TENCON 2023 - 2023 IEEE Region 10 Conference (TENCON)
31 Oct - 3 Nov 2023. Chiang Mai, Thailand

Coffee Leaf Disease and Severity Prediction Using Deep Learning

Harshit Kumar Chaubey[1], Siri Arelli[2], Aryan Raj[3] and Nenavath Srinivas Naik[4]

Abstract— **Coffee production is a vital industry in many countries, but diseases affecting coffee leaves can lead to significant losses for farmers. To mitigate these losses, timely disease detection and accurate assessment of disease severity are crucial. This work proposes a deep learning approach for classifying coffee leaf diseases based on their severity levels. The proposed methodology involves several steps. Initially, U^2Net removes the background from the coffee leaf images. Subsequently, the background-removed images are converted into BGR format to identify the diseased regions. DeepLabV3 is then trained to extract and mark the diseased portions of the leaves in red. Using these annotated images, various convolutional neural network (CNN) models, including VGG-16, InceptionV3, and MobileNetV2, were trained to classify the diseases based on their severity levels. These models are carefully modified and fine-tuned with hyperparameters to achieve the best performance metrics. Upon model training, the modified MobileNetV2 model performs better than the other CNN models, achieving an impressive F1-score of 97.99%. This outcome highlights the effectiveness of this paper's approach in accurately classifying coffee leaf diseases according to their severity levels. The proposed methodology has significant implications for coffee farmers, enabling them to swiftly detect diseases and assess their severity, allowing for timely and appropriate actions. Furthermore, the findings indicate the superiority of the modified MobileNetV2 model in achieving high accuracy in disease severity classification. This research contributes to advancing deep learning techniques for agricultural applications, offering practical solutions for disease management in the coffee industry.**

I. Introduction

Nowadays, AI techniques and smartphone applications have rapidly increased and become famous in automated crop production. These intelligent technologies have been used to address the riddle found on coffee crops originated by pests and pathogens. Machine learning algorithms, like convolutional neural networks (CNN), have been proven helpful in image recognition tasks that can automatically locate and categorize coffee diseases and pests. Moreover, a few obstacles are associated with the complexity of the image background and image capture circumstances, for example, variation in brightness, angle, and other appliances used. These aspects can lead to misinterpretations in observing and categorizing coffee diseases and pests.

Further, detecting lesions is vital for accurate finding and classification, and combining segmentation and detection with a classification task can generate more precise results. Regardless of these challenges, only a few studies in the previous work used deep learning algorithms to find, segment, and categorize diseases in different crops, including wheat, cucumber, and tomato leaves. In this research, deep learning algorithms have been applied to evolve a model that can find and classify diseases in crop images with a high degree of correctness.

Generally, AI techniques are assuring key for auto-mated crop protection. These technologies have the aptitude to boost crop yield and minimize losses drastically. Moreover, there is still a barrier to overcome, primarily related to the complexity of the image background and capture circumstances. To deal with these difficulties, additional research is required to cultivate more accurate and reliable deep-learning algorithms to find and categorize crop diseases and pests under diverse circumstances.

The main objectives of this paper are:

- To train a classification model that will distinguish between healthy and rusted leaves on different severity levels requires training and testing a modified CNN model on a RoCoLe dataset, i.e., robusta coffee leaf images dataset, which is the on-field dataset.
- To ensure good fitting of the classification model by tuning various hyperparameters such as learning rate and batch size and evaluating the model's performance on some metrics such as accuracy, precision, recall, and F1-score.
- To compare various CNN models to check which is the best-performing architecture for this problem and dataset, which involves evaluating model performance on the test dataset.
- To preprocess field dataset so that model can ignore noise present in images and focus only on target leaf.

This paper is organized as follows: Section II provides a brief of the current research, including the techniques used, the outcome derived, and the conclusion drawn. Section III introduces the dataset used in this research, with the deep learning network architectures employed and the proposed model. Section IV presents the research's outcome and offers a detailed discussion. Finally, the study concludes with Section V.

