A Hybrid Model for the Classification of Sunflower Diseases Using Deep Learning

Akash Sirohi
Research Scholar
Lovely Professional University
Punjab
akashsirohi7@gmail.com

Dr. Arun Malik
Department of Computer Science
Lovely Professional University
Punjab
arun.17442@lpu.co.in

2021 2nd International Conference on Intelligent Engineering and Management (ICIEM) | 978-1-6654-1450-0/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICIEM51511.2021.9445342

***Abstract* - Prediction and Recognition of plant disease in the early stage is one of the most essential needs to increase agriculture, which plays an important role in our country's economy and helps to feed a large population. And with the help of earlier detection, we can save the plants and avoid losses. Deep learning techniques are used widely to classify or predict diseases by using images. This paper proposed a hybrid model of deep learning to classify the sunflower diseases, i.e. Alternaria leaf blight, Downy mildew, Phoma blight, and Verticillium wilt. To make a hybrid model I used the stacking ensemble learning technique and combine two models i.e. VGG-16 and MobileNet, We also make our own dataset with Google images, and our proposed model gave 89.2% accuracy on our dataset, which is better than the other models.**

Keywords: Deep learning, Models, Diseases, Agriculture, Hybrid, Stacking ensemble.

I. INTRODUCTION

Sunflower originates in America Mexico in 2100 BCE and it is also known as Helianthus. Sunflower is the species which planted mostly in tropical countries and it grows up to 3 to 4 m, Genetic makeup and background are the main factors in the growth of sunflower and according to the Journal Of Environmental Management, nitrogen-based fertilizer gives various impact on a sunflower. Due to the diseases, production of sunflower can be reduced which affect many things because sunflower and its seed have several uses in the term of nutrients and medicines .i.e. it can be used as a food due to nutrients present in its seed and leaves, it also used to make natural dyes, it's root have the ability to soak that's why it is used to soak radioactive substances. Sunflower is the good source of vitamins and it is also used in the treatment like Malaria, Arthritis, Gastroenteritis, Chest Pain, Respiratory tract. It can also cure the bites of insects, snakes, and spiders. Sunflower leaves also used to cure bladder disorders due to their diuretic properties. As Sunflower has many uses so it's become important to save the sunflower by earlier detection of diseases with the help of computer vision.

Previously traditional method was used to classify or recognize the diseases which took more time and that was also costly. Those methods have high chances of error because all the work did by manually. So to overcome this situation and try to get high accuracy with less amount of time, traditional machine learning algorithms come into existence. Traditional machine learning algorithms better than manual work but their result were not satisfactory means they provided results with less accuracy. That's why Deep learning comes into existence as an image classifier.

Deep learning is the subset of Machine learning which is used widely for the detection and recognition of disease with the help of images. In Deep learning technique there is no need to design the feature extraction function and classifier manually because it study the hierarchical feature automatically. Deep learning methods perform well and better than machine learning as an image classifier and the best one is CNN. CNN is considered the best and effective image classifier with small and large size dataset and CNN is also considered the basic architecture of all deep learning models. Basically in the selection of deep learning models, the dataset plays an important role i.e. if there is a small dataset then pre-trained models are best suited, and if there are multi-classification with a balanced dataset then the deep learning method with cross-entropy loss function performed well. Deep learning is used in all most every field or we can say everything shifting towards computer or automation. Deep learning areas are: Fraud detection, Natural language processing, Visual recognition, Healthcare, game playing, Image classification and etc.

As sunflower have many benefits so in order to save the sunflower plant we proposed a hybrid model by using stacking ensemble learning technique on Vgg-16 and MobileNet. In this work we classify 4 diseases of sunflower i.e. Alternaria leaf spot, Downy Mildew, Phoma Blight and Verticillium wilt, and one healthy leaf.


Fig. 1.1: Alternaria Lef Spot


Fig.1.2: Downy Mildew


Fig1.3:VerticilliumWilt


Fig1.4:PhomaBlight

This paper further divided into different sections i.e.(2) Related Works, (3) Material & Method, (4) Result & Analysis and (5) Conclusion.

