Explainable AI(XAI) using LIME and Disease Detection in Mango Leaf by Transfer Learning Approach

Vishakha Mistry

Head of Department, Department of Information Technology, 360 Research Foundation, Tumkaria, Bihar, India

---***---

Abstract - *Agriculture is an essential means of earning income for a significant percentage of the worldwide population. As a result, the productivity of crops has become crucial all over the world. Farmers will get more benefits from using modern digital tools for autonomous disease detection. Because agriculture is a complex field, it is particularly essential to improve the interpretability of agricultural ML models. First, the article proposes to identify leaf disease in Mango plants via pre-trained deep-learning architecture. Secondly, for the purpose of demonstrating the interpretability of my model's choice, I made use of Local Interpretable Model-Agnostic Explanations (LIME), an explainable AI (XAI) tool.*

Key Words: leaf disease detection, Transfer learning, explainable Artificial Intelligence (XAI), Local Interpretable Model-agnostic Explanations (LIME).

1. INTRODUCTION

Artificial intelligence and machine learning are now used in diverse agricultural applications. The mango fruit tree is a highly cultivated crop that is economically important across the majority of the world. In general, it is highly prized because of its beneficial nutrient content and mouthwatering taste, and it plays an important part in the lives of millions of farmers. However, this crop is susceptible to a variety of illnesses, which can reduce the yield and overall quality of the plants. The quick identification of these diseases in mango plants is needed to prevent the spread of disease. The original method for identifying mango leaf disease was visual examination by agricultural professionals. Which has time commitment and knowledge-dependent limitations. Automatic leaf disease identification is accomplished via the use of machine learning and computer vision disciplines.

The goal of this research is to identify mango leaf disease using a pre-trained transfer learning-based machine learning system. Model interpretability must also be researched as the number of convolution neural networks grows and the framework black box interpretability problem becomes more relevant. In order to make the models more transparent and interpretable, explanatory artificial intelligence (XAI) reveals its importance, which is considered to be at the highest level of explainability, accuracy, and performance [1]. Researchers are better able to comprehend the reasoning behind the outcomes of deep learning frameworks when the output of the LIME model is visually represented.

2. LITERATURE SURVEY

Several academics offered numerous machine learning (ML) and deep learning (DL) methodologies for detecting various diseases in plant leaves.

Adi Dwifana Saputra , Djarot Hindarto, Handri Santoso [2] have presented rice leaves disease classification using the Convolutional Neural Network algorithm with DenseNet architecture. The accuracy of DenseNet211 was 91.67%, that of DenseNet169 was 90%, and that of DenseNet201 was 88.33%. The training duration of the model was 24 seconds.

Authors in [3] have used PlantVillage and PlantDoc dataset to identify leaf disease in the corn plant. EfficientNetB0 architecture was used in this research article. The performance of the proposed architecture is compared with Inception V3, VGG16, Resnet50, Resnet101, and Densenet121. The proposed approach achieved an accuracy of 98.85% and a precision of 88% and it is more computationally efficient.

H. Amin et al. [4] have used two pre-trained CNN architectures namely EfficientNetB0 and DenseNet121. The Authors have applied feature fusion techniques features extracted from two models. The proposed model achieved 98.56% accuracy which is the highest amongst Resnet152, InceptionV3, and Densenet121.

Authors in [5] have performed classification of leaf disease on different fruit leaves. The average accuracy of VGG, GoogLeNet and ResNet are compared and ResNet have shown best accuracy amongst all. Explainability testing is done using GradCAM, LIME, and SmoothGrad techniques on convolution-based neural networks.

3. DATASET DESCRIPTION

My dataset was gathered from Kaggle. The dataset includes pictures of 32 different types of Indian mango leaves. The collection includes 768 photos of 32 Indian mango leaf species, with 24 photos of each species taken at various orientations and angles. The dataset's sample images are seen in Figure 1.

