Chinese Journal of Evidence-Based Medicine

Ensemble Deep Learning Applied in Precision Farming Through Controlled Fertilisation, Disease Mitigation, Optimised Harvesting and Marketing

Pushpalatha M	**Rajeswari V**	**Aparna C M**	**Gowtham P**	**Pradeep A V A**	**Vasu Prem M**
Assistant Professor	Professor	Final Year/CT	Final Year/CT	Final Year/CT	Final Year/CT
Department of CT	Department of CT	Department of CT	Department of CT	Department of CT	Department of CT
Karpagam College of Engineering	Karpagam College of Engineering	Karpagam College of Engineering	Karpagam College of Engineering	Karpagam College of Engineering	Karpagam College of Engineering
Coimbatore, India	Coimbatore, India	Coimbatore, India	Coimbatore, India	Coimbatore, India	Coimbatore, India
pushpalatha.m@kce.ac.in	rajeswari.cst@gmail.com	cmaparna25@gmail.com	gow8004@gmail.com	avapradeep3@gmail.com	vasuprem.m@gmail.com

***Abstract*— Agriculture is considered as largest economy in the world especially in India as it provides huge no of employment and nutrient benefits of various crop and plant for human survival. However, due to reduced no of research in agriculture field has impacted the farmer on basis of crop yield. In order to enhance the yield of crop, various researches has been conducted in past decade especially in terms of detecting the crop diseases. In specific, it becomes mandatory to analyze the suitable crop to specific soil condition and environmental condition. Meanwhile application of fertilizer is significant to increase the soil condition for crop cultivation and increase crop yield, Thus, a ensemble deep learning architecture is designed and implemented with functionalities of crop recommendation. Plant disease identification is becoming significant challenge all over the world, therefore automatic detection and arrangement of the disease of plant leaf is very important in monitoring the plant. In this paper, ecommerce application is developed using ResNet architecture towards plant disease prediction. Additionally, aggregation of Random Forest algorithm and Support vector Machine is used for fertilizer recommendation on basis of soil and environmental conditions. Convolution neural network is modeled to classify the leaf image into multiple disease classes. However convolution layer and max pooling layer of the architecture produces the features map in order to increases the classification efficiency. Further ReLu activation function is employed in fully connected layer to avoid the over fitting issues to enhance the scalability and accuracy of the model on classify feature map into various diseases classes of the plant leafs as leaf blight, gray mold, powdery mildew. Experimental analysis of the model is carried out using plant-village dataset. Finally efficiency of the model is evaluated on basis of accuracy and loss to the training and validation data.**

Keywords*— *Plant Disease, Convolution Neural Network, Random Forest, Support Vector Machine, Fertilizer Recommendation, Deep learning.

I. INTRODUCTION

Plant is represented as cultivated crops in the agriculture regions around the world due to its high nutritional content and commercial value. Hence it is mandatory to monitor the soil condition , environment condition for effective crop cultivation to obtain high yield. However, plant disease have increased owing to variation in the cultivation systems and pathogen changes such as leaf scorch, gray mold, Crown leaf blight, anthracnose crown, leaf spot, leaf blight round spot and rust. In order to manage those complications, manual diagnosis of the plant disease in laboratories by experts on basis of disease severity has been carried traditionally whereas it is highly costly and time-consuming. Therefore automatic method for identification leaf disease on characterizing the disease using image processing and recognition techniques is very crucial in diagnosing the plant leaf with respect to the leaf appearance and diseases symptoms. Further diagnosis is indispensable to control and manage the plant production.

Machine learning based data processing system is considered as accurate and inexpensive tool for predicting and recommending suitable crop and fertilizer on basis of the soil parameter and environment parameters of the particular agriculture region. Especially Supervised learning model like Regression model, Random forest and support vector machines (SVMs) has been employed in large extent towards the prediction of fertilizer and suitable crop to particular agriculture region. Recently, Deep learning has gained popularity and immense advances on processing the images in accurate and efficient manner through extracting convenient feature of the input images for discriminant class representation. Especially deep learning model like Convolutional Neural Network analysis and artificial neural network are considered as highly capable models for classifying the various category of plant disease.

Initially random forest algorithm is used to the process the soil parameter and environment parameter to construct a decision tree with feature with highest information gain. Decision tree is composed of bagging approach to combine the relevant parameters on splitting and replacement of suitable parameter. Next , Decision Tree is processed using Support Vector Machine to recommend the suitable fertilizer to the feature of the soil. Convolution neural network, a deep learning technique has been proposed in this research work by optimizing the activation function with ReLu function to eliminate the network over fitting issues and to provide a timely and accurate detection on plant diseases into various classes. The proposed model has capability in enhancing the precision agriculture along surging the precision of plant protection.