978-1-6654-1450-0/21/$31.00 ©2021 IEEE

II. RELATED WORK

S. Ramesh & D. Vydeki in [1] classify and recognise the paddy leaf diseases with the help of optimized deep NN with jaya Algorithm. They collected their own dataset of rice leaf images from the farm field. In the end, they concluded that their proposed method gives the accuracy of 98.9%, 95.78%, 92%, 94% and 90.57% for blast affected, bacterial blight, sheath rot, brown spot and normal leaf respectively.

Miaomiao Ji et al. in [2] Classified the diseases of common-grape leaf by using their proposed method i.e. United Model (combinations of multiple CNN) with plant-village dataset and they concluded that their proposed united model gives the accuracy of 99.17% for validation and 98.57% for test.

Utpal Barman et al. in [3] proposed a method based on CNN for the detection of potato disease, they used both augmented and non-augmented images to train their model separately and concluded that their proposed method gives the validation accuracy of 96.98% for non-augmented and 96.75% for the augmented dataset.

Surampalli Ashok et al. in [4] proposed a method, to detect the tomato plant leaf disease with the help of image processing techniques i.e. segmentation, clustering, and open-source algorithms and they concluded that their proposed method gives an accuracy of 98.12%.

Vinod Kumar et al. in [5] proposed a method to classify the plant leaf diseases by using Residual network i.e. ResNet34 with a dataset of 15200 images collected from the palnt-village database and they concluded that their model gave 99.40% accuracy for test set.

S. Gayathri et al. in [6] proposed a method to classified the tea plant diseases with the help of LeNet and the leaf images collected from the plant village database and they concluded that their proposed method gives an accuracy of 90.23%.

Yong Zhong & Ming Zhao in [7] proposed a method for the classification of the apple leaf diseases with the help of the DenseNet-121 model and they concluded that their proposed model gives an accuracy of 93.30%.

Jose G.M. Esgario et al. in [8] proposed an effective method with the help of the ResNet50 model, to detect and to estimate the coffee leaves stress severity, caused by biotic agents. They used the dataset form the plant-village database and they concluded that their proposed method gave 95.24% accuracy and 86.51% accuracy of biotic stress and severity estimation respectively.

Aditya Karlekar & Ayan Seal in [9] proposed a method to classify the soybean leaf diseases with the dataset "Image Database of Plant Disease Symptoms". Their model consists of two modules one for feature extraction and the other for classification by CNN and in the end, they concluded that their model gives an accuracy of 98.14%.

Junde Chen et al. in [10] proposed an method to classify the plant leaf diseases by using pre-trained models i.e. VGGNet and Inception with their own dataset and in the end, they concluded that their model gives an accuracy of 91.83%.

Abdul Waheed et al. in [11] proposed a method to classify and recognize the corn leaf diseases with the help of an optimized DenseNet model and in the end, they concluded that their model gives an accuracy of 98.06%, which is better than other models.

Lawrence C. Ngugi et al. in [12] proposed a method based on a mobile application for the segmentation of tomato leaf with the help of the CNN model. They used their own dataset of 1408 images and they concluded that their proposed model gave an accuracy of 96%.

M.Nagaraju & Priyanka Chawla in [13] reviewed some existing deep learning models, sources of data acquisition, and the image processing techniques used for classification and prediction purposes.

Miaomiao ji et al. in [14] classified the plant leaf diseases and do severity estimation with the help of CNN i.e. BR-CNN which was based upon ResNet50. They concluded that their model gives the highest accuracy of 86.70%.

M.P.Vaishnave et al. in [15] proposed a DCNN method to classify the groundnut diseases that work automatically without any human supervision. They used the stochastic gradient descent momentum for the training of the dataset which was collected from the plant-village database and in the end, they concluded that their proposed method gave an accuracy of 99.28%.

