

Potato Leaf Disease Detection Approach Based on Transfer Learning with Spatial Attention

Rima Grati¹, Emna Ben Abdallah², Khouloud Boukadi² and Ahmed Smaoui²

¹Zayed University, Abu Dhabi, U.A.E.

²Mir@cl Laboratory, Sfax University, Sfax, Tunisia

Keywords: Smart Agriculture, Classification, Potato Leaf Disease, Transfer Learning, Explainable Artificial Intelligence.

Abstract: Agricultural productivity is vital to global economic development and growth. When crops are affected by diseases, it adversely impacts a nation's financial resources and agricultural output. Early detection of crop diseases can minimize losses for farmers and enhance production. Symptoms of diseases may take form in different parts of plants. However, the leaves, especially those of potatoes, are most commonly used in disease detection because they are buried deep in the ground. Deep learning-based CNN methods have become the standard for addressing most technical image identification and classification challenges. To improve training performance, the attention mechanism in deep learning helps the model concentrate on the informative data segments and extract the discriminative properties of inputs. This paper investigates spatial attention, which aims to highlight important local regions and extract more discriminative features. Moreover, the most popular CNN architectures, MobileNetV2, DenseNet121, and InceptionV3, were applied to transfer learning for potato disease classification and then fine-tuned by the publicly available dataset of PlantVillage. The experiments reveal that the proposed Att-MobileNetV2 model performs better than other state-of-the-art methods. It achieves an identification F-measure of 98% on the test dataset, including images from Google. Finally, we utilized Grad-CAM++ in conjunction with the Att-MobileNetV2 method to provide an interpretable explanation of the model's performance. This approach is particularly effective in localizing the predicted areas, clarifying how CNN-based models identify the disease, and ultimately helping farmers trust the model's predictions.

1 INTRODUCTION

Today, with the increase in the world's population, the need for food has risen considerably, making traditional agricultural methods insufficient. The global food system will have to provide healthy, nutritious food for a population that will rise from 7.5 billion today to almost 10 billion by 2050 (Bahar et al., 2020). Plants are threatened by diseases caused by microorganisms: viruses, bacteria and fungi. These diseases cause major yield losses in food, fruit, vegetable and ornamental crops, particularly in tropical and warm-temperate zones. In some cases, entire harvests or even entire industries are wiped out. And new diseases regularly emerge due to mutations in pathogens or their adaptation to new environments. As phytosanitary products are ineffective against bacteria and viruses, we need to find other ways of reducing the risk of infection, particularly by avoiding the introduction of these microorganisms.

Among the many crops affected by diseases, the potato occupies a central place in the world's diet. However, diseases such as late blight and common scab represent significant threats to potato production, resulting in major yield losses and considerable economic impact (Afzaal et al., 2021).

To meet these challenges, intelligent agriculture has emerged, aimed at reducing processes, cutting costs and improving the quality of agricultural production. This new approach injects intelligence into traditional farming practices, notably through intelligent detection or irrigation systems. These systems combine physical equipment with software applications incorporating advanced technologies such as Deep Learning, a promising Artificial Intelligence (AI) technology (Afzaal et al., 2021). The development of artificial intelligence, endowed with the ability to learn from experience, has seen significant growth. Deep learning applications have emerged useful in various sectors and industries, including in-

telligent agriculture. This approach aims to optimize the use of agricultural resources by calculating plants' water and nutrient requirements (Koné et al., 2023a), (Koné et al., 2023b). In addition, smart agriculture includes plant disease detection and phytosanitary diagnostics, which allows for more efficient chemical management. Monitoring plant health is critical to ensuring long-term output. Advanced technologies, such as AI-driven analytics and IoT sensors, are increasingly being used to offer real-time agricultural status. These advances not only increase production but also encourage environmentally friendly farming practices.

In this context, image analysis is invaluable for detecting plant diseases. Images play a crucial role in automatic disease identification, offering diverse conditions and symptom characteristics encountered in practice (Afzaal et al., 2021). Several recent studies have demonstrated the effectiveness of convolutional neural networks (CNNs) in solving these disease detection problems by extracting complex features from images to identify the type of infection. Different architectures, such as AlexNet, GoogleNet, LeNet, MobileNet and VGG16, have been used for this purpose (Liang et al., 2019; Kamal et al., 2019; Mahum et al., 2023).