The residual portion of the article is sectionized as; segment 2 represents the review of literature work including the ML techniques and techniques of deep learning for disease identification and cataloguing of plant. Section 3 to define the proposed methodology for plant disease classification using Convolution Neural Network, fertilizer recommendation using random forest and support vector machine. Section 4 discusses the experimental result of the model. Section 5 ends the articles.

II. Related Work

Review of the literature on plant disease identification and recognition on different disease categories using different deep learning and machine learning architectures that process the dataset almost exactly like the suggested architecture was covered below, along with a description of various public datasets, is presented in this section.

A. *Leaf Disease Identification using DMS-Robust AlexaNet*

Identification of leaf diseases has been carried out using DMS-Robust Alexnet, a deep learning technique that combines image enhancement and recognition by avoiding extreme deformation of the transformed images. The feature extracted is considered an important abstraction to categorize the images of well-being and damaged leaves. It maintains a correlation of the features on its spatial and temporal aspects using the connectivity pattern of the neural network on adjacent layers. The ReLu activation unit was used to generate the class label.

B. *Diseases Classification using Multi-Scale Convolutional Global Pooling Neural Network*

Multiscale Convolutional Neural Network has been modeled on the basis of the AlexaNet to identify maize diseases. Features extraction is carried out in the convolution layers to avoid over-fitting issues as the disease region contains a large number of parameters. Global pooling layer is processed such that it transforms the original fully-connected layer to acquire the methods of the transfer learning to manage insufficient sample data with epoch tuning to increase the recognition performance of plant diseases.

III. Proposed System

In this section, ensemble deep learning architecture is modeled for crop and fertilizer recommendation to the specified agriculture region using random forest, logistic regression and support vector machine on using soil parameters and environmental parameters. Furthermore plant disease prediction is modeled using Convolutional Neural Network on both the real time image or through plant village dataset.

A. *Fertilizer Recommendation*

Fertilizer Recommendation is carried out on acquiring the parameter details of the soil and environment. Acquired details are processed in the machine learning classifier

Chinese Journal of Evidence-Based Medicine

represented as Random forest and Support Vector Machine to identify the suitable fertilizer on processing the fertilizer index to the condition of the soil in the particular region.

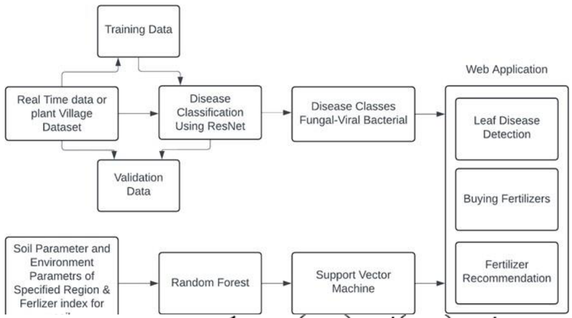

Fig. 1. Architecture of the proposed Model

- *Random Forest*

Random forest algorithm is processed with acquired soil and environment parameters to generate the decision tree. Decision Tree is obtained on aggregating the features through bagging and bootstrapping method with majority voting and high variance computations.

$$CCM_{m,n}(I,j) = \sum_x \sum_y^n x \begin{cases} 1 & m(x,y) = i(x,y) = j \\ 0 otherwise \end{cases}$$

Majority voting computation uses correlation to determine the soil condition of the agriculture region. It is given by,

$$\text{Classes of the Soil} \quad p\sigma^2 = \frac{1-p}{B}\sigma^2$$

where b is a random variable, p is correlation and σ is variance.

- *Support Vector Machine*

Support Vector machine is used to predict the suitable fertilizer classes for decision tree classes of the random forest. Decision Tree class variable containing soil and environment features and fertilizer index value is distributed in the hyperplane.

$$G_{func}(u) = \Delta.O_{data}(u) + (1 - \Delta).MFO_{reg}(u)$$

On Hyperplane maximum margin separator is used for predict the fertilizer to the specified soil conditions. Support Vector is determined and it is provides suitable fertilizer to the soil condition of the specified region.

$$\text{Support Vector} = \sum w_s K(x^s)$$

B. Neural Network ResNet Convolution CNN

CNN is designed to classify the features of segmented infected regions into various classes of diseases. In order to perform classification using the Convolutional Neural Network, optimization is carried out in the activation layer of the ResNet model with ReLu function. In this work, more parallel convolutional layers of the segment with a size of 2
× 2 and 5 × 5 and max pooling layer for the acquiring of dissimilar features continuously has been included.