Sinan Uguz & Nese Uysal in [16] classify the diseases of olive leaf by using CNN and they studied and compared their model with VGG16 and VGG19. In the end, they concluded that their model gives an accuracy of 95% with augmented image and an accuracy of 88% without augmentation.

Maheswari prabhakar et al. in [17] proposed a method to detect the disease of tomato leaf and to estimate the severity level. They used the ResNet101 model with the dataset collected from plant-village and compared this model with VGG 16, VGG 19, GoogLe-Net, Alex-Net, and ResNet 50 and concluded that ResNet 101 gives the highest accuracy of 94.6%.

Sachin B. Jadhav et al. in [18] proposed an efficient method for the classification of soybean diseases with the help of pre-trained networks i.e. AlexNet and GoogLeNet. They used the dataset of 1199 images for the training purpose and in the end,

they concluded that their models give an accuracy of 98.75% and 96.25% for AlexNet and GoogLeNet.

REFERENCE NO.	AUTHOR	CROP	TECHNIQUE	ACCURACY
1	Ramesh, S. & D. Vydeki.	Rice	Optimized DCNN + Jaya Algorithm	98.9%,95.78%,92%,94% and 90.57%
2	Miaomiao ji. et al.	Grape	Multi-CNN	98.57%
3	Utpal Barman et al.	Potato	CNN	96.98% & 96.75%
4	Surampalli Ashok et al.	Tomato	Segmentation & Clustering	98.12%
5	Vinod Kumar et al.	Plants	ResNet-34	99.4%
6	S. Gayathri et al.	Tea	LeNet	90.23%
7	Yong Zhong & Ming Zhao	Apple	DenseNet-121	93.30%
8	Jose G.M. Esgario et al.	Coffee	ResNet50	95.24% and 86.51%
9	Aditya karlekar&Ayan Seal	Soybean	CNN	98.14%
10	Junde Chen et al.	Plant	VGGNet+Inception	91.83%
11	Abdul Waheed et al.	Corn	DenseNet	98.06%
12	Lawrence C. Ngugi et al.	Tomato	CNN	96%
14	Miaomiao Ji. Et al.	Plant	ResNet-50	86.70%
15	M.P.Vaishnnave et al.	Groundnut	DCNN	99.28%
16	Sinan Uguz & Nese Uysal	Olive	CNN	95%
17	Maheswari Prabhakar et al.	Tomato	ResNet-101	94.6%
18	Sachin B. jadhav et al.	Soybean	AlexNet & GoogLeNet	98.75% & 96.25%

III. PROPOSED MODEL

The objective of this paper is to proposed a hybrid model of deep learning that can classify the sunflower diseases using images with the help of ensemble learning techniques. The process of model give in fig: 3.1

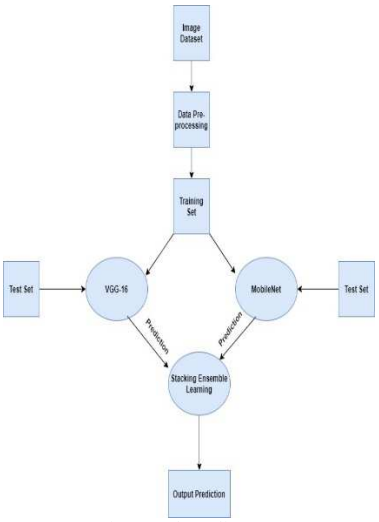

Fig 3.1 Flow diagram

A. Algorithm

1. Read images with Cv2 library and convert each image into array with keras function.
2. Make label list with LabelBinarizer function
   image_labels= LabelBinarizer().fit_transform(label_list)
3. convert image into grayscale:
   np_image_list=np.array(image_list,dtype=np.float16)/225.0
4. Now divide dataset into training and testing set as:
   x_train,x_test,y_train,y_test=train_test_split()
5. Now use ImageDataGenerator for the Data augmentation as:
   aug=ImageDataGenerator() aug.fit(x_train)
6. Then implement Vgg-16 and MobileNet from keras library
7. Now start training of both models using :
   model[i].fit_generator()
8. Apply stacking ensemble learning technique to create stacked model.
9. Now combine predicted values of both model in an array to form input of stacked model.
10. Now train stacked model with predicted value and testing set.
11. Now get predict value of stacked model in yhat and then evaluate accuracy as:
    acc = accuracy_score(testy, yhat)