However, most studies from the literature have concentrated on potato crop diseases, training their models solely on the PlantVillage dataset, failing to assess the algorithms' accuracy on previously unreported datasets. Furthermore, the authors failed to prioritize post-hoc explanations of the models, which is a critical oversight. In the context of smart agriculture, it is critical to detect damaged leaves and clearly explain these findings to the farmer to gain trust in the model's suggestions.

This study investigates the transfer learning for deep CNNs and modifies the network structure to improve the learning capability of minute lesion features. We select the most popular backbone models MobileNet-V2, DenseNet121 and InceptionV3. Based on the transfer learning, we transfer the common knowledge of the three pre-trained models on ImageNet and incorporate the Convolutional Block Attention Module (CBAM) (Woo et al., 2018) to create new networks for identifying potato plant diseases.

Furthermore, a gradient-based visualisation technique – grad-CAM++ (Chattopadhyay et al., 2018) has been integrated with the attention-based CNN models to deal with deep learning "black box" problems and to assist farmers with disease visualisation. The experimental results clearly demonstrate the efficiency of the proposed methodology, which successfully and

accurately completes the classification of potato plant diseases.

The rest of the paper is organized as follows. The related work is summarized in Section 2. Our proposed methodology is described in detail in Section 4, and then experimental results with their analysis are reported in Section 5. Finally, this study's discussion and conclusion are presented in Section 7.

2 RELATED WORK

In recent years, many researchers have worked on crop disease detection while relying on PlantVillage dataset developed in the USA and Switzerland. Potatoes' diseases vary from region to region due to differences in leaf shapes, varieties, and environmental factors (Baker and Capel, 2011). Researchers have tried to build their custom models compatible with the species they have in their respective countries.

Geetha Ramani and Pandian (Geetharamani and Pandian, 2019) proposed a deep CNN model to differentiate between healthy and unhealthy leaves of multiple crops. The model was trained using the PlantVillage dataset, which included 38 crops with disease leaf images, healthy leaf images, and background images. The focus of the model was not on single potato crop diseases. The model is also trained on specific region datasets in the USA and Switzerland, which failed to detect potato leaf diseases in the Pakistani region. Kamal et al. (Kamal et al., 2019) developed plant leaf disease identification models named Modified MobileNet and Reduced MobileNet using depth-wise separable convolution instead of convolution layer by modifying the MobileNet (Howard et al., 2017). The proposed model was trained on multiple crops of the PlantVillage dataset, where the plant leaf images were collected from a specific world region.

In (Liang et al., 2019), Liang et al. proposed a plant disease diagnosis and severity estimation network based on a residual structure and shuffle units of ResNet50 architecture (He et al., 2016). Khalifa et al. (Khalifa et al., 2021) proposed a CNN model to detect early and late blight diseases and a healthy class. The researchers trained their model on the PlantVillage dataset for specific regions' crops only. In the same direction, Rozaqi and Sunyoto (Rozaqi and Sunyoto, 2020) proposed a CNN model to detect the early blight and late blight diseases of potatoes and a healthy class. They trained the model on the PlantVillage dataset to detect the diseases of a specific region. Similarly, Sanjeev et al. (Sanjeev et al., 2020) proposed a Feed-Forward Neural Network (FFNN) to detect early blight and late blight diseases and healthy

leaves. The proposed method was trained and tested on the PlantVillage dataset. Barman et al. (Barman et al., 2020) proposed a self-build CNN (SBCNN) model to detect early blight, late blight potato leaf diseases, and healthy class. The PlantVillage dataset was also used to train the model for a specific region.