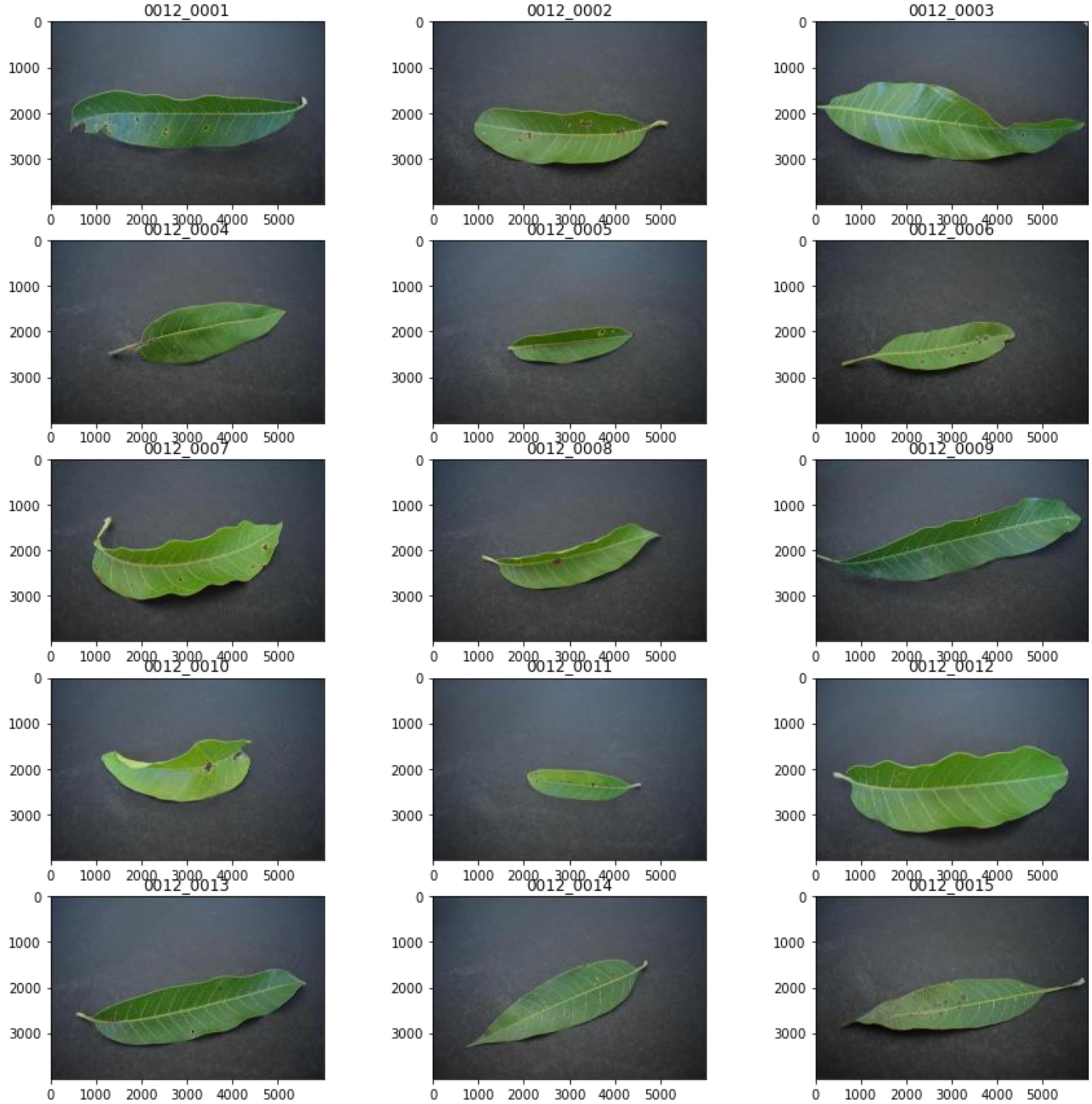


Fig -1: sample dataset diseased images

With a ratio of 70:20:10, I broke down the data into train, validation, and test sets. My model has been trained and validated using the training and validation set. To achieve pixel normalization, I divide each image pixel by 255. In order to make sure that my network, once trained, sees new variants of data at each and every epoch, I applied on-the-fly data augmentation over the training samples.

4. PROPOSED TRANSFER LEARNING APPROACH DENSENET169

For image detection, one of the recent developments in neural networks is the DenseNet architecture. ResNet and DenseNet are fairly similar, yet there are several key distinctions. While DenseNet concatenates (.) the output of the previous layer with the future layer, ResNet employs an additive approach (+) to combine the previous layer (identity) with the future layer [2]. There are various variants of DenseNet, including DenseNet-121, DenseNet-169, and DenseNet-201. The number of Densenet represents the neural network's layer count. The convolution layer, pooling layer, and fully linked layer are the three layers that make up the DenseNet-169 Architecture. By applying the convolution layer, pooling layer, batch normalization, and nonlinear activation layer, an output from the previous layer serves as an input for the second layer. DenseNet (Dense Convolutional Network) has several advantages: they alleviate the vanishing-gradient, problem, strengthen feature propagation, encourage feature reuse, and substantially reduce the number of parameters [2].