Specifically, the proposed deep learning incorporates two fully connected layers, two Convolutional layers, Activation

$$L(f) = \frac{1}{n}\sum_{k=0}^{L-1}\{\sum_{x}[y\log(p(i,j))+(1-y)\log(1-P(i,j)$$

layer, Soft max layer and a loss layer. Next, every model is composed of activation function which uses *ReLu* operation and softmax function containing classifier to classify the disease features. Further correlation among the various features from the max pooling layers will be combined to yield high recognition accuracy. A dropout function is incorporated to reduce over-fitting issues of CNNs. The *ReLu* function is used to learn non linear features.

- *Convolution Layer*

Convolution layer is the significant layer which generates feature maps on employing kernel functions of the Neural Network. The convolution layer maps the segmented pixel on the basis of the continuous sliding convolution window. In the proposed architecture, a feature map is generated to input segments. For segmented feature in the Convolutional layer, Feature Map is calculated as

$$H_{ic} = f(w_i * x)$$

Where Hic convolution operation considers the convolution layer kernels, and f considers the activation function.

$$W=[W_i^1, W_i^1, W_i^{1,\ldots,} W_i^k]$$

K denotes the quantity of kernels within the convolutional layer.

- *Activation Function*

The Activation function will employ ReLu operations to process the feature vector in matrix form.. It is considered a nonlinear function which combines operation of sigmoid and tanh function to calculate the categorization of the feature map on the basis of features. This feature is considered as disease types with fungal, viral and bacterial categories.

- *Pooling Layer*

On Increase of the Convolutional layers for the feature processing, the network parameters will enhance the model for classification. The pooling layer is employed to minimize the number of features while simultaneously reducing the network parameter. Further it computes the efficient feature for softmax function which operates as classification components. In addition, it considers only statistical characteristics of a segment instead of entire characteristics of the segment.

- *Completely linked film*

The function of the fully connected layer is to categorize the feature maps produced by the neural networks' different levels. The softmax method will generate the corresponding class structure for each feature map. Subsequent class structures distribute the feature weight at a comparable rate, assigning dissimilar features to distinct classes. Ultimately, the dropout layer is added to improve model generalization and lessen over-fitting problems.

- *Loss Function*

On the computed classes, the loss function computes the discriminant features. The algorithm of stochastic gradient descent (SGD) is used to remove discriminant features for every class.

Chinese Journal of Evidence-Based Medicine

The P represented a probability matrix of the feature of the convolution and pooling layer, n is considered as no of features to fully connected layer, i, feature index is j, and the class index is k. Network training computes the feature weight W which helps to reduces the loss function E where W is iteratively iterated and is represented as

$$W_k = W_{k-1}\partial(\partial E(w)/\partial w$$

Where ∂E(w) is referred to as a significant parameter to calculate the size feature vector to each layer. Figure 2 represents the disease infected region.

Fig. 2. Disease infected segments of the plant

- *Hyper Parameter Tuning*

In this part, hyperparameters of the Densenet169 model were fine tuned to generate the increased recognition accuracy on employing the Adam optimizer as it is highly capable in modifying the base learning rate of the network during various iterations of the classifier. Further Model optimization was carried out using a stochastic gradient descent approach.

TABLE I. HYPERPARAMETER TUNING OF THE CONVOLUTION NEURAL NETWORK

Hyperparameter	***Value***
Model Learning rate	0.02
Epoch	35
Activation Function	ReLu
Loss Function	Sigmoid
Model Batch Size	15

In this work, ResNet architecture of convolution Neural Network family was utilized in this research to attain the increased appreciation correctness on discovery of plant illness on processing the feature of the segmented regions into viral, bacterial and fungal classes. The proposed techniques provide good results on determining the varied size of disease regions of the plant using feature distribution and its feature characteristics using particle swarm optimization.