Tiwari et al. (Tiwari et al., 2020) used a pre-trained model VGG19 to extract the features and used multiple classifiers (KNN, SVM and neural network) for classification. The model was also trained on the PlantVillage dataset to detect potato leaves' early blight and late blight disease. Islam et al. (Islam et al., 2017) proposed a segment-based and multi-SVM-based model to detect potato diseases, such as early blight, late blight and healthy leaves. Their method also used the PlantVillage dataset, which needs improved accuracy. Another initiative has been suggested in (Mahum et al., 2023), which proposes a novel framework for potato leaf disease detection utilizing an efficient deep-learning model. This framework leverages advanced convolutional neural networks (CNN) to identify and classify various diseases affecting potato leaves accurately. It is designed to be computationally efficient, making it suitable for real-time applications and deployment on devices with limited processing power. The model is trained and tested on the PlantVillage dataset, showcasing an accuracy of 97.2%.

As illustrated in Table 1, most studies have focused on potato crop diseases, training their models mainly using the PlantVillage dataset. However, these studies typically evaluated their models only within the context of this dataset, without assessing their accuracy on previously unseen data, limiting the generalizability of their findings. Furthermore, none of the cited works addressed the critical aspect of explaining the models' results, resulting in a significant gap in understanding and trust in the models.

3 THEORETICAL BACKGROUND

3.1 Convolutional Block Attention Module (CBAM)

CBAM is an attention mechanism that integrates in series two attention modules, channel attention followed by spatial attention module (Woo et al., 2018) (see Figure 1). The channel attention module was used to generate two feature maps using average and maximum pooling layers from the intermediate layer. Then, both feature maps were input to the shared multilayer perceptron (MLP), and the output feature maps were added before normalizing using the sigmoid

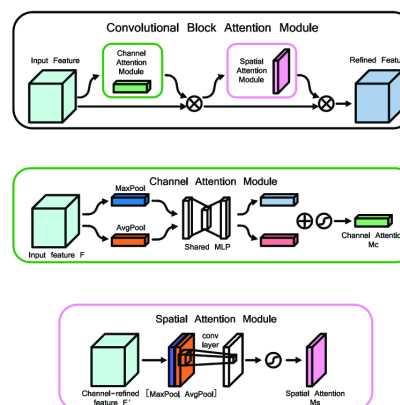


Figure 1: Convolutional block attention module (CBAM) architecture (Woo et al., 2018).

function. The multiplied features between the channel attention module and convolutional layer were applied to the spatial attention module to determine the position of the important features in the image.

3.2 Grad-Cam++

Grad-CAM is a technique for visualizing important regions for available classes using guided propagation. It uses the gradient of any targeted class, passing into the final CNN layer to highlight important regions in the image for prediction (Selvaraju et al., 2017). Grad-CAM computes the gradient concerning the feature map of a convolutional layer by calculating the effect of each area of the image on the final output based on the gradient of the parameter of the final convolutional layer and calculates the degree of influence, represented by a heatmap. Grad-CAM could be applied to several CNN models such as Xception, ResNet and MobileNet without architectural changes or re-training.

To improve object localization and to detect various object instances existing in a single image, Chattopadhyay et al. (Chattopadhyay et al., 2018) proposed Grad-CAM++ technique that aims to visualize predictions from the CNN models better. Grad-CAM++ uses a combined weighted average of the positive partial derivatives of the feature mappings from the final convolutional layer.

4 METHODOLOGY

The main objective is to develop an innovative and effective solution to help farmers quickly and accurately diagnose diseases affecting their crops, thus facilitating the implementation of appropriate management measures. The process we propose for develop-

Table 1: Summary of Related Work.

Study	Methodology	Plant Name	Disease	Dataset	Accuracy (%)
(Geetharamani and Pandian, 2019)	Deep CNN	Multiple (Potato)	Multiple	PlantVillage	96.46
(Kamal et al., 2019)	Modified MobileNet	Multiple (Potato)	Multiple	PlantVillage	98.34
(Liang et al., 2019)	ResNet50	Multiple (Potato)	Multiple	PlantVillage	98
(Khalifa et al., 2021)	CNN	Potato	Early Blight, Late Blight	PlantVillage	98
(Rozaqi and Sunyoto, 2020)	CNN	Potato	Early Blight, Late Blight	PlantVillage	92
(Sanjeev et al., 2020)	FFNN	Potato	Early Blight, Late Blight	PlantVillage	96.5
(Barman et al., 2020)	SBCNN	Potato	Early Blight, Late Blight	PlantVillage	96.75
(Tiwari et al., 2020)	SVM, KNN and Neural-Net	Potato	Early Blight, Late Blight	PlantVillage	97.8
(Mahum et al., 2023)	Advanced CNN	Potato	Early Blight, Late Blight	PlantVillage	97.2

