

Explainable AI techniques for Deep Convolutional Neural Network based plant disease identification

Abstract—Deep learning-based computer vision has shown improved performance in image classification tasks. Due to the complexities of these models, they have been referred as opaque models. As a result, users need justifications for predictions to enhance trust. Thus, Explainable Artificial Intelligence (XAI) provides various techniques to explain predictions. Explanations play a vital role in practical application, to apply the exact treatment for a plant disease. However, application of XAI techniques in plant disease identification is not popular. This paper discusses the key concerns and taxonomies available in XAI and summarizes the recent developments. Also, it develops a tomato disease classification model and uses different XAI techniques to validate model predictions. It includes a comparative analysis of XAI techniques and discusses the limitations and usefulness of the techniques in plant disease symptom localization.

Index Terms—Explainable Artificial Intelligence, Plant Disease, Deep Learning

I. INTRODUCTION

Novel advancements in Machine Learning (ML) algorithms create Artificial intelligence (AI) systems that can learn, decide, perceive, and act independently. These algorithms have shown remarkable accomplishments which include autonomous vehicles, health care, legal, finance, and military [1]. In addition, these improvements have significantly impacted daily lives over the past few decades. Further, the use of Deep Learning (DL) based architectures has drastically improved the performance of speech recognition and computer vision exceptionally [2]. The use of various models such as VGG, GoogleNet, ResNet, and EfficientNet has shown powerful capabilities in image classification tasks. However, these models are considered as opaque, despite having high performance. Further, the performance of ML and explainability seem to be inherently at odds with one another. Generally, the most accurate approaches are the least explainable [3]. Initial AI systems were easily understood, but in recent years, opaque decision-making systems, such as Deep Neural Networks (DNN), have become more prevalent. Their practical success is a result of large parametric space and efficient learning methods. These models have millions of parameters and hundreds of layers, which make them complicated black-box models. A transparent model, which enables a clear comprehension of a model's operating mechanism, is the opposite of a black-box model [4]. The requirement for transparency plays a key role in black-box models as these models are mainly employed in making significant predictions in crucial scenarios [5]. The DNN's opaque nature makes it difficult to explain the inner workings of the trained models. The increase of interest

in using ML algorithms in crucial fields namely disease identification, autonomous vehicles, and robotics has enabled numerous benefits, however, their usefulness is constrained by the incapacity to provide justifications for the decision. As AI-based systems become more prevalent, concerns regarding their ethics, fairness, and safety are becoming more important [6]. It is vital to know the justifications for the critical decision that could change a person's life, such as disease diagnosis. It is clear from this point how important it is to explain AI results. With the focus on high performance, Explainable Artificial Intelligence (XAI) tries to develop a set of methods that makes the model more explainable [1]. XAI models are used to comprehend, trust, and manage decisions. In the beginning, AI systems were primarily logical and symbolic and they performed logical inferences and offered steps involved in inference as the explanation. However, these explanations could not provide substantial justifications for their decisions. Due to rapid success of ML, XAI has become a major component in academic libraries. It is a collection of methods, procedures, and tactics that have been used by libraries to ensure that ML serves librarianship, the academy, and the public interest [7]. According to DARPA goal of XAI is to "create a suite of new or modified ML techniques that produce explainable models that, when combined with effective explanation techniques, enable end users to understand, appropriately trust, and effectively manage the emerging generation of AI systems" [3].

Agriculture plays a vital role in increasing the economic growth of a country by providing employment opportunities and ensuring food security. However, numerous factors such as soil moisture, climate, weeds, humidity, and plant diseases mainly affect agriculture production. Among those plant diseases cause significant losses, reducing agricultural productivity. Early disease identification plays a vital in preventing the spread of the disease, reducing production loss, and limiting the use of pesticides. Generally, the identification of plant disease is carried out by experts such as agronomists or plant pathologists. However, the use of manual identification techniques has a few drawbacks. Fortunately, Deep Convolution Neural Network (DCNN) based plant disease identification techniques automate disease recognition providing higher performance. But, the complexity of these models makes them less transparent. Thus, these models are unable to offer justifications for the classification result. However, justifications are required to gain the user's trust as they are keener on knowing the symptoms that lead to a certain disease classification. Disease symptoms help experienced farmers to enhance their