- *Algorithm 1: Convolution Neural Network*

Input: Optimal feature set F={f1, f2,…fN}
Output: Target Disease Class C={c1,c2..cN}
Process*:* For (M= Feature Vector[s], M++ , M<threshold)
Calculate Convolution () for segmented feature as 3*3 layers for Hic =P(wi*x)

Feature Map = Kernel (H_{ic})
Calculate High discriminant feature using Max pooling ()
Reduce (Feature Map)
High Discriminant Feature Map HFm = Reduced Feature(F)
Fully connected layer for disease classification Dropout layer ()
Eliminate the over fitting feature to classes Activation Function
Regulated Feature R = ReLu(HFm)
Softmax function C = random forest (R)
Loss function ()
Apply stochastic gradient descent (C(f))
W(f)= Weighted matrix of the Classes with Over fitting features
Adam classifier(w(f))
Class Labels = disease region of the plants and Disease type of Plant- Viral / Fungal

IV. EXPERIMENTAL RESULTS

Experimental analysis of proposed model employed to e commerce application incorporates the fertilizer recommendation and crop recommendation using python flask. Different design layouts are constructed to detect the disease of the plant , fertilizer recommendation and crop recommendation on acquiring soil parameters and environment parameters.

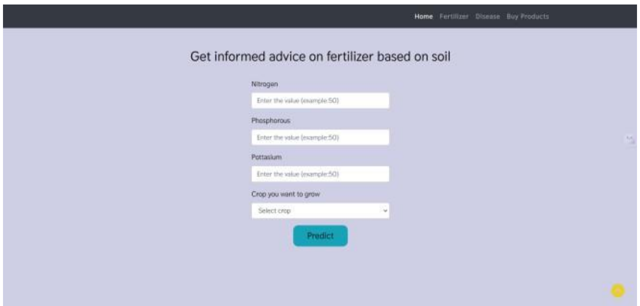

Fig. 3. Page Layout of the Fertilizer Suggestion

The layout represented in figure 3 provides fertilizer suggestions on obtaining the soil parameter and environment parameter.

Experimental analysis of the proposed architecture is carried out using village plant dataset. Proposed architecture classifies the plant disease into three discriminant classes. Those classes were considered disease classes of the plant images represented as test images. The proposed architecture is modeled in a python environment. Towards model validation, 10 fold validations were employed on utilizing the confusion matrix. Cross fold validation improves the model performance.

A. *Performance metrics*

With the use of accuracy and loss measurements, the suggested model's plant disease classification performance is evaluated. By using the plant's volumetric changes at various phases of growth, the proposed model is able to identify plant diseases.

- *Model Accuracy*

It is computed by emphasizing the differences between the categorized results in order to provide ground truth information for processing the segmented features.

Furthermore, it can be computed in relation to the results of confusion matrices, including true positive, true negative, false positive, and false negative classification output values. It is mentioned as

$$\text{Accuracy} = \frac{2TP}{2TP+FP+FN}$$

- *Model Loss*

It estimates the percentage of True Positive which is considered as disease features which is correctly classified among the other features. It is provided as

$$\text{Loss} = \frac{TP}{TP+FN}$$

The analysis on different test plant images on real time images or from dataset computes the disease type. Further performance evaluation values of the proposed model and existing model are tabulated in table 2.

TABLE II. PERFORMANCE EVALUATION

Technique	***Accuracy***	***Loss***
ResNet Proposed- Training	0.968	0.6
ResNet- Proposed- Validation	0.978	0.5

It has been proved that the proposed architecture classifies the plant disease accurately and it is capable of classifying the various plant growth stages. Further performance of the proposed Convolution neural Network represents good performance results. The proposed architectures provides enhanced performance without any additional parameters [20].

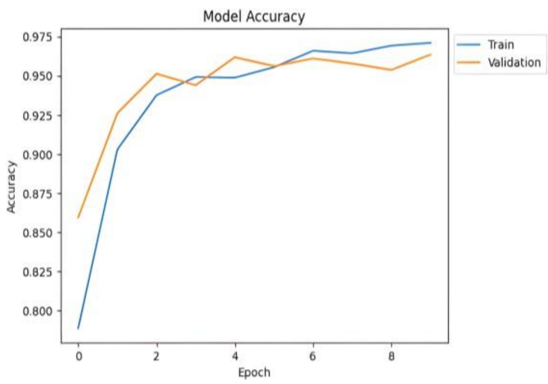

Fig. 4. Performance analysis of Model accuracy against training and testing data

Model accuracy evaluated to disease classes is represented in figure 4. Regarding the loss of the model, the model reduces it on hyperparameter tuning. Figure 5 illustrates the outcomes of the loss to the obtained classes .