Performance analysis of the proposed architecture is presented in Figure 2. Line plots of the accuracy of train and test sets and loss of train and test sets show that there is no sign of underfitting or overfitting. Researchers now choose to assess their models in conjunction with accuracy using the F1 score, which combines precision and recall using their harmonic mean. As a result, I also computed the F1 score as a performance metric. My model has a good separability as seen by the AUC of 0.9984 in Figure 2.

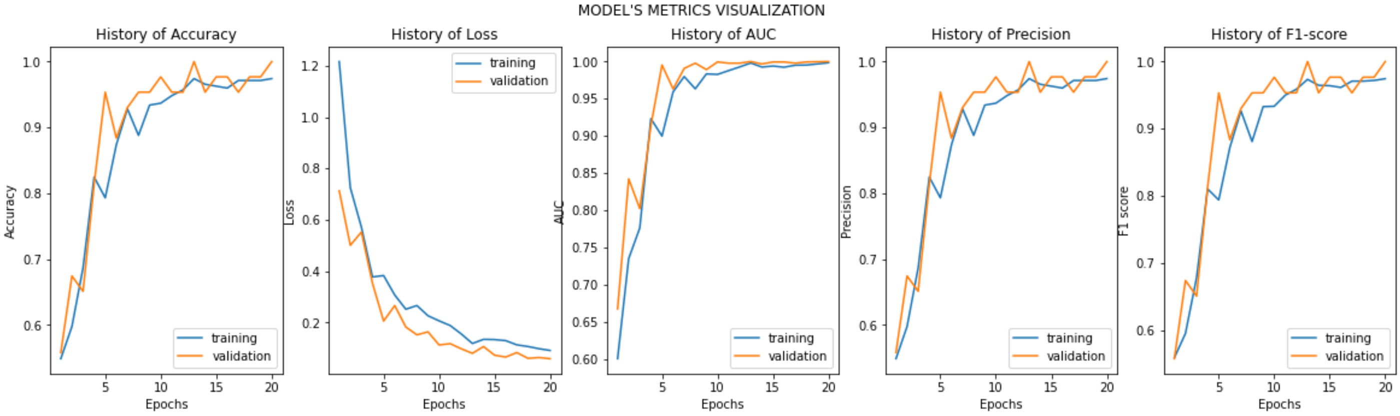


Fig -2: Graph of Accuracy, AUC, Precision, F1-score

5. EXPLAINABLE DEEP LEARNING FRAMEWORK(LIME)

5.1 Why LIME

Shapley and gradient-based explainability are two alternative approaches to black box explanation. These techniques explain individual input samples by taking the partial derivatives of a model's output with regard to its inputs [7]. Such gradient-based techniques offer a significant amount of output features, which contributes to the high-dimensionality issue. In addition to being computationally expensive, Shapley values produce a large number of output features with complex explanations [8]. LIME is more focused on providing individualized predicted explanations than these two approaches, and it is also easier to understand while requiring less computing power and effort.