trust in the model meanwhile it helps inexperienced farmers to make accurate decisions. Further, XAI models will help to identify the symptoms that were not earlier considered as important. This work initially presents the key concerns and key concepts in XAI. Further, it introduces novel XAI techniques and their application in plant disease identification. A model on a tomato dataset was trained and justifications were generated for the predictions using Gradient-weighted Class Activation Map++(Grad-CAM++), Score-CAM, and Local Interpretable Model Agnostic Explanations(LIME).

Our contributions,

- Discusses the key concerns and key concepts in XAI and developments in plant disease identification using DCNN.
- Implementation of a classifier for tomato disease identification.
- Generation of explanations using several XAI techniques and these explanations localize the disease symptoms that are responsible for the classification.
- Conduct a comparative analysis between the explanations generated by the XAI models.

II. BACKGROUND

A. Explainable AI

An interpretable system enables users not only to observe but also to analyze and comprehend the mathematical mapping of inputs to outputs [8]. In the area of ML, interpretability is more favored over explainable and is used interchangeably [9], [10]. The terms understandability [11], comprehensibility, and transparency have been also used to refer to the same issue [4]. Explanations depend on the context. An explanation's "quality" depends on the demands and objectives of both the explainer and the user. In most cases, explanations do not need to be complete or even be entirely accurate [7]. Therefore, a good explanation must consider who it is intended for. Critical fields require complete and more accurate explanations. Explanations enhance the user's trust, confidence and determine the accuracy of the prediction [12]. However, it is challenging to develop models and methods that are both precise and offer substantial justifications that satisfy customers' needs [1]. To define taxonomies of XAI many surveys have been carried out in different sectors [1] [4] [6], [13], [14].

Based on the methodology utilized, the types of data used, scope, stage, and applicability of the model, various authors have provided their perspectives on how to interpret the XAI landscape. A few authors further break these criteria down to provide a more detailed breakdown. These criteria's help researchers to easily identify the XAI landscape and apply the best technique to their problem in hand. Table I summarizes the important terms in XAI landscape. The recent XAI models were studied based on the data type they could be applied to, the scope of the explanation, the nature of the technique, and the method. Table II summarizes the existing XAI models.

III. PLANT DISEASE IDENTIFICATION

There are a number of factors that limit food productivity among those plant diseases cause a significant decrease in

TABLE I
XAI TERMS

XAI Term	Definition
Post-hoc	Provide explanations for inputs after the model training
Model-agnostic	Generate explanations without accessing the model internals
Model-specific	Could be applied only to certain models
Local explanations	Explains an output of a single instance
Global explanations	provides an overall understanding of the model predictions
Interpretability	Ability to explain or to present in understandable terms to a user [6]

both crop quality and yield. These factors directly impact the economy which results in higher food prices and low income of the farmer. Further, it may also lead to food shortages that may result in hunger and starvation. Earlier, plant diseases were identified through physical inspection by professionals. However, it is a time-consuming process and practically impossible to analyze huge growing areas. Further, lack of professional assistance causes frequent delays in disease identification which may directly impact the growth of the plant. Thus, disease identification in the early stage is critical as it increases production and reduce the use of higher amount of chemicals. Disease occurrences mainly depend on climate conditions, plant variety, soil, and the season. Resistance level for any disease depends on the variety of the plant. Recently, computer vision has provided promising results to effectively identify plant diseases. These models need huge datasets to train the models as a solution researchers use data augmentation [15] and transfer learning techniques [16]–[20]. In addition, many researchers have implemented custom DCNN models to identify plant diseases [21]–[25]. These models mainly focus on the implementation and optimization of the model.