ing the predictive model for potato leaf diseases is illustrated in Figure 2. This process follows the CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology (Wirth and Hipp, 2000) and comprises several key steps, starting with data collection and preparation of images of potato leaves affected by various diseases, together with images of healthy leaves for reference. Secondly, we use transfer learning techniques to exploit pre-trained deep learning models on large image databases by adding the mechanism of spatial attention to these models in order to extract relevant features from images of potato leaves.

These features are then used to form classification models capable of distinguishing healthy leaves from those affected by disease and specifically identifying

the type of disease present. Once the models have been trained, we rigorously evaluate them using appropriate performance measures, such as precision, recall, and F-measure. We also optimize model hyperparameters to improve performance and generalizability.

An important step in the detection model development process is to select the best model trained using transfer learning. This selection is based on two different test bases: one similar to the training data and the other completely different. This ensures that the selected model is robust and generalizable. Finally, the selected model will be used by the mobile application we propose as a practical tool for farmers to monitor and manage the health of their potato crops

effectively and proactively. In addition, the application provides a detailed description of the predicted disease, enabling farmers to make informed decisions on the measures to be taken to prevent the spread and protect their crops. The following sections will detail the various classification and interpretation methodology stages.

4.1 Data Collection

Following the analysis and exploring existing situations, our study is based on the PlantVillage database. This database contains 54,305 images of plant leaves, divided into 38 classes. These images, collected under controlled conditions, represent healthy and diseased plant leaves. Among these images, we find representations of 14 crop species, including apples, blueberries, cherries, grapes, oranges, peaches, bell peppers, potatoes, raspberries, soybeans, squash, strawberries and tomatoes. The dataset covers 17 basic diseases, four bacterial diseases, two diseases caused by fungi (oomycetes), two viral diseases and one mite disease. In addition, for 12 crop species, images of healthy leaves not visibly affected by disease are also included.

Additionally, we collected 100 potato leaf images photographed under practical field scenarios featuring heterogeneous background conditions and varying lighting intensities by downloading potato crop images from popular search engines such as Google. The potato leaf image annotation is carried out by an agronomist who is regarded as an expert.

4.2 Data Understanding

The potato leaf disease dataset is an exhaustive collection of comprehensive images carefully classified into three distinct categories: early blight, late blight and healthy leaf. Each category represents a specific condition affecting potato crops, offering researchers, agricultural researchers, and agricultural experts the opportunity to explore the nuances of disease identification, progression and management. The potato-based data set consists of 2152 image instances and is broken into three distinct classes, each representing a specific condition of potato leaves. Figure 3 illustrates an image example of each class. Figure 4 shows the classes' distribution in the PlantVillage data set. Of the 2152 images, 1000 are of the Early_Blight class, 1000 of the Late_Blight class and 152 of the healthy class.

4.3 Image Pre-Processing

The images used in our study were subjected to several pre-processing operations to make them suitable for input to the CNN models. These operations help improve data quality and facilitate the learning process. In particular, we apply the intensity normalization technique. Secondly, we apply a resizing technique to standardize the shape of all data into 256 x 256 pixels. Thereafter, the images were categorized according to their respective classes. Each image is associated with a label indicating the class to which it belongs using a one-hot encoding technique. Moreover, we used prefetch and cache operations to improve data reading performance during training and evaluation. For the data splitting, we have allocated 70% of our data for the training set, 20% for the validation set and 10% for the test set.

4.4 Data Augmentation and Balancing

Data augmentation artificially enlarges a dataset's size by applying random transformations to existing images. This allows the model to generalize better and reduce overfitting, as different variations of the same image have been seen. Our study uses data augmentation to account for the unbalanced data. In particular, we apply *RandomOverSampler* to deal with the class imbalance problem in our training data.

Table 2: Before and after data balancing.