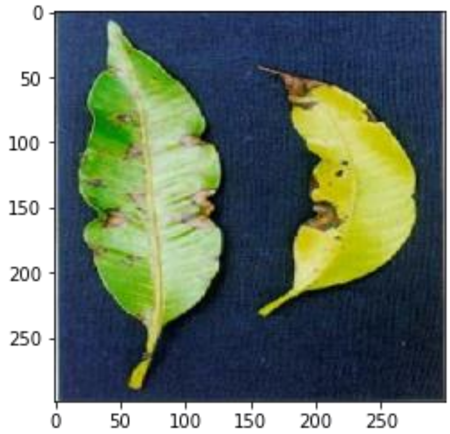
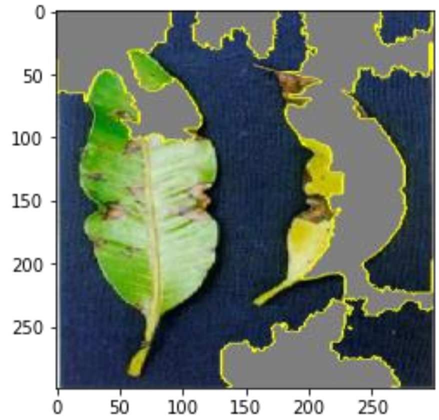
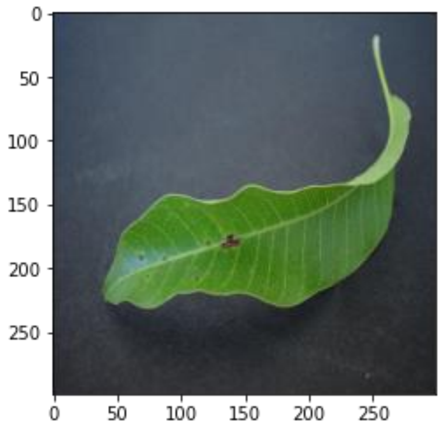
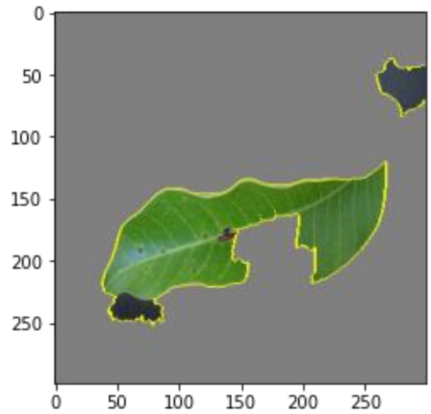
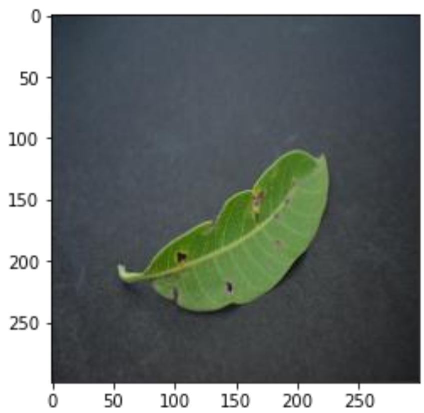
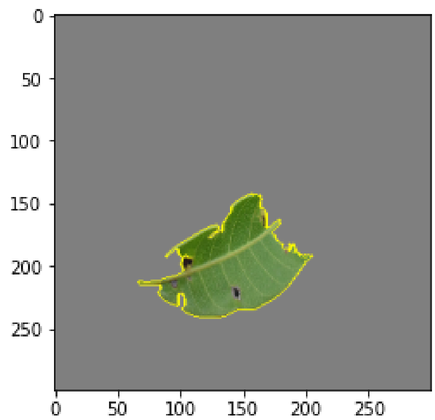
5.2 LIME interpretation of DenseNet169

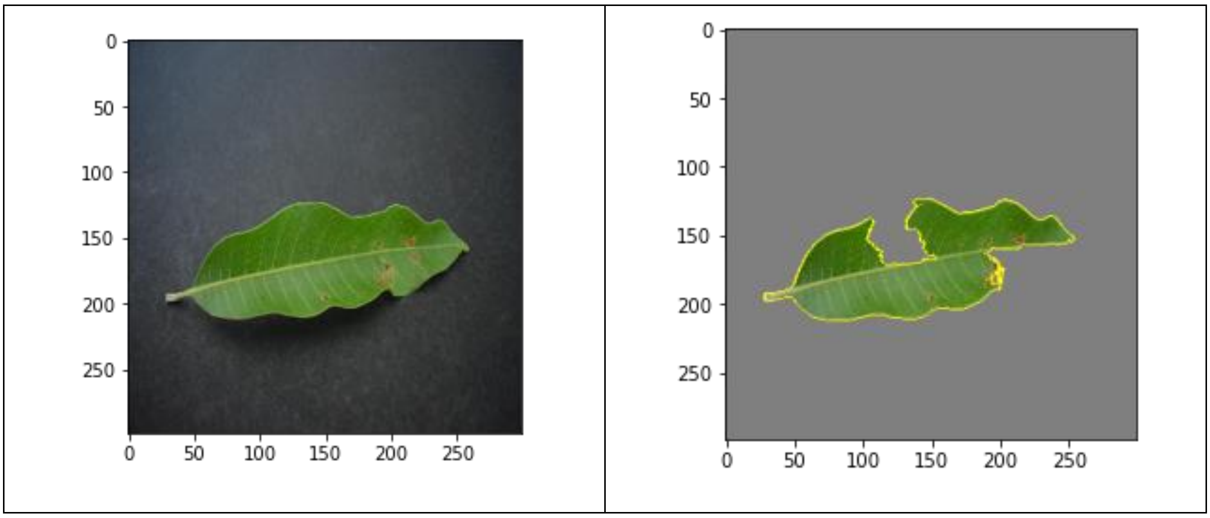
Explainable AI is a collection of procedures and techniques designed to make judgments made by AI and machine learning models intelligible to humans [6]. There are numerous methods for interpreting the machine Learning results. Here, Local Interpretable Model-agnostic Explanations (LIME)—a tool for interpreting machine learning models—have been utilized. This paper examines the classification of mango leaf diseases with further justifications based on Local Interpretable Model-agnostic Explanations (LIME).