IV. DEEP CNN MODEL TO IDENTIFY PLANT DISEASES

This section describes the details of the implemented DCNN-based tomato disease identification model. The tomato images were extracted from the plant village dataset to train the model.

A. Data set preparation

The dataset contains ten disease categories including healthy images. The total number of images considered were 16012. The original images were 256*256 and they were resized to 224*224 before the training. Images were divided into train, test, and validation sets ratio of 70:20:10.

B. Model Training and Evaluation Results and Discussion

EfficientNet architecture has shown powerful capabilities in plant disease identification [26]–[28]. Therefore, in this experiment, the EfficientNet-B0 model was employed with pre-trained weights of ImageNet. The model was trained using tomato images by unfreezing a few sequential layers in the bottom and the classifier layer in the model. This model was trained in Google Colab using Pytorch framework with

TABLE II
SUMMARY OF XAI MODELS

Ref	XAI	Scope	Method	Technique	Data
[30]	CAM	Local	MS	Post-hoc	Img
[31]	Grad-CAM	Local	MS	Post-hoc	Img
[32]	Grad-CAM++	Local	MS	Post-hoc	Img
[33]	LIME	Both	MA	Post-hoc	Any
[34]	SHAP	Both	MA	Post-hoc	Any
[35]	Score-CAM	Local	MS	Post-hoc	Img
[36]	Occlusion	Local	MS	Post-hoc	Img
[37]	RISE	Local	MS	Post-hoc	Img
[38]	LRP	Both	MS	Post-hoc	Img
[39]	Integrated Gradients	Local	MS	Post-hoc	Img
[40]	Saliency map	Local	MS	Post-hoc	Img
[41]	Deep Lift	Local	MS	Post-hoc	Img

^a MS-model specific, MA-model agnostic, Both-Local and global.

GPU acceleration. The trained model achieved 99.6% training accuracy and 99.1% test accuracy after 30 epochs. Fig.3 shows the classification report of the trained model.

C. Explaining DCNN model decisions

To improve the transparency of black-box models requires justifications to support prediction results. As a result, explanations were needed to guard prediction results. There is no proper definition for a good explanation. However, “Some say a good explanation depends on the question” [14]. To ensure algorithmic fairness, to spot potential bias or errors in the training data, and to make sure that algorithms are working as planned, explanations are therefore of utmost important [14]. Interpretability and completeness properties could be used to evaluate explanations. Interpretability could be defined as the ability to explain or present in a way that a human can understand [29]. The main focus of interpretability is to explain the internals of the model in a way it is comprehensible for humans. Accurately describing a system’s operation is the aim of completeness. For example, when explaining a DNN model, it is always possible to provide a complete explanation by disclosing every mathematical operation and system parameter [14]. Toja et.al [13] has introduced two major categories of interpretability, namely “perceptive interpretability and interpretability by mathematical structures”. The interpretability methods which are easily understood by humans are categorized under perceptive interpretability (e.g.: LIME, CAM, Deep Lift). The latter explains a model by using the available mathematical knowledge. According to [29] even though ML models provide higher predictive performance classification accuracy alone is not sufficient to explain real-world problems. Therefore, interpretability methods are required to improve the

trust of the model. Highly interpretable models could provide sufficient evidence for any model prediction to enhance trust in the model. Also, these evidences provide clear clarifications for humans to understand the reasons for certain predictions as these are more critical in some domains. The following characteristics show the interpretability measures of ML models. [29].

- Fairness: Makes sure the model is fair(unbiased) and does not implicitly or explicitly discriminate against protected groups.
- Privacy: Ensure that the data’s sensitive information is protected.
- Reliability/Robustness: Make sure that a slight change in the input does not result in a significant change in the prediction.
- Causality: Make sure that the model is only picking up information from causal relationships.
- Trust: Humans are more likely to trust a model that provides explanations for its decisions(predictions) than a black box model that simply produces predictions.