II. Literature survey

In this section, other research on the given problem statement is discussed. Very few of those researchers speak about the severity prediction or classification of crop diseases.

Earlier, some research studies were conducted where they developed ML models for crop disease detection. In study [1], they used deep learning models to classify diseases in cucumber crops. In this study, they trained their model on a large dataset containing over 5000 images of cucumber leaves with various diseases. The model achieved an accuracy of 95% in classifying the diseases. Another study [2] talks about classifying diseases present in wheat crops where they

[1]Computer Science and Engineering, IIIT Naya Raipur, Raipur, India harshit21100@iiitnr.edu.in
[2]Data Science and Artificial Intelligence, IIIT Naya Raipur, Raipur, India siri21102@iiitnr.edu.in
[3]Computer Science and Engineering, IIIT Naya Raipur, Raipur, India aryanr21100@iiitnr.edu.in
[4]Computer Science and Engineering, IIITDM Kurnool, Kurnool, India srinu@iiitk.ac.in

979-8-3503-0219-6/23/$31.00 ©2023 IEEE

have trained their deep learning model on over 10000 images of wheat leaves containing different types of diseases. Their CNN model was able to detect disease with an accuracy of 90%. Then some studies discuss visualizing diseases in coffee [3], which used techniques like Score-CAM, Grade-CAM, and Grad-CAM++ for visualization. These techniques can help identify misclassification and propose a guided approach for classifying coffee disease. They have made a guided approach that achieved an accuracy of 98%, which was way higher than the naive approach on the Robusta coffee leaf dataset. This study is the base of this research. Some papers [4] introduced an integrated framework to automate the detection and classification of coffee leaf diseases from in-field smartphone images. The framework combines instance and semantic segmentation to achieve promising disease classification results. This can be used for developing mobile apps for automating the task of disease classification. Some paper like [5] discusses the development of a multi-task system for identifying and estimating diseases on coffee leaves. It integrates convolution and pooling layers in CNN models to classify and assess disease severity on coffee leaves. In the study [6], they compared results for disease classification on field collected and laboratory datasets. They have used different CNN architectures where results were excellent on laboratory datasets but failed on on-field datasets. This implies the technical issue generally observed in applying deep learning models on on-field datasets. Some studies like [7] explores how deep learning models can accurately estimate plant disease severity. They have trained CNN models on annotated apple black rot images from the plant village dataset. Where they found VGG16 model, which is trained on transfer learning, was performing best with an accuracy of 90.4% which gives promising potential to utilize deep learning for increasing the yield of the crops. Some papers like [8] developed a model based on instance segmentation and object detection for developing a deep learning model for classifying the disease present in segment apple trees in orchards. They trained their model on over 3000 images where apple trees were in various growth stages. They achieved an accuracy of over 90%.

A. *Limitations of Existing Work*

A few drawbacks from existing works are listed below :

- Most deep learning models are trained on laboratory datasets that need a more accurate representation of actual datasets. This may result in poor performance of models on on-field datasets.
- Most models consider single-leaf images, but in reality, multiple leaves exist.
- Most of the research focuses on the classification of diseases, but only some can estimate the severity of the disease in the leaf.

To tackle the limitations of previous studies, this work recommends a unique deep-learning semantic segmentation strategy for verifying plant diseases and approximating their severity. This technique aims to strengthen the accuracy and authenticity of managing plant disease.

B. *Issues discussed in this paper*

- This work utilizes the on-field robusta coffee leaf images dataset (RoCoLe) to make an app-ready DL model that farmers can use.
- This work is mainly focused on classifying diseases with different levels of severity present in the coffee leaf.

III. Proposed Work

We have implemented this work in basically three stages, i.e., data collection, data preprocessing, model training, and evaluation of the model.