LIME's main concept is to divide the input image into superpixels, which can be considered as new examples. We generated these superpixels by applying the Quickshift method [9]. Perturbations are produced by randomly sampling the superpixels and the prediction for each perturbation is computed. To ensure explainability, weights—a measure of the cosine distance—are calculated between each randomly generated perturbation and the input image that is given to the LIME. The weighted linear regression model is fitted with these weights, perturbations, and predictions and the model yields interpretable coefficients. Finally, LIME output gives superpixels that have a greater influence on the label prediction.

Table 1 displays the output that LIME generated to explain the predictions made by the Machine Learning Model. Images in the first column are the original image with the disease in Mango leaves. The features that enable the DenseNet model to predict diseased mango leaves are visualized in the second column of Table 1. These features contribute positive results to the prediction. We can observe from the prediction interpretations that the model mostly focuses on the leaf regions that are impacted when making predictions.

Table -1: LIME interpretation of DenseNet169 on diseased Mango leaves

Original Image	LIME output
	
	
	



6. CONCLUSION

In this paper, I implemented a pre-trained Deep learning model DenseNet169 for Mango leaf Disease detection. I have evaluated the model based on Accuracy, Precision, AUC, and F1 score. The model had the best performance, with a 97.41% accuracy rate. In the future to improve the generalization of my model, I will try to increase my dataset to cover a larger spectrum of plant leaf diseases.

REFERENCES

[1] Daglarli, Evren, "Explainable Artificial Intelligence (xAI) Approaches and Deep Meta-Learning Models for Cyber-Physical Systems", *Artificial Intelligence Paradigms for Smart Cyber-Physical Systems,* edited by Ashish Kumar Luhach and Atilla Elçi, IGI Global, 2021, pp. 42-67.

[2] Adi Dwifana Saputra , Djarot Hindarto, Handri Santoso, "Disease Classification on Rice Leaves using DenseNet121, DenseNet169, DenseNet201*", Sinkron : Jurnal dan Penelitian Teknik Informatika* Volume 8, Issue 1, January 2023, DOI : https://doi.org/10.33395/sinkron.v8i1.11906

[3] Fathimathul Rajeena, Aswathy S, Mohamed A. Moustafa and Mona A. S. Ali, "Detecting Plant Disease in Corn Leaf Using EfficientNet Architecture—An Analytical Approach", *Electronics* 2023, https://doi.org/10.3390/electronics12081938

[4] HASSAN AMIN, ASHRAF DARWISH (Member, IEEE), ABOUL ELLA HASSANIEN AND MONA SOLIMAN. "End-to-End Deep Learning Model for Corn Leaf Disease Classification", *IEEE Access*, ", Volume 10, 2022, Digital Object Identifier 10.1109/ACCESS.2022.3159678

[5] Kaihua Wei, Bojian Chen, Jingcheng Zhang , Shanhui Fan, Kaihua Wu, Guangyu Liu and Dongmei Chen, "Explainable Deep Learning Study for Leaf Disease Classification", *Agronomy* 2022, 12, 1035, https://doi.org/ 10.3390/agronomy12051035

[6] E. Tjoa and C. Guan, "A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 32, no. 11, pp. 4793–4813, 2021

[7] D. Baehrens, T. Schroeter, S. Harmeling, M. Kawanabe, K. Hansen, and K. R. Müller, "How to explain individual classification decisions," *J. Mach. Learn. Res.*, vol. 11, pp. 1803–1831, 2010

[8] S. M. Lundberg and S. I. Lee, "A unified approach to interpreting model predictions," *Adv. Neural Inf. Process. Syst.,* vol. 2017-Decem, no. Section 2, pp. 4766–4775, 2017

[9] A. Vedaldi and S. Soatto, "Quick Shift and Kernel Methods for Mode Seeking", *Lecture Notes in Computer Science*, pp. 705–718, 2008, https://link.springer.com/chapter/10.1007%2F978-3-540-88693-8_52.

BIOGRAPHY

Vishakha is working as a HOD of the Department of IT at 360 Research Foundation, she establishes a research agenda in **AI/ML**, helps researchers in the field of empowerment and livelihood, and guides research work to engineering and computer science students. She earned her M.Tech. from (NIT), Surat, Gujarat. She is an experienced Assistant Professor with a variety of engineering educational institutes including NIT-Surat and R V College of Engineering, Bangalore. She is passionate about AI/ML, speech and audio signal processing, and Digital Signal Processing.