In order to explain the internal workings of the black-box models and to enhance transparency, XAI techniques were introduced. Hence, these techniques improve the interpretability of the DNN models. Generally, explanations should be simple and concise. Thus, end users would be able to understand the reasons for the classification easily. Simple forms of explanations comprise visualizations, decision rules, and text. A domain like agriculture needs justifications for a prediction to apply the exact treatment for plant disease. XAI models provide visual explanations by highlighting the key regions that positively contribute to the classification. These explanations are very useful for farmers as they can understand the reasons for the decision-making process based on the highlighted infected regions in the plant leaf. However, the application of XAI techniques to provide justification for prediction is not very popular in the agriculture field. In plant disease symptom identification, researchers have mainly employed class activation map-based techniques. These techniques mainly highlight the symptoms that are positively contributing to the classification. Among Class Activation Mapping(CAM)-based techniques, the most commonly applied technique is Grad-CAM. This technique has been frequently applied as it is simple and cost-effective. Table III summarizes the studies that have employed XAI models. The Class Activation Mapping is the most primary technique among other CAM-based methods, which is architecture-sensitive. It identifies the key regions of the input image by projecting the prediction layer weights back onto the convolutional feature map [30]. This technique has been rarely applied in plant disease classification models. Grad-CAM is a generalized version of CAM that could be applied to any model. It assigns importance to each activation map by computing gradients of the target class with respect to the activation map and finally computes the saliency map as a linear weighted sum of the importance score and the activation map [31]. However, this model cannot locate

the multiple instances of the same class in a single image. Grad-CAM++ addresses this limitation available in Grad-CAM and enables the identification of multiple symptoms of a single image. It assigns importance to each activation map as a weighted linear sum of positive partial derivatives that enables the identification of multiple occurrences of the same class [32]. Techniques such as Grad-CAM and Grad-CAM++ use gradient information to generate the explanation. Therefore, these models suffer from two main drawbacks gradient saturation and false confidence [35]. The score-CAM addresses the limitations available in gradient-based techniques by generating the saliency map as a linear weighted sum of the forward passing score and the activation map. LIME is considered the most prominent approach in the XAI domain. However, it has been rarely applied in the agriculture domain as it requires a higher computational cost. It alters the input image by giving variations into input features and generates a new set of images. Then perform forward passes for perturb images and next, assign weights to each image based on the proximity to the original image. LIME finally fits a simple interpretable model for the predictions, weights, and the new image sets and approximates the behavior of the complex model [33]. Fig.2 summarizes the Score-CAM explanations generated for tomato diseases. Score-CAM explanations identify most of the symptoms accurately. However, when plant disease symptoms appear all over the plant leaf, it can not locate all the symptoms accurately. For example, yellow leaf curl symptoms(curling and yellowing) appear all over the plant leaf according to the figure explanation identifies only part of the symptoms and it highlights the entire region as the curliness is spread all over that region. In the case of leaf mold, symptoms have spread all over the leaf as small yellow flakes. Score-CAM highlights the entire region without highlighting the symptoms separately. Fig. 4 shows the explanations generated for the LIME method. According to the explanations LIME explanation size is quite large. Thus, it identifies irrelevant image regions in addition to the relevant regions as the explanation. Also, explanations do not locate all the symptoms. Fig.1 compares the explanations generated using Grad-CAM++, Score-CAM, and LIME. In some cases, Grad-CAM++ performs similarly to Score-CAM. Sometimes Score-CAM outperforms Grad-CAM++ by identifying disease symptoms more accurately. However, explanations generated by both techniques are noisy. Therefore, these explanations are not understandable for farmers. In the case of LIME, though the explanations identify relevant regions accurately in most of the explanations, in addition, it identifies irrelevant regions as the explanation. Compared to other methods, LIME explanation size is not sensitive for plant disease symptom identification. Further, the large explanation size in LIME is not suitable for plant disease symptom localization. However, even though the size of the explanation could be configurable, decreasing the explanation size in LIME will improve the computational time.