A. ***Dataset Collection***

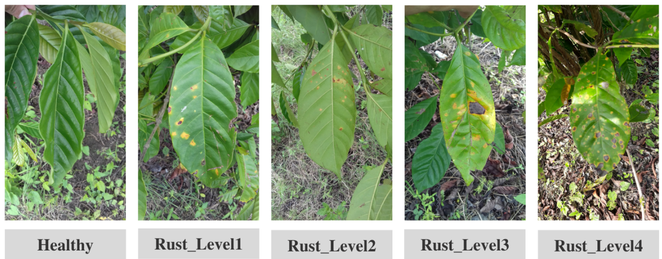

Fig. 1. RoCoLe dataset and its sample image of each class

This study utilizes the dataset from the Mendeley database, i.e., RoCoLe dataset [9] as shown in figure 1. It is a collection of 1393 high-quality images of Robusta coffee leaves, divided into five categories. These categories include healthy leaves and those affected by varying levels of rust damage from level 1 to level 4. These severity levels are classified based on the percentage of portion diseased as shown in table II. The dataset is helpful in this study for detecting and classifying coffee leaf diseases.

TABLE I
Sample Distribution of RoCoLe Dataset

Classes	Number of Images
healthy	791
rust_level_1	344
rust_level_2	166
rust_level_3	62
rust_level_4	30
Total Images	1393

TABLE II
Description about severity levels of Rust disease

Rust Levels	% of portion diseased
rust_level_1	1-5
rust_level_2	6-20
rust_level_3	21-50
rust_level_4	greater than 51

B. ***Data Preprocessing***

The dataset is first preprocessed to ensure a suitable and more meaningful format for further analysis or model training. This step involves background noise removal, color segmentation, and semantic segmentation to highlight the