Balancing	Before	After
Image number	1722	2151

4.5 Transfer CNN Model with Spatial Attention Mechanism

We have selected three convolutional neural network (CNN) models renowned for their performance in image recognition, namely MobileNetV2 (Sandler et al., 2018), DenseNet121 (Huang et al., 2017) and InceptionV3 (Szegedy et al., 2016) to select the best performed model.

We exploit the transfer learning technique to deal with limited data and maximize the efficiency and performance of the resulting model. Transfer learning aims to exploit previously acquired knowledge on large datasets to avoid overfitting while reducing the time and resources needed to train a model from scratch (Chen et al., 2020). Moreover, we added custom layers to the output of each architecture to adapt the model to our particular problem. These additional layers include normalization, regularization, and clas-

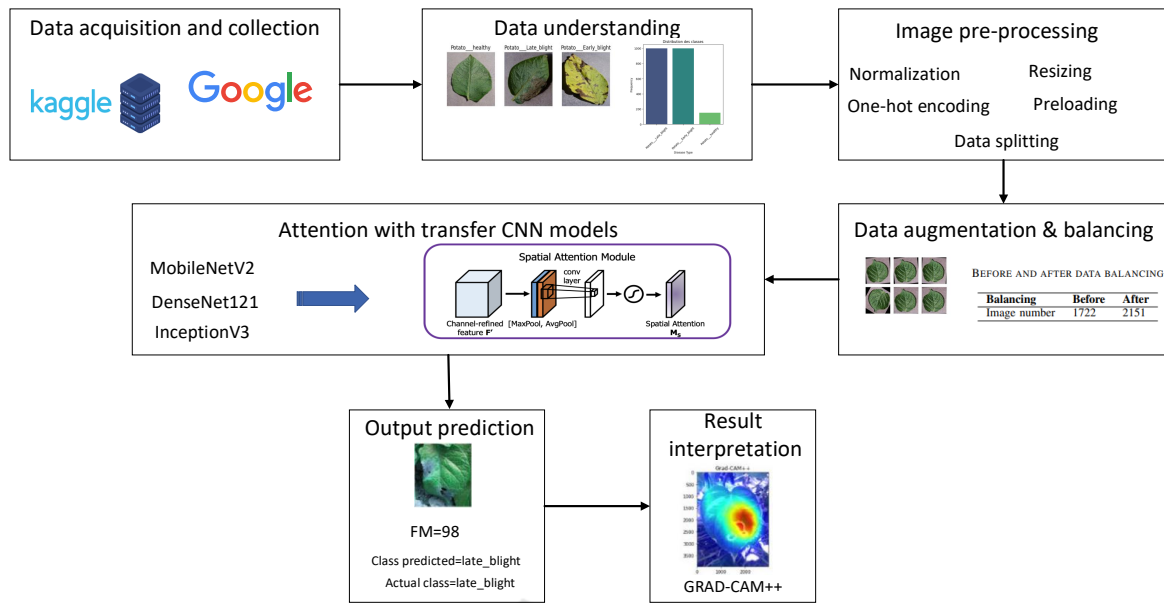


Figure 2: The proposed methodology for potato leaf disease detection.

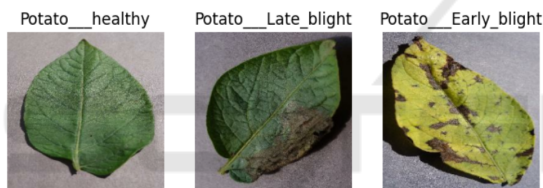


Figure 3: Image example for the three classes.

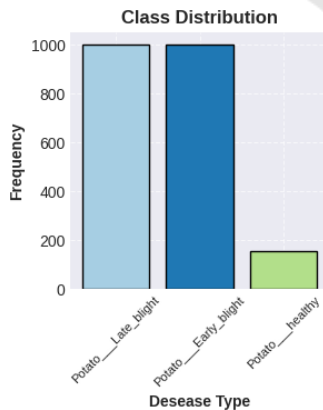


Figure 4: Distribution of classes.

sification operations to extract information specific to the characteristics of potato leaves and use it to predict the corresponding classes.