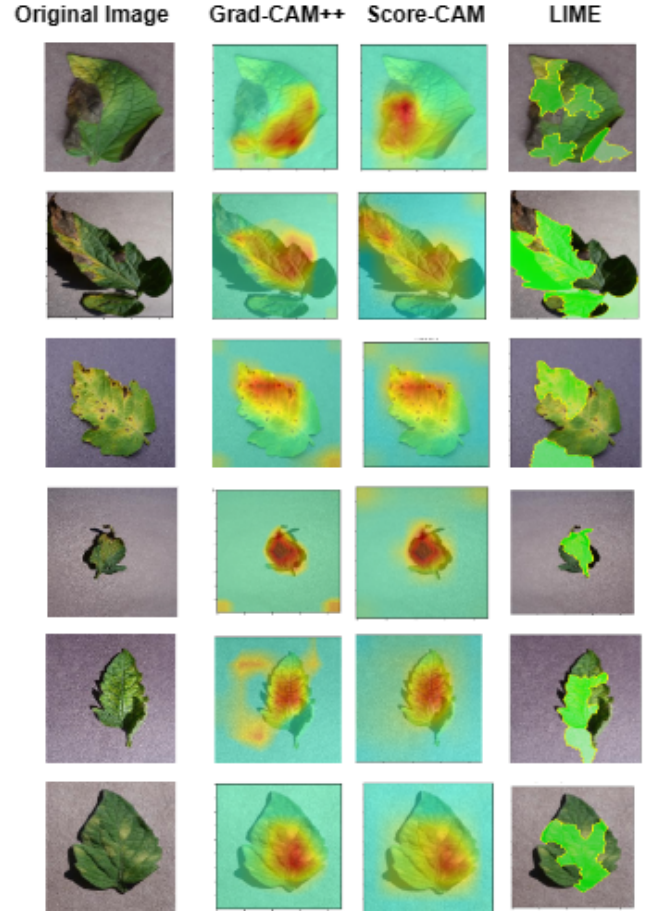
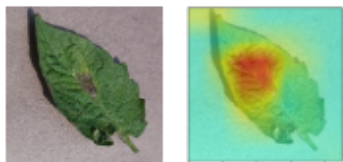


Fig. 1. Comparison of explanations generated using Grad-CAM++,Score-CAM,and LIME.

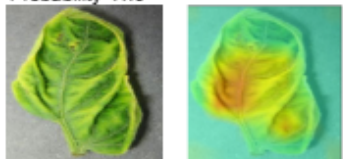
TABLE III
SUMMARY OF THE MODELS EMPLOY XAI MODELS

Ref	Dataset	Visualization technique
[42]	Cucumber	Grad-CAM
	Wheat , google images	Grad-CAM
[43]	coffee	Grad-CAM,ScoreCAM,Grad-CAM++
[44]	Beans	Grad-CAM
[45]	Grapes	CAM
[46]	Plant Village	Grad-CAM
[26]	Wheat	Grad-CAM
[47]	PlantVillage	Saliency map.
[48]	Soybean	Saliency map
[49]	PlantVillage and pepper	Grad-CAM
[50]	Tomato(PlantVillage)	Occlusion
[18]		

predicted class : Tomato Late Blight
Probability : 0.937048614025116



predicted class : Tomato Leaf Curl Virus
Probability :1.0



predicted class : TomatoSeptoria_leaf_spot
Probability :0.9999996423721313



predicted class : Tomato Leaf Mold
Probability :1.0

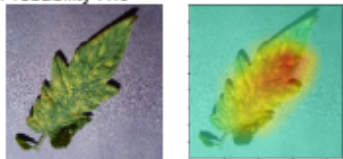


Fig. 2. Explanations generated using Score-CAM.