979-8-3503-0219-6/23/$31.00 ©2023 IEEE

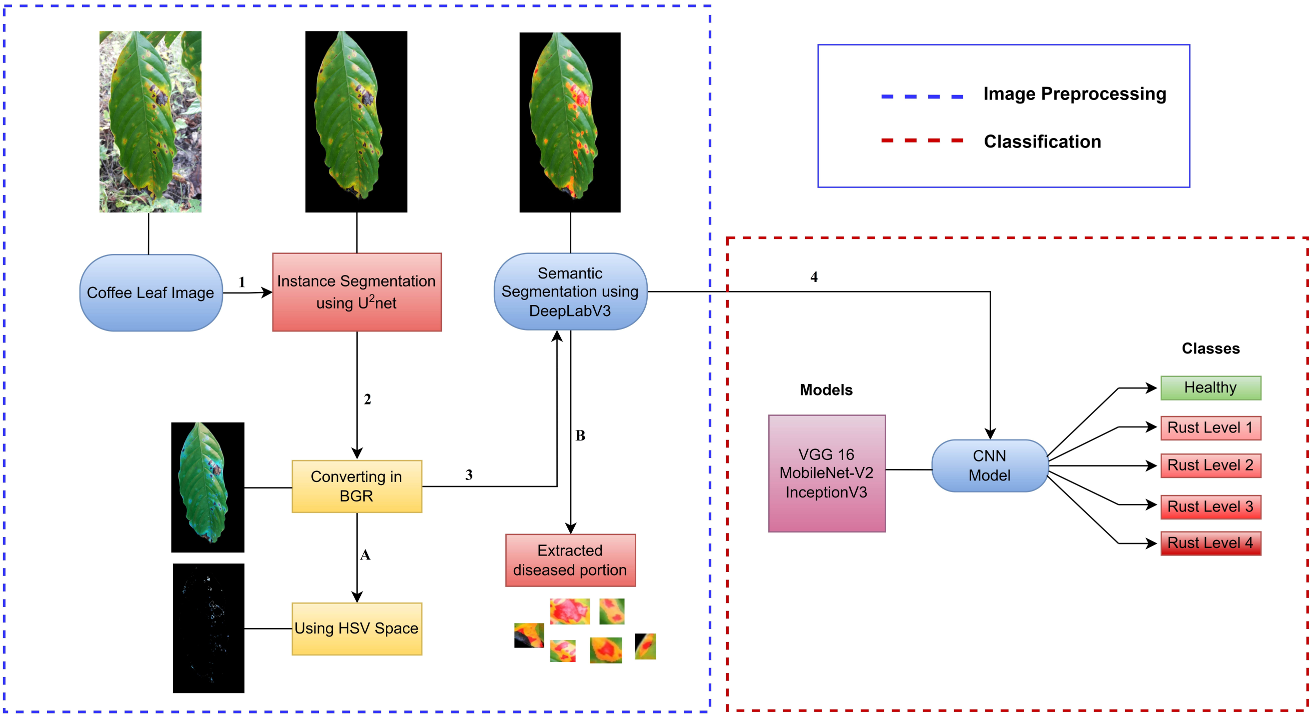

Fig. 2. Proposed methodology - 1,2,3,4 are steps involved in image preprocessing and model training; A, B are visualization of disease in HSV color space and cropped diseased portion

diseased portion, data balancing using oversampling, and data augmentation techniques.

U^2Net is a deep neural network architecture for image segmentation [10]. It extends the U-Net model and achieves state-of-the-art performance. The Bitwise AND operator created an output image with pixels in both input images. It helps in removing the background by combining the original image with the U^2Net-generated image.

Color segmentation is a process of separating pixels in an image based on their color. It is often used in computer vision applications such as object detection, tracking, and image processing. One way to perform color segmentation is to use the BGR color space. To segment diseased parts of a leaf image, a similar approach to the one used for leaf segmentation can be implemented. However, instead of segmenting the entire leaf, the focus can be on segmenting only the parts of the leaf that show signs of disease.

In OpenCV, the HSV color space can be easily converted from the RGB or BGR color spaces using the cvtColor() function. In the HSV color space, the 'Hue' component represents the color information and is described as an angle around a color wheel. The 'Saturation' component represents the purity or intensity of the color, while the 'Value' component represents the luminance or brightness of the color. After converting the image to HSV color space, color-based segmentation sets a color threshold on the Hue, Saturation, and Value components.

DeepLabV3 is a convolutional neural network architecture designed for semantic segmentation tasks, which is used to classify every pixel in an image into a specific category [11]. The architecture of DeepLabV3 is based on a modified ResNet backbone that extracts features from the input image.

The available dataset is imbalanced as the majority class contains 791 images while the minority classes include less than 700, respectively, as seen in table III. This imbalance in data can cause bias in the model during training, resulting in poor performance and inaccurate predictions. There was a need for data balancing techniques such as data augmentation and oversampling to address this issue. Data augmentation is used in machine learning to increase the size and diversity of a dataset by creating new variations of existing data. The idea is to generate recent examples by applying various transformations to the original data, such as flipping, rotating, cropping, zooming, or changing the brightness or contrast of images. The final balanced training dataset used for model training can be seen in table IV.

TABLE III
SAMPLE DISTRIBUTION OF TRAINING DATASET

Classes	Number of Images
healthy	474
rust_level_1	206
rust_level_2	99
rust_level_3	37
rust_level_4	18

TABLE IV
SAMPLE DISTRIBUTION OF TRAINING DATASET AFTER DATA BALANCING

Classes	Number of Images
healthy	360
rust_level_1	360
rust_level_2	360
rust_level_3	360
rust_level_4	360

C. Model Training

Convolutional Neural Networks (CNNs) have revolutionized computer vision by effectively analyzing and understanding images. They consist of convolutional layers that extract key features using filters, pooling layers that reduce dimensionality, batch normalization layers for stabilization, activation function layers for non-linearity, and fully connected layers for classification.

CNNs eliminate the need for manual feature engineering as they automatically learn essential features through training. By optimizing weights and biases using backpropagation, CNNs achieve remarkable performance in image analysis tasks. CNNs provide a robust framework for extracting meaningful information from images, making them invaluable in computer vision applications.

We have trained various models on the RoCoLe dataset and compared their outcomes and predictions. A few of the models have been described below:

1) VGG16: VGG-16 is a 16-layer deep CNN architecture with 13 convolutional layers and three fully linked layers [12]. The convolutional layers use 3x3 filters with a stride of 1 pixel, followed by ReLU activation and 2x2 max pooling layers. The final fully connected layers perform classification based on the extracted features. VGG-16's strength lies in its simplicity, depth, and effective use of small filters and max pooling to prevent overfitting.