B. Working

The first step of this process is Data pre-processing in which, it convert the images of size 224*224*3 into array then make a list of labels of images with the help of LabelBinarizer function and then with the help of ImageDataGenerator augment the images. After data pre-processing it divide the

dataset into training and testing set with the help of train_test_split (), in the ratio of 80:20 and with the help of random_state parameter of train_test_split () it select every 3 images for the splitting. After dividing the dataset, now this time for training the models(Vgg16 & MobileNet) with the help of model.fit_generator() with 25 epochs. After completion of training, combine the predict values of both models into an array with the help of numpy library and model. Predict function of keras. Now train the stacked model which is formed with the help of Logistic Regression model of machine learning with combined predict values(stacked dataset) and after the training store the predict values of stacked model with stacked dataset and at the end with that predict value evaluate the accuracy of stacked model with stacked dataset and testing set.

- *Parameters we used*

```
ImageDataGenerator(rotation_range=25,
width_shift_range=0.1,
height_shift_range=0.1,
shear_range=0.2,zoom_range=0.2,
horizontal_flip=True, fill_mode="nearest")

model.fit_generator (
ImageDataGenerator.flow(x_train, y_train, batch_size=32)
validation_data=(x_test,y_test),
steps_per_epoch=len(x_train)//BS, epochs=25,
callbacks=[ReduceLROnPlateau(monitor='loss',
patience=3, factor=0.1)],verbose=1)
```

IV. RESULT & DISCUSSION

Figure 4.1 shows the comparison between our proposed hybrid model and some existing deep learning models i.e. Vgg16, MobileNet, AlexNet, InceptionV3, and DenseNet121, in which our proposed hybrid model gives the best accuracy among them i.e. 89.2%. We used the stacking ensemble technique to combine or make an optimal hybrid model of Vgg16 and MobileNet with the properties which are given in table 4.1.

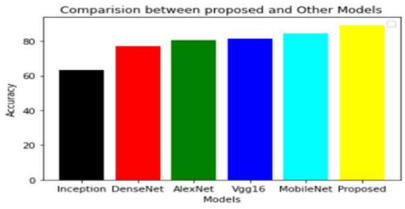

Fig. 4.1

PROPERTY	VALUE
Window	10
Processor	Intel i3, 6th generation
RAM	4gb
Language	Python
Software	Jupyter Notebook

Table 4.1: System Properties

Figure 4.2, shows the comparison between a proposed hybrid model with other ensemble technique i.e. Averaging model which give the accuracy of 83% which is less than our model i.e. 89.2% that's why we select stacking ensemble technique instead of Averaging model technique of ensemble learning.

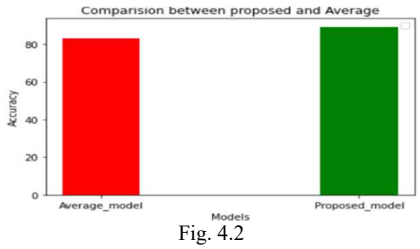

Fig. 4.2

V. CONCLUSION

In this work, we proposed an optimal hybrid model to classify the diseases of sunflower leaf with the help of the stacking ensemble learning technique that combine Vgg16 and MoileNet models. In This work we used this model to classify 5 classes of the sunflower which consist of 4 diseases i.e. Alternaria leaf spot, Downy Mildew, Phoma Blight, and Verticillium wilt, and 1 is healthy leaf class. We tried to apply two techniques of ensemble learning i.e. stacking and weighted average but stacking provide the better result then weighted average as shown in fig(4.2) we choose stacking technique.The strength of our model is Vgg-16 previous weight of imagenet which helps to extract the features more accurately and depth of MobileNet.The result and analysis shows that our method gave an accuracy of 89.2% which is better than the other deep learning models i.e. AlexNet, Inception_V3, DenseNet-121, Vgg-16, and MobileNet on our own dataset.