	accuracy	precision	recall	f1_score
Tomato_Bacterial_spot	0.998	1.000	0.986	0.993
Tomato_Early_blight	0.997	0.990	0.968	0.975
Tomato_late_blight	0.998	0.990	0.990	0.990
Tomato_leaf_mold	0.999	0.980	1.000	0.990
Tomato_Septoria_leaf_spot	0.999	0.994	1.000	0.997
Tomato_Spider_mites_two_spotted_spider_mite	0.997	0.982	0.988	0.985
Tomato__Target_Spot	0.997	0.979	0.986	0.982
Tomato__Tomato_YellowLeaf__Curl_Virus	0.999	0.997	0.997	0.997
Tomato__Tomato_mosaic_virus	0.999	0.974	1.000	0.987
Tomato_healthy	0.999	0.994	0.994	0.994

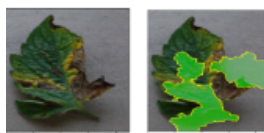
Fig. 3. Classification report

V. CONCLUSION

In this work, the key concerns in XAI and taxonomy of interpretable models in ML were discussed comprehensively. In addition, recent plant disease identification models were studied and many researchers have focused only on developing classification models and optimizing the model, few researchers have employed XAI techniques to validate the model. Further, the study discussed the recent developments in the field of XAI and their applicability in plant disease identification. Moreover, a tomato disease classification model is implemented using a pre-trained architecture Efficient-NetBo, and generated explanations for the predictions. The

predicted class : Tomato Early Blight

probability :1.0



predicted class: Tomato Early Blight
Probability :0.996356725692749



predicted class : Tomato Septoria leafspot
Probability 0.9999996423721313

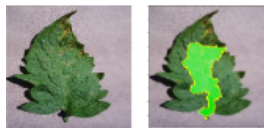


Fig. 4. Explanations generated using LIME.

comparative analysis between the XAI explanations depicted that explanations are unable to locate the disease symptoms more accurately when plant disease symptoms appear all over the plant leaf. In addition, CAM-based explanations are noisy. Thus, explanations are not comprehensible for the farmers. In addition, the explanation size of LIME is not sensitive for plant disease symptom identification. Therefore, further experiment is needed to develop novel XAI techniques to explain plant diseases.

REFERENCES

- [1] A. Adadi and M. Berrada, "Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)," *IEEE Access*, vol. 6, pp. 52138–52160, Sep. 2018.
- [2] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, "A Survey of Deep Neural Network Architectures and Their Applications," *Neurocomputing*, vol. 234, pp. 11–26, 2017.
- [3] D. Gunning and D. W. Aha, "DARPA's Explainable Artificial Intelligence Program Deep Learning and Security," *AI Magazine*, pp. 44–58, 2019.
- [4] A. B. Arrieta et al., "Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI," *Information fusion*, vol. 58, pp. 82–115, Oct. 2020.
- [5] A. Preece, D. Harborne, D. Braines, R. Tomsett, and S. Chakraborty, "Stakeholders in Explainable AI," Sep. 2018.
- [6] T. Speith, "A Review of Taxonomies of Explainable Artificial Intelligence (XAI) Methods," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Jun. 2022, pp. 2239–2250.
- [7] M. Ridley, "Explainable Artificial Intelligence (XAI)," *Information Technology and Libraries*, vol. 41, no. 2, Jun. 2022.
- [8] D. Doran, S. Schulz, and T. R. Besold, "What Does Explainable AI Really Mean? A New Conceptualization of Perspectives," Oct. 2017.
- [9] M. Bojarski et al., "Explaining How a Deep Neural Network Trained with End-to-End Learning Steers a Car," Apr. 2017.
- [10] P. Wei Koh and P. Liang, "Understanding Black-box Predictions via Influence Functions," 2017, pp. 1885–18894.
- [11] A. Andrzejak, F. Langner, and S. Zabala, "Interpretable Models from Distributed Data via Merging of Decision Trees," in *IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*, 2013.