2) InceptionV3: InceptionV3 is an advanced convolutional neural network architecture that enhances image classification performance [13]. It incorporates parallel convolutional blocks, each with its own convolutional and pooling layers, which are then concatenated for classification. Inception modules with different filter sizes enable multi-scale feature extraction, while 1x1 convolutional filters reduce model parameters. Batch normalization is employed for faster training and improved network stability.

3) MobileNetV2: MobileNetV2 is a deep convolutional neural network designed for mobile and embedded vision applications with limited computational resources and memory [14]. It achieves efficiency by utilizing depthwise separable convolutions and inverted residual blocks. Depthwise separable convolutions separate spatial and channel-wise operations, reducing computation and memory requirements. Inverted residual blocks use shortcut connections to bypass lightweight convolutional layers.

The MobileNetV2 architecture introduces linear bottleneck blocks, which combine 1x1 and 3x3 convolutions to reduce computations while preserving complex feature representation. This innovation enhances model accuracy while maintaining efficiency, making it suitable for resource-constrained devices.

IV. Experimental Analysis and Results

This section will discuss the performance metrics employed for the classification problem and the results obtained from the experiment. The results will provide insights into the model's performance and capability to capture relevant patterns and make accurate predictions for disease classification in different severity levels.

A. Performance Metrics Used

1) Confusion Matrix: The performance of a model is evaluated using a confusion matrix. The following terms are associated with the confusion matrix:

True Positive (TP): The disease is predicted in coffee leaf when it actually has a disease.

True Negative (TN): No disease is predicted in coffee leaf when it actually does not have a disease.

False Positive (FP): No disease is predicted in coffee leaf, but it has a disease.

False Negative (FN): The disease is predicted in coffee leaf when it actually does not have a disease.

2) Precision: Precision specifically focuses on the relevance of the model's predictions regarding the presence of coffee leaf disease.

$$\text{Precision } = \frac{(TP)}{(TP+FP)} \tag{1}$$

3) Recall: It refers to the proportion of correctly identified instances of the disease among all the instances of coffee leaves infected with it.

$$\text{Recall } = \frac{(TP)}{(TP+FN)} \tag{2}$$

4) Accuracy: It refers to how well a classification model correctly classifies coffee leaves as either infected or non-infected with the disease.

$$\text{Accuracy } = \frac{(TP+TN)}{(FP+FN+TN+TP)} \tag{3}$$

5) F1-score: F1-score is a useful metric to assess the overall performance of a classification model. A high F1 score indicates a balanced performance in correctly identifying infected leaves while minimizing misclassifications.

$$\text{F1-score } = \frac{2(Precision * Recall)}{(Precision + Recall)} \tag{4}$$

Hyperparameters of CNN model:

Table VI indicates the hyperparameters used for training the deep learning model. The optimizer and loss function used are Adam and cross-entropy, respectively. With a batch size of 30, the CNN model was trained for 100 epochs. The learning rate was set at 0.05, and no weight decay was considered during training. These hyperparameters were chosen to optimize the model's accuracy and reduce the risk of overfitting.

B. Results

After tuning hyperparameters, the modified CNN models were ready for training. After model training, the results from Table VII, it can be seen that MobileNetV2 outperformed the other two CNN models. MobileNetV2 was able to acheive an F1-score of 97.99%, whereas the other two models, VGG16 and InceptionV3 achieved an F1-score of 96.49% and 94.84%, respectively. The figures 4, 5, 6 represent training accuracy vs. testing accuracy for CNN models, i.e., VGG16, MobileNetV2, and InceptionV3, respectively. Then, figure 7 represents test accuracy across all three CNN models.