REFERENCES

[1] Ramesh, S., and D. Vydeki. "Recognition and classification of paddy leaf diseases using Optimized Deep Neural network with Jaya algorithm." *Information processing in agriculture* 7.2 (2020): 249-260.

[2] Ji, Miaomiao, Lei Zhang, and Qiufeng Wu. "Automatic grape leaf diseases identification via UnitedModel based on multiple convolutional neural networks." *Information Processing in Agriculture* 7.3 (2020): 418-426.

[3] Barman, Utpal, et al. "Comparative Assessment of Deep Learning to Detect the Leaf Diseases of Potato based on Data Augmentation." *2020 International Conference on Computational Performance Evaluation (ComPE)*. IEEE, 2020.

[4] Ashok, Surampalli, et al. "Tomato Leaf Disease Detection Using Deep Learning Techniques." *2020 5th International Conference on Communication and Electronics Systems (ICCES)*. IEEE, 2020.

[5] Kumar, Vinod, Hritik Arora, and Jatin Sisodia. "ResNet-based approach for Detection and Classification of Plant Leaf Diseases." *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*. IEEE, 2020.

[6] Gayathri, S., et al. "Image Analysis and Detection of Tea Leaf Disease using Deep Learning." *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*. IEEE, 2020.

[7] Zhong, Yong, and Ming Zhao. "Research on deep learning in apple leaf disease recognition." *Computers and Electronics in Agriculture* 168 (2020): 105146.

[8] Esgario, José GM, Renato A. Krohling, and José A. Ventura. "Deep learning for classification and severity estimation of coffee leaf biotic stress." *Computers and Electronics in Agriculture* 169 (2020): 105162.

[9] Karlekar, Aditya, and Ayan Seal. "SoyNet: Soybean leaf diseases classification." *Computers and Electronics in Agriculture* 172 (2020): 105342.

[10] Chen, Junde, et al. "Using deep transfer learning for image-based plant disease identification." *Computers and Electronics in Agriculture* 173 (2020): 105393.

[11] Waheed, Abdul, et al. "An optimized dense convolutional neural network model for disease recognition and classification in corn leaf." *Computers and Electronics in Agriculture* 175 (2020): 105456.

[12] Ngugi, Lawrence C., Moataz Abelwahab, and Mohammed Abo-Zahhad. "Tomato leaf segmentation algorithms for mobile phone applications using deep learning." *Computers and Electronics in Agriculture* 178 (2020): 105788.

[13]Nagaraju, M., and Priyanka Chawla. "Systematic review of deep learning techniques in plant disease detection." *International Journal of System Assurance Engineering and Management* 11 (2020): 547-560.

[14] Ji, Miaomiao, et al. "Multi-label learning for crop leaf diseases recognition and severity estimation based on convolutional neural networks." *Soft Computing* 24.20 (2020): 15327-15340.

[15] Vaishnnave, M. P., K. Suganya Devi, and P. Ganeshkumar. "Automatic method for classification of groundnut diseases using deep convolutional neural network." *Soft Computing* 24.21 (2020): 16347-16360.

[16] Uğuz, Sinan, and Nese Uysal. "Classification of olive leaf diseases using deep convolutional neural networks." *Neural Computing and Applications* (2020): 1-17.

[17] Prabhakar, Maheswari, Raja Purushothaman, and Durga Prasad Awasthi. "Deep learning based assessment of disease severity for early blight in tomato crop." *Multimedia Tools and Applications* 79.39 (2020): 28773-28784.

[18] Jadhav, Sachin B., Vishwanath R. Udupi, and Sanjay B. Patil. "Identification of plant diseases using convolutional neural networks." *International Journal of Information Technology* (2020): 1-10.