficult to say how many layers are required. Training should be done gradually. It's difficult to comprehend. [3]

Rajeswary et al 2015 Pest identification: Method with Static images, Method which uses Sticky Traps, Method which uses pan-tilt camera with 20 x zoom. It has good results and accuracy, It is excellent for identifying the pests, This PTZ camera has 2.1 megapixels and 3600 degree rotation. Now let us see the drawbacks of the method, the scanner only uses it and it is excellent difficult, if insects move away from the camera while taking the photo it is difficult to tilt the camera and zoom, and it gives very poor image quality.[4]

Rupesh et al 2013 Pest identification: Method with static images, method of sticky traps. It has good results and accuracy. It is a great method for identifying pests and great in identifying the flies. Let us see the drawbacks of this work. There is a chance of improper scanning and assisting in false information, It is only for identifying the insects but the damage is done and also has very poor picture quality.[5]

R Rani et al 2017 Pest Identification: Image Acquisition, Image Processing, K-Means Clustering Algorithm, SVM Classifier. It is used to select pictures of the leaves which are non-defective or defective, It is an image enhancing technique that increases the dynamic range of the intensity values in an image, It is an algorithm to categorize the image segregation, and also it is an unsupervised technique that separates the unwanted data from the background, It is a strong binary classifier and can generate extremely fast classifier. Let us see the drawbacks of this work, It will take only the photos and then divide them into leaf images and it is a long process, it has good results when it is implemented before the K-Means algorithm, It is an unsupervised learning method.[6]

Chinnayan et al 2019 IoT-based leaf diseases identification: Leaf Disease classification, Data Acquisition, Image Edge Detection, Image Optimization, and comparison. It is used for selecting and classifying the images, RGB images are stored in jpg format and then they are changed to greyscale, because of data acquisition it is great because most of the images are greyscale and easy to classify. Let us see the drawbacks of this work, Its great disadvantage is selecting the pictures

will match to the pictures which are defined by the user, it can only be stored in .jpg or .png format, It is useless for the noisy images and also high-frequency pictures.[7]

DevenJ Patel et al 2021 Pest Detection: Faster R-CNN Architecture, Image Augmentation, Implementation. It is a pathway between CNN it enhances the images very quickly and sends to CNN, To enlarge the image proportion and also the angle of the picture in both Horizontal and Vertical it is used, All the experiments were performed by a high-speed Quadro P5000 which makes us easy to take photos by our phone. Let us see the drawbacks of this work, R-CNN sometimes is very slowest algorithm due to the size and quality of the image, some pests are seasonal and they are hard to identify this kind of pests, In the implementation technique, all the images are stored in an XML file which makes us difficulty to open or access the file.[8]

3. Methodology and Description

CNN is introduced to classify crop disease by means to predict which category the pest belongs to. By CNN we can achieve remarkable accuracy around 96.08% by training our algorithm. CNN is based on a deep learning algorithm lot of parameters are required to train our testing data set and by convolution of all hidden layers we get desired output. Deep learning is a sort of machine learning that learns from many representations of the same input set using deep neural architectures with a large number of hidden layers[9]. The same approach may be applied to categorize multiple data sets in different fields [10]. Our CNN model can efficiently learn intrinsic properties of images from our training data set for classification.

In this paper performance of CNN is compared with traditional ML technique SVM and performance of CNN is evaluated along with XGBoost boosting algorithm for sugarcane crop disease identification.

CNN: CNN (convolutional neural networks) is an Image recognition process in machine learning. It is deep learning neural network architecture for processing structured designs in images such as lines, gradients, circles or even small features. CNN do not need any preprocessing like any other methods CNN can directly process the image. We wish to use an algorithm other than CNN to handle all filters. In comparison to traditional neural networks, CNNs perform better with image, speech, and audio signal inputs. They are made up of three primary layers.

1) Convolution

2) Polling

3) Fully connected layer (FC)

3.1. Convolution Layer

The convolution layer is an important part of CNN, as it is where the majority of the processing happens. It requires input data, a filter, and a feature map, among other things. Assume the input is a 3-Dimension colour image composed of a matrix of pixels, which means the image will have three dimensions: width, height, and RGB. We also have a feature identifier, sometimes known as

a filter, that moves through the image's matrix of pixels to detect the presence or absence of a feature. Convolution is the term for this procedure.

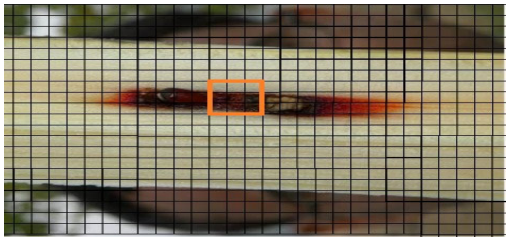

Figure 1.Filter across pixels of image to extract features

The filters are made up of a two-dimensional array of weights that each represent a section of the image. The filter size is usually a 3x3 matrix that corresponds to the size of the field. After that, the filters are represented by an area of the image, and the dot product between the input pixels and the filters is determined. The output array receives this dot product, and the filters are moved by a stride, and the procedure is repeated until the filters have covered the entire image.

Each process' output is compiled into a matrix, which we refer to as a feature map.