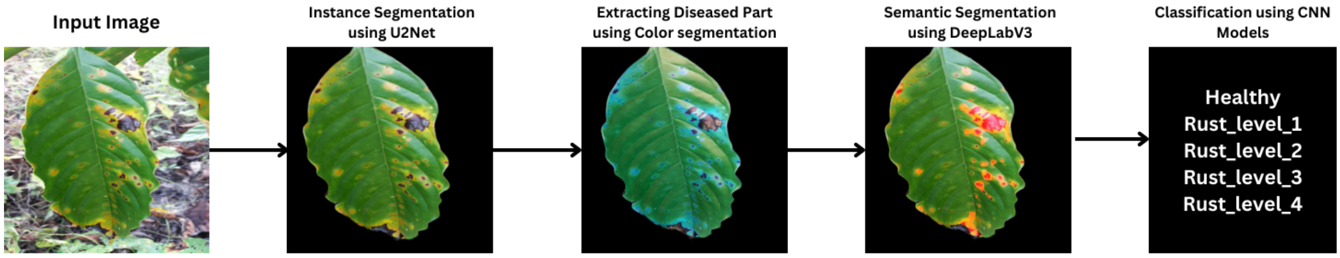

Fig. 3. Steps illustrated schematically in this paper

TABLE V

COMPARISION OF PREVIOUS METHODS AND PROPOSED WORK

Title	Dataset	CNN Model	Dataset Size	No. of classes	Accuracy
Coffee Disease Visualization and Classification [3]	RoCoLe (Coffee Leaf Dataset)	Modified ResNet	1560	2	98.0%
A deep learning approach combining instance and semantic segmentation to identify diseases and pests of coffee leaves from in-field images [4]	RoCoLe (Coffee Leaf Dataset)	PSPNet	1560	6	94.25%
Deep Learning for Classification and Severity Estimation of Coffee Leaf Biotic Stress [5]	BRACOL (Coffee Leaf Dataset)	ResNet50	1685	5	84.13%
Proposed Work	RoCoLe (Coffee Leaf Dataset)	Modified MobileNetV2	1393	5	98.82%