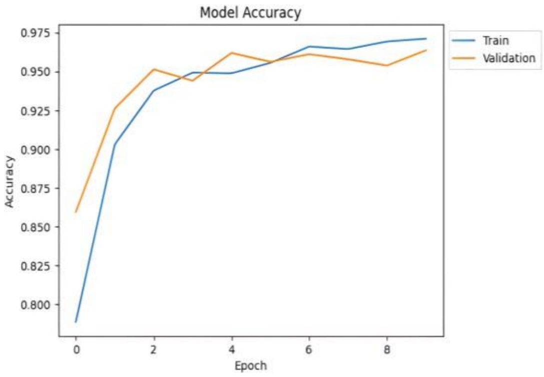

Fig. 5. Performance analysis of Model loss against training and testing data

The efficiency of the model is defined on analyzing the performance on trained models as it is highly capable of adapting to feature changes. For enhancing the classification accuracy, proposed architecture is optimized by changing the depth and kernel size in the hyper parameter tuning.

V. CONCLUSION

The design and implementation of a novel ensemble deep learning framework using random forests for fertilizer recommendation, support vector machines for disease classification and prediction, and convolutional neural networks for disease prediction are presented in this research. With improved recognition accuracy, the proposed model can identify the type of plant disease. Additionally, it produces positive outcomes for datasets containing volumetric variations, imaging artifacts, anatomical variability, shifting contrast characteristics, and inadequate image registration.

REFERENCES

[1] C. Dechant, T. Wiesner-Hanks, S. Chen, E. L. Stewart, J. Yosinski, M. A. Gore, R. J. Nelson, and H. Lipson, "Automated identification of northern leaf blight-infected maize plants from field imagery using deep learning," Phytopathology, vol. 107, no. 11, pp. 1426-1432, Nov. 2017.

[2] S.P. Mohanty, D. P. Hughes and M. Salathé, "Using Deep Learning for Image-Based Plant Disease Detection," Frontiers in Plant Science, vol. 7, pp. 1419, Sep. 2016.

[3] K. Song, X. Y. Sun, J. W. Ji, "Corn leaf disease recognition based on support vector machine method," Transactions of the Chinese Society of Agricultural Engineering, vol. 23, no. 1, pp. 155-157, Jan. 2007.

[4] S. Zhang, X. Wu, Z. You, and L. Zhang, ``Leaf image based cucumber disease recognition using sparse representation classification," Computer Electronics of Agriculture., vol. 134, pp. 131-141, Mar. 2017.

[5] C. Dechant, T. Wiesner-Hanks, S. Chen, E. L. Stewart, J. Yosinski, M. A. Gore, R. J. Nelson, and H. Lipson, "Automated identification of northern leaf blight-infected maize plants from field imagery using deep learning," Phytopathology, vol. 107, no. 11, pp. 1426-1432, Nov. 2017.

[6] D. Oppenheim, G. Shani, "Potato Disease Classification Using Convolution Neural Networks," in the 11th European Conference on Precision Agriculture (ECPA 2017), 2017, pp. 244-249.

[7] A. Krizhevsky, I. Sutskever, G. Hinton, "ImageNet classification with deep convolutional neural networks," in International Conference on Neural Information Processing Systems (NIPS'12), Lake Tahoe, Nevada USA, 2012, pp. 1097-1105.

[8] S. Chetlur, C. Woolley, P. Vandermersch, J. Cohen, J. Tran, B. Catanzaro, E. Shelhamer, "cuDNN: Efficient Primitives for Deep Learning," in Deep Learning and Representation Learning Workshop (NIPS2014), 2014, pp. 1-9.

[9] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, "Dropout: a simple way to prevent neural networks

from overfitting," Journal of Machine Learning Research, vol. 15, no. 1, pp. 1929-1958, Jun. 2014.

[10] D. K. Berger, ``Grey leaf spot disease of maize and food security research," South Afr. J. Botany, vol. 100, no. 109, p. 327, Jan. 2017.

[11] S. Ren, K. He, R. Girshick, and J. Sun, ``Faster R-CNN: Towards realtime object detection with region proposal networks," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137-1149, Jun. 2017.

[12] P. Xiao, G. K. Venayagamoorthy, and K. A. Corzine, "Combined training of recurrent neural networks with particle swarm optimization and backpropagation algorithms for impedance identification," in Proceedings of the IEEE Swarm Intelligence Symposium (SIS '07), pp. 9–15, April 2007

[13] K. R. Aravind, P. Raja, K. V. Mukesh, R. Anirudh, R. Ashwin, and C. Szczepanski, ``Disease classification in maize crop using bag of features and multiclass support vector machine," in Proc. 2nd Int. Conf. Inventive Syst. Control (ICISC), Jan. 2018, pp. 1191-1196