The outcome can be influenced by the number of filters used. We have n feature maps if we have n filters. After each convolution, CNN does the RELU (rectified linear unit) operation, which renders our feature map non-linear.

3.2. Pooling Layer

Dimensionality reduction is also carried out through the pooling layer. The pooling procedure applies to all filters throughout the whole input, with the exception of this filter, which has no weights. In the pooling layer, a lot of data is lost. The efficiency of our model increases and our model complexity decreases when we employ the pooling layer.

3.3. Fully Connected Layer

To appropriately identify inputs, the fully linked layer employs the SoftMax activation function. It generates a probability ranging from 0 to 1.

3.4. SVM (Support Vector Machine)

SVM is one of the most used supervised learning algorithms. SVM is useful for both regression and classification. The purpose of SVM is to create a decision boundary that can divide n-dimensional spaces into classes, allowing future data points to be labelled into the appropriate category. A hyper plane is a term used to describe the optimal decision boundary.

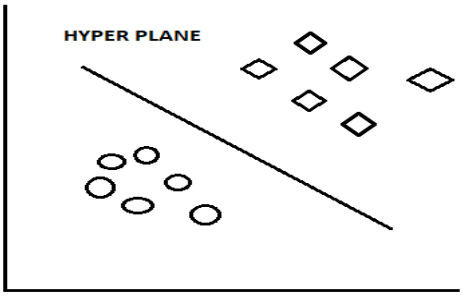

Figure 2. Hyperplane separating two classes

The data points nearest to hyperplanes are called support vectors, and they have an effect on the hyperplane's position.

3.5. XGBooster

XGBoost algorithm is to boosting of framework it speeds our model and it utilizes memory efficiently so XGBoost makes an important role in classification and regression. XGBoost is a better algorithm than any other machine learning algorithms to provide a good solution; It has become the "state-of-the-art" algorithm for dealing with structured data.

4. Diseases Classified

4.1. Grassy Shoot

Grassy shoot appears at 2nd month after plantation. The disease is classified by the production of enormous lanky tillers from the base of the affected shoots. The primary spread of this disease is due to the cutting knifes and diseases setts.

Figure 3. Grassy shoot

4.2. Scorching the Leaf

Browning of plant tissues, including leaf edges and tips, as well as yellowing or darkening of veins, is known as leaf scorch, and it can cause wilting and leaf abscission. They typically appear 3 to 8 days after inoculation. Leaf scorch is produced by stagonospara sachhari infections, which results in a reduction in photosynthetic activity.

Figure 4. Leaf scorch

4.3. Red Rot

Red rot is very known disease in sugarcane worldwide it is also known as cancer of sugarcane it is caused by Colletotrichum falcatum Went. 29.4c to 31c temperature is favorable for red rot disease.

Figure 5. Red rot

4.4. WILT

This disease is mainly occurring at the end phase of the growth, the leaves turn yellowish green and loose firmness and eventually dry this disease starts from the base crowns turn white or yellowish midrib surrounded by a pale green blade. The affected canes are stunted and lose weight.

Figure 6. Wilt

5. Results and Discussion

Table.2.ResultAnalysis of different algorithms on sugarcane crop disease dataset

Method	Accuracy
CNN	94.56
CNN+SVM	97.56
CNN+XGBoost	98.54

The above table shows the performance of various machine learning algorithm for sugarcane crop disease prediction a dataset of 1500 images were used to train various algorithm and we are able to identify that CNN performs well along with the gradient boosting framework and is able to predict sugarcane disease with better accuracy and also the performance of CNN is improved using XGBoost framework.

References

1. Xi Cheng, Youhua Zhang, Yiqiong Chen, Yunzhi Wu, Yi Yue, Pest identification via deep residuallearning in complex background, Computers and Electronics in Agriculture, Volume 141, 2017

2. Feng Cen, Xiaoyu Zhao, Wuzhuang Li, Guanghui Wang, Deep feature augmentation for occluded image classification, Pattern Recognition, Volume 111, 2021, 107737, ISSN 0031-3203
3. Sivasakthi, S. (2019). The capitalized title is Plant Leaf Disease Identification Using Image Processing and SVM, ANN Classifier Methods.
4. Rajeswary, &Divya (2015). Identification and Classification of Pests in Greenhouse Using Advanced SVM in Image Processing.
5. Mundada, Rupesh G. and V. Gohokar. "Detection and Classification of Pests in Greenhouse Using Image Processing." IOSR Journal of Electronics and Communication Engineering 5 (2013): 57-63.
6. Rani, R.. "Pest Identification in Leaf Images using SVM Classifier." (2017).
7. Reddy, H.S., Hedge, G., &Chinnayan, P.D. (2019). IoT based Leaf Disease Detection and Fertilizer Recommendation.
8. Patel, Deven J. and N. Bhatt. "Improved accuracy of pest detection using augmentation approach with Faster R-CNN." IOP Conference Series: Materials Science and Engineering 1042 (2021): n. pag.
9. Suganya, E., S. Sountharrajan, Shishir K. Shandilya and M. Karthiga. "IoT in Agriculture Investigation on Plant Diseases and Nutrient Level Using Image Analysis Techniques." (2019).
10. Chauhan, Rahul, K. Ghanshala and R. Joshi. "Convolutional Neural Network (CNN) for Image Detection and Recognition." 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC) (2018): 278-282.