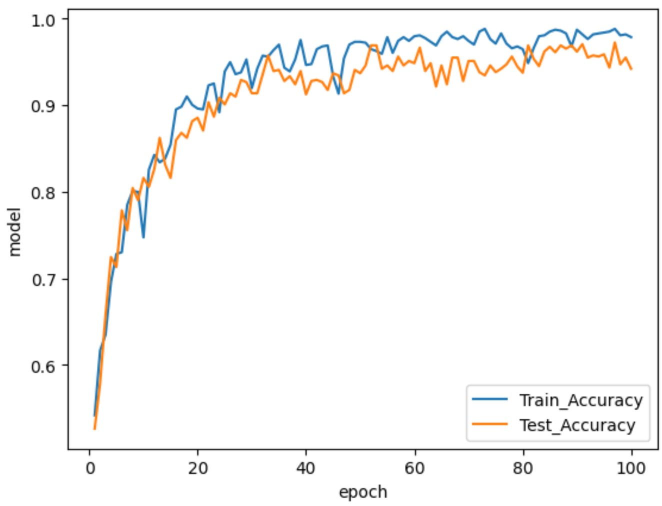

Fig. 4. Train vs. Test Accuracy for VGG16

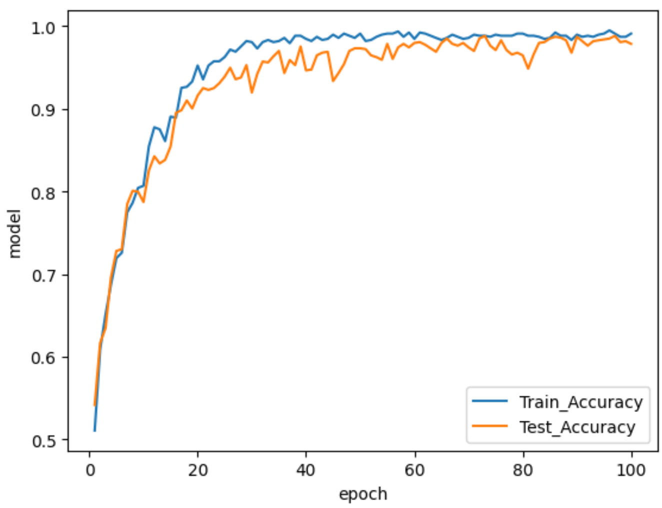

Fig. 5. Train vs. Test Accuracy for MobileNetV2

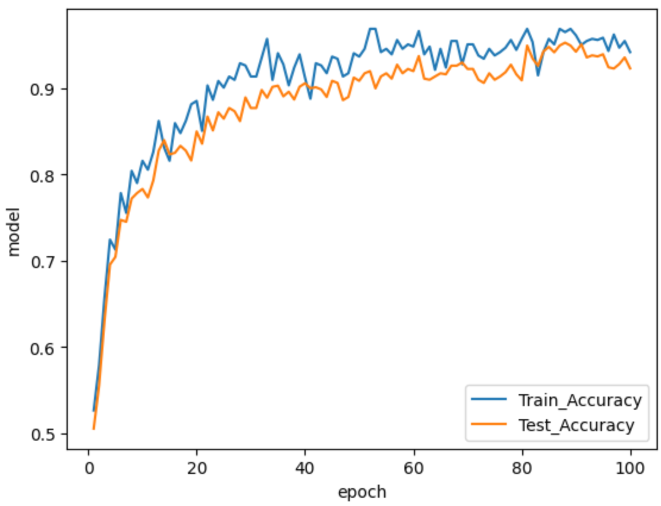

Fig. 6. Train vs. Test Accuracy for InceptionV3

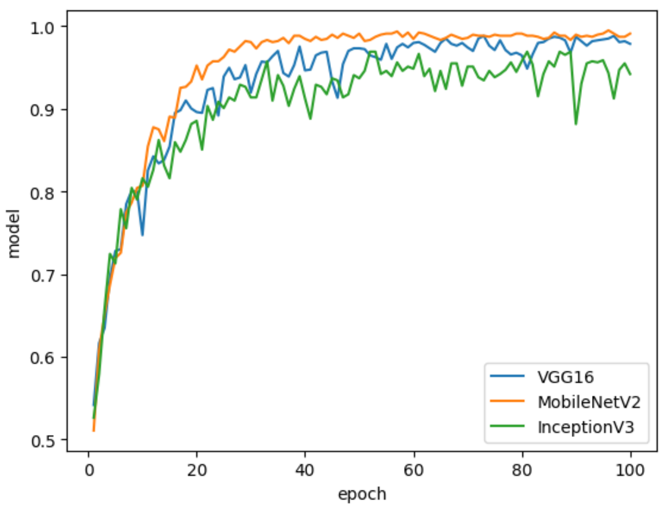

Fig. 7. Test accuracy comparison across model

TABLE VI

CNN TRAINING HYPERPARAMETERS

Hyperparameter	Value
Learning Rate	0.05
Loss Function	Crossentropy
Weight Decay	None
Epochs	100
Optimizer	Adam
Batch Size	30

TABLE VII

COMPARISON OF METRICS FOR MODELS

Model	Accuracy	Precision	Recall	F1 Score
VGG 16	97.25%	96.20%	96.80%	96.49%
MobileNetV2	**98.82%**	**97.81%**	**98.20%**	**97.99%**
InceptionV3	95.30%	94.60%	95.10%	94.84%

From the above results, an inference can be taken that this proposed methodology has the potential to detect diseases in coffee leaves at an early stage which will be helpful for the farmer to cut down their loss and increase the crop yield. It can also be observed in table V that this proposed method has better results than other work to classify diseases.

V. CONCLUSION AND FUTURE WORK

In conclusion, a deep learning model has been developed that can accurately classify coffee leaves as healthy or diseased while also predicting the severity level of the identified diseases. This involved several steps, including data preprocessing, feature extraction, model training, disease detection, and severity prediction. The proposed model exhibited impressive performance, achieving an F1 score of 97.99%. The successful development of this deep learning model and its potential integration into a user-friendly mobile application holds promise for disease detection and severity assessment in the coffee industry. It has the potential to enhance farmers' capabilities, reduce crop losses, and contribute to sustainable and efficient coffee production practices.