- [12] O. Biran N-Join and C. Cotton N-Join, "Explanation and Justification in Machine Learning: A Survey," in *IJCAI-17 workshop on explainable AI (XAI)*, 2017, pp. 8–13.
- [13] E. Tjoa and C. Guan, "A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI," *IEEE Trans Neural Netw Learn Syst*, vol. 32, no. 11, pp. 4793–4813, Nov. 2021.
- [14] L. H. Gilpin, D. Bau, B. Z. Yuan, A. Bajwa, M. Specter, and L. Kagal, "Explaining Explanations: An Overview of Interpretability of Machine Learning," in *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*, May 2018, pp. 80–89.
- [15] G. Geetharamani and A. P. J., "Identification of plant leaf diseases using a nine-layer deep convolutional neural network," *Computers and Electrical Engineering*, vol. 76, pp. 323–338, Jun. 2019.
- [16] J. G. Arnal Barbedo, "Plant disease identification from individual lesions and spots using deep learning," *Biosyst Eng*, vol. 180, pp. 96–107, Apr. 2019.
- [17] A. K. Rangarajan, R. Purushothaman, and A. Ramesh, "Tomato crop disease classification using pre-trained deep learning algorithm," in *Procedia Computer Science*, Elsevier B.V., 2018, pp. 1040–1047.
- [18] M. Brahimi, K. Boukhalfa, and A. Moussaoui, "Deep Learning for Tomato Diseases: Classification and Symptoms Visualization," *Applied Artificial Intelligence*, vol. 31, no. 4, pp. 299–315, Apr. 2017.
- [19] X. Zhang, Y. Qiao, F. Meng, C. Fan, and M. Zhang, "Identification of maize leaf diseases using improved deep convolutional neural networks," *IEEE Access*, vol. 6, pp. 30370–30377, 2018.
- [20] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," *Comput Electron Agric*, vol. 145, pp. 311–318, Feb. 2018.
- [21] P. Jiang, Y. Chen, B. Liu, D. He, and C. Liang, "Real-Time Detection of Apple Leaf Diseases Using Deep Learning Approach Based on Improved Convolutional Neural Networks," *IEEE Access*, vol. 7, pp. 59069–59080, 2019.
- [22] M. Arsenovic, M. Karanovic, S. Sladojevic, A. Anderla, and D. Stefanovic, "Solving current limitations of deep learning based approaches for plant disease detection," *Symmetry (Basel)*, vol. 11, no. 7, Jul. 2019.
- [23] Y. Lu, S. Yi, N. Zeng, Y. Liu, and Y. Zhang, "Identification of rice diseases using deep convolutional neural networks," *Neurocomputing*, vol. 267, pp. 378–384, Dec. 2017.
- [24] C. DeChant et al., "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.
- [25] K. KC, Z. Yin, M. Wu, and Z. Wu, "Depthwise separable convolution architectures for plant disease classification," *Comput Electron Agric*, vol. 165, Oct. 2019.
- [26] N. Shukla, S. Palwe, Shubham, M. Rajani, and A. Suri, "Plant Disease Detection and Localization using GRAD-CAM," *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 8, no. 6, pp. 3069–3075, Mar. 2020.
- [27] P. Zhang, L. Yang, and D. Li, "EfficientNet-B4-Ranger: A novel method for greenhouse cucumber disease recognition under natural complex environment," *Comput Electron Agric*, vol. 176, Sep. 2020.
- [28] Ü. Atila, M. Uçar, K. Akyol, and E. Uçar, "Plant leaf disease classification using EfficientNet deep learning model," *Ecol Inform*, vol. 61, Mar. 2021.
- [29] F. Doshi-Velez and B. Kim, "Towards A Rigorous Science of Interpretable Machine Learning," Feb. 2017.
- [30] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning Deep Features for Discriminative Localization," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2921–2929.
- [31] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization," in *Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy: 2017 IEEE International Conference on Computer Vision (ICCV)*, Oct. 2017, pp. 618–626.