In addition, we introduced spatial attention mechanisms in each model, allowing the network to focus on the most relevant parts of the image when making a decision. This attention mechanism im-

proves the capability of modeling spatial information in CNNs. Spatial attention is widely used with great success (Woo et al., 2018). The spatial attention focuses on 'where', which is an informative part. To compute the spatial attention, we rely on CBAM attention module (Woo et al., 2018). In particular, we first apply average-pooling and max-pooling operations along the channel axis and concatenate them to generate an efficient feature descriptor. On the concatenated feature descriptor, we use a convolution layer to develop a spatial attention map which encodes where to emphasize or suppress.

4.6 Result Interpretation

Explaining how these "black-box" models predict such disease is required for establishing trust in advanced systems depending on CNN networks. To assist farmers in making decisions, we rely on GRAD-CAM++ as an XAI technique to locate the symptoms responsible for the disease. GRAD-CAM++ is recognized for its precise localization of important image features.

In this study, Grad-CAM++ was implemented on potato leaf images from the test dataset to extract the key regions of the image that contributed most to the model's prediction. In particular, we consider only diseased classes while avoiding healthy classes. The objective is to interpret diseased leaves by highlighting exclusively the defective regions. To do this, Grad-CAM generates a heatmap that high-

lights the critical regions of the potato leaf image based on the gradients. As shown in Figure 9, the Input image (early blight disease) feeds to CNN (Att-MobileNetV2) model to detect the disease, and the grad-CAM model is applied to the last convolution layer of Att-MobileNetV2 for disease visualisation.

5 EXPERIMENT

5.1 Experiment Protocol

To identify the best model for potato disease classification in terms of performance, we compare the three CNN architectures MobileNetV2, DenseNet121 and InceptionV3 to find the best one with optimal hyperparameters. For doing so, the modified MobileNetV2, DenseNet121 and InceptionV3 are re-trained. These models are gradually enhanced by applying several techniques, such as data balancing and augmentation. Moreover, during the training phase, the callback tool is considered for early stopping to prevent the model from over-fitting. In particular, we use callbacks from *Keras*. By using hyper-parameter tuning from *Keras-tuner*, we optimize the number of parameters to the maximum. It should be noted that the performance of the different generated models' classification was then tested on over 272 test images, including 172 from PlantVillage and 100 from Google.

To evaluate the performance of the different CNN models, we rely on the confusion matrix for the test data and well-known performance metrics such as accuracy and F-measure.

5.2 Experimental Results and Discussion

5.2.1 Hyperparameter Tuning Impact

During the training of the different models, hyperparameter tuning and optimization of the model are applied. This technique aims to find the best hyperparameters, such as learning rates, batch sizes, and regularization techniques, to optimize the network's performance. This ensures that the network effectively learns and represents the most relevant features for accurate object recognition and classification. Table 3 presents the basic models' accuracy (without attention mechanism). The first column presents the performance of the models where each hyperparameter is selected manually. As for the second column, it plots the performance of the models trained with hyperparameter tuning based on different ranges for

each hyperparameter. The result illustrates the positive impact of hyperparameter tuning. It enhances the three models' accuracy. For instance, it improves the accuracy by 2.99% for the MobileNetV2. Hence, we consider hyper-parameter tuning for the rest of the experiments.

Table 3: Hyperparameter tuning impact assessment based on accuracy metric.

	Manual	Hyperparameter tuning
<i>MobileNetV2</i>	0.95	0.97
<i>DenseNet</i>	0.97	0.98
<i>InceptionV3</i>	0.95	0.95

5.2.2 Transfer CNN Models Comparison

This experiment concerns the performance results from the transfer learning-based models. The results indicate the excellent performance (0.98) of the proposed DenseNet121 compared to the other transfer learning-based models, MobileNetV2 and InceptionV3. However, we cannot deny the high accuracy of the MobileNetV2 model since it achieves 0.97 of accuracy.

5.2.3 Spatial Attention Assessment

Table 4: Models' performance.