There are several avenues for future research and application of this work. One crucial direction is validating the model's performance on a more extensive and diverse on-field dataset to assess its accuracy and generalization across coffee leaves. Furthermore, integrating this model into a mobile application could provide immense value to farmers, allowing them to determine their crop condition in real-time conveniently. This integration would entail designing a user-friendly interface and ensuring compatibility with various mobile devices. By providing actionable insights and guidance, the application could assist farmers in making informed decisions and taking proactive measures to preserve the health and productivity of their coffee crops.

ACKNOWLEDGMENT

This research work has received funding from the Science and Engineering Research Board (SERB), Department of Science Technology (DST), Government of India under the Empowerment and Equity Opportunities for Excellence in Science (EMEQ) scheme - 2019, with File no. EEQ/2019/000439.

REFERENCES

[1] S. M. Omer, K. Z. Ghafoor, and S. K. Askar, 'An Intelligent System for Cucumber Leaf Disease Diagnosis Based on the Tuned Convolutional Neural Network Algorithm', Mobile Information Systems, vol. 2022, p. 8909121, Dec. 2022.

[2] M. Khalid, M. S. Sarfraz, U. Iqbal, M. U. Aftab, G. Niedbała, and H. T. Rauf, 'Real-Time Plant Health Detection Using Deep Convolutional Neural Networks', Agriculture, vol. 13, no. 2, p. 510, Feb. 2023.

[3] M. Yebasse, B. Shimelis, H. Warku, J. Ko, and K. J. Cheoi, 'Coffee Disease Visualization and Classification', Plants, vol. 10, no. 6, p. 1257, Jun. 2021.

[4] L. M. Tassis, J. E. Tozzi de Souza, and R. A. Krohling, 'A deep learning approach combining instance and semantic segmentation to identify diseases and pests of coffee leaves from in-field images', Computers and Electronics in Agriculture, vol. 186, p. 106191, 2021.

[5] J. G. M. Esgario, R. A. Krohling, and J. A. Ventura, 'Deep learning for classification and severity estimation of coffee leaf biotic stress', Computers and Electronics in Agriculture, vol. 169, p. 105162, 2020.

[6] G. Fenu and F. M. Malloci, 'Evaluating Impacts between Laboratory and Field-Collected Datasets for Plant Disease Classification', Agronomy, vol. 12, no. 10, p. 2359, Sep. 2022.

[7] G. Wang, Y. Sun, and J. Wang, 'Automatic Image-Based Plant Disease Severity Estimation Using Deep Learning', Computational Intelligence and Neuroscience, vol. 2017, p. 2917536, Jul. 2017.

[8] S. Sood, H. Singh, and S. Jindal, 'Rust Disease Classification Using Deep Learning Based Algorithm: The Case of Wheat', in Food Systems Resilience, A. I. Ribeiro-Barros, D. S. Tevera, L. F. Goulao, and L. D. Tivana, Eds. Rijeka: IntechOpen, 2022.

[9] J. Parraga-Alava, K. Cusme, A. Loor, and E. Santander, 'RoCoLe: A robusta coffee leaf images dataset for evaluation of machine learning based methods in plant diseases recognition', Data in Brief, vol. 25, p. 104414, 2019.

[10] X. Qin, Z. Zhang, C. Huang, M. Dehghan, O. R. Zaiane, and M. Jagersand, 'U2-Net: Going deeper with nested U-structure for salient object detection', Pattern Recognition, vol. 106, p. 107404, Oct. 2020.

[11] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, 'Rethinking atrous convolution for semantic image segmentation', arXiv preprint arXiv:1706. 05587, 2017.

[12] K. Simonyan and A. Zisserman, 'Very deep convolutional networks for large-scale image recognition', arXiv preprint arXiv:1409. 1556, 2014.

[13] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, 'Rethinking the inception architecture for computer vision', in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 2818–2826.

[14] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, 'Mobilenetv2: Inverted residuals and linear bottlenecks', in Proceedings of the IEEE Conference on computer vision and pattern recognition, 2018, pp. 4510–4520.