- [32] A. Chattopadhyay, A. Sarkar, P. Howlader, and V. N. Balasubramanian, "Grad-CAM++: Generalized gradient-based visual explanations for deep convolutional networks," in *Proceedings - 2018 IEEE Winter Conference on Applications of Computer Vision, WACV 2018, Institute of Electrical and Electronics Engineers Inc.*, May 2018, pp. 839–847.
- [33] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should i trust you? Explaining the predictions of any classifier," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Association for Computing Machinery*, Aug. 2016, pp. 1135–1144.
- [34] S. M. Lundberg, P. G. Allen, and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," in *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems*, 2017, pp. 4768–4777.
- [35] H. Wang et al., "Score-CAM: Score-Weighted Visual Explanations for Convolutional Neural Networks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2020, pp. 24–25.
- [36] M. D. Zeiler and R. Fergus, "LNC8 8689 - Visualizing and Understanding Convolutional Networks," in *European Conference on Computer Vision*, Springer, 2014, pp. 818–833.
- [37] V. Petsiuk, A. Das, and K. Saenko, "RISE: Randomized Input Sampling for Explanation of Black-box Models," Jun. 2018.
- [38] S. Bach, A. Binder, G. Montavon, F. Klauschen, K. R. Müller, and W. Samek, "On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation," *PLoS One*, vol. 10, no. 7, Jul. 2015.
- [39] M. Sundararajan, A. Taly, and Q. Yan, "Axiomatic Attribution for Deep Networks," Mar. 2017.
- [40] K. Simonyan, A. Vedaldi, and A. Zisserman, "Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps," Dec. 2013.
- [41] A. Shrikumar, P. Greenside, and A. Kundaje, "Learning Important Features Through Propagating Activation Differences," in *ICML'17: Proceedings of the 34th International Conference on Machine Learning -*, Apr. 2017, pp. 3145–3153.
- [42] E. Fujita, H. Uga, S. Kagiwada, and H. Iyatomi, "A Practical Plant Diagnosis System for Field Leaf Images and Feature Visualization," *International Journal of Engineering and Technology*, vol. 7, no. 11, pp. 49–54, 2018.
- [43] M. A. Genaev, E. S. Skolotneva, E. I. Gulyaeva, E. A. Orlova, N. P. Bechtold, and D. A. Afonnikov, "Image-based wheat fungi diseases identification by deep learning," *Plants*, vol. 10, no. 8, Aug. 2021.
- [44] M. Yebasse, B. Shimelis, H. Warku, J. Ko, and K. J. Cheoi, "Coffee disease visualization and classification," *Plants*, vol. 10, no. 6, Jun. 2021.
- [45] P. Sahu, A. Chug, A. P. Singh, D. Singh, and R. P. Singh, "Deep Learning Models for Beans Crop Diseases: Classification and Visualization Techniques," *International Journal of Modern Agriculture*, vol. 10, no. 1, pp. 796–812, Mar. 2021.
- [46] B. Liu, Z. Ding, L. Tian, D. He, S. Li, and H. Wang, "Grape Leaf Disease Identification Using Improved Deep Convolutional Neural Networks," *Front Plant Sci*, vol. 11, Jul. 2020.
- [47] E. Ennadifi, S. Laraba, D. Vincke, B. Mercatoris, and B. Gosselin, "Wheat Diseases Classification and Localization Using Convolutional Neural Networks and GradCAM Visualization," in *2020 International Conference on Intelligent Systems and Computer Vision, ISCV 2020, Institute of Electrical and Electronics Engineers Inc.*, Jun. 2020.
- [48] M. Brahimi, M. Arsenovic, S. Laraba, S. Sladojevic, K. Boukhalfa, and A. Moussaoui, "Deep Learning for Plant Diseases: Detection and Saliency Map Visualisation," 2018, pp.93–117.
- [49] K. Nagasubramanian, S. Jones, A. K. Singh, S. Sarkar, A. Singh, and B. Ganapathysubramanian, "Plant disease identification using explainable 3D deep learning on hyperspectral images," *Plant Methods*, vol. 15, no. 1, Aug. 2019.
- [50] M. Ahmad, M. Abdullah, H. Moon, and D. Han, "Plant Disease Detection in Imbalanced Datasets Using Efficient Convolutional Neural Networks with Stepwise Transfer Learning," *IEEE Access*, vol. 9, pp. 140565–140580, 2021.