Model	Accuracy	F-measure
MobileNetV2	0.97	0.92
DenseNet121	0.98	0.93
InceptionV3	0.95	0.87
Att-MobileNetV2	0.99	0.98
Att-DenseNet121	0.98	0.93
Att-InceptionV3	0.94	0.87

The third experiment aims to illustrate the impact of spatial attention on potato disease classification. Table 4 shows the performance of the attention-based models. The table clearly highlights the significant improvement in model accuracy achieved by integrating the spatial attention module. In particular, the Att-MobileNetV2 surpasses the performance of MobileNetV2 by 2% of accuracy and 6% of F-measure. Moreover, the confusion matrices presented in Figures 5, 6, and 7 clearly indicate an overall performance using Att-MobileNetV2. However, it is not the case for the InceptionV3 architecture, where the Att-InceptionV3 decreases performance by 1% of accuracy.

Table 5 presents the performance metrics for Att-MobileNetV2, Att-DenseNet121, and Att-InceptionV3. The table reveals that Att-MobileNetV2 demonstrates exceptionally high classification per-

Table 5: Performance per class of attention-based models.

Model	Class	Accuracy	F-measure
Att-MobileNetV2	Early_blight	1.00	1.00
	Late_blight	1.00	0.99
	Healthy	0.88	0.94
Att-DenseNet121	Early_blight	0.99	0.99
	Late_blight	0.98	0.96
	Healthy	0.78	0.85
Att-InceptionV3	Early_blight	0.97	0.97
	Late_blight	0.98	0.92
	Healthy	0.56	0.71

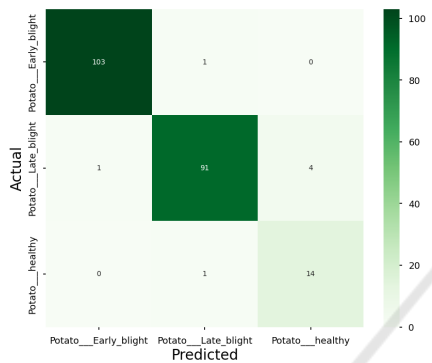


Figure 5: Confusion matrix on the test data using Att-DenseNet.

formance, indicating the model’s strong capability in accurately identifying leaf diseases with minimal errors. Notably, the model achieves a 100% accuracy rate in classifying both early and late blight classes. However, for the healthy class, the classification error increases across all three models, likely due to the limited number of only 152 images available in the PlantVillage dataset for this class.

To sum up, the experiments reveal the outperformance of Att-MobileNetV2 compared to the other experimented models and also compared to related works (see Table 1).

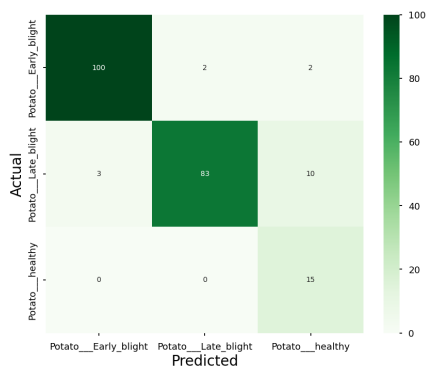


Figure 6: Confusion matrix on the test data using Att-Inceptionv3.

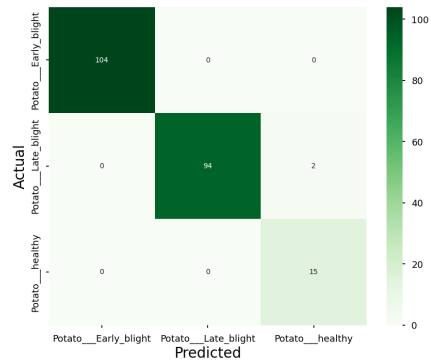


Figure 7: Confusion matrix on the test data using Att-MobileNetV2.

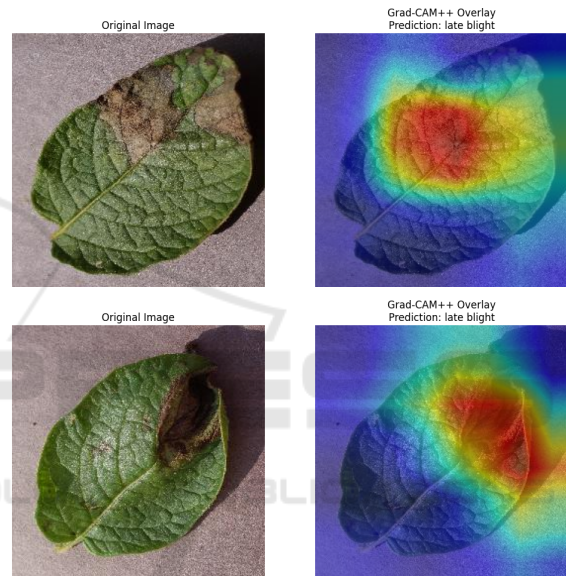


Figure 8: GradCam++ output for late-blight disease prediction based on Att-MobileNetV2 model

5.2.4 Result Interpretation Assessment

The Grad-CAM++ outputs have been illustrated in Figures 8 and 9 for visual explanations. Red indicates higher attention values, and blue indicates lower attention values. These figures clearly show the discriminative regions of different images. Hence, two conclusions could be drawn: the model effectively locates the region of the disease, which increases the trust and confidence of the farmers in the proposed model predictions.

5.3 Proof of Concept: Mobile Application

We developed a mobile application with Flutter (Flu, online) as a proof-of-concept. Flutter has been used because of its cross-platform capabilities, allowing us

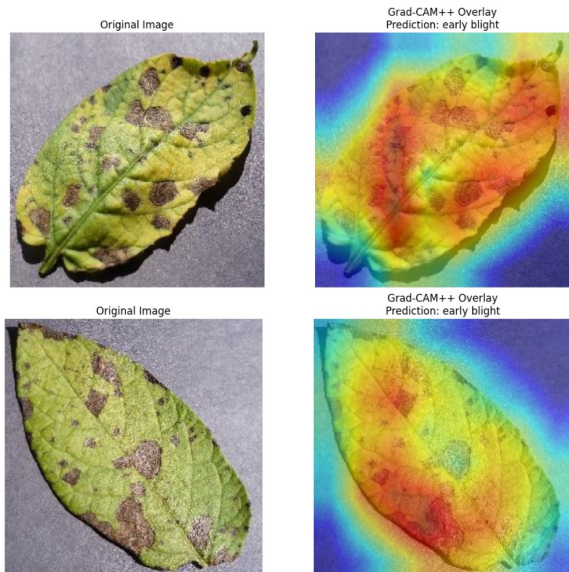


Figure 9: GradCam++ output for early-blight disease prediction based on Att-MobileNetV2 model

to provide a uniform and responsive user experience on Android and iOS devices with a single codebase.

The developed mobile application aims to empower farmers by allowing them to take images of potato leaves and seek analysis straight from their mobile devices (see Figure 10). The analysis is carried out using the Att-MobileNetV2 classification model, which is hosted on a private server. The server processes the images given by the application, and communication between the Flutter application and the server is enabled using RESTful API queries. An example of the application running is depicted in Figure 11.

6 DISCUSSION

The experimental findings clearly show Att-MobileNetV2's remarkable classification performance, emphasizing the model's strong capacity to correctly identify leaf diseases with minimal errors, outperforming state-of-the-art models. However, it is crucial to discuss the limitations of our model. It was primarily trained on potato leaf images captured in controlled environments, specifically within the PlantVillage dataset. Although we tested the model on 100 images sourced from Google in uncontrolled environments, this effort needs to be extended to a larger set of images. Images taken in uncontrolled environments where variations in lighting, angles, and background noise are more pronounced can significantly challenge the model's disease identification and classification accuracy. This underscores the im-

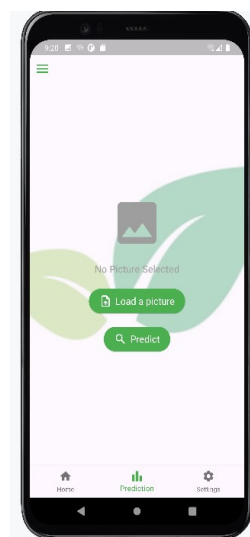


Figure 10: Homepage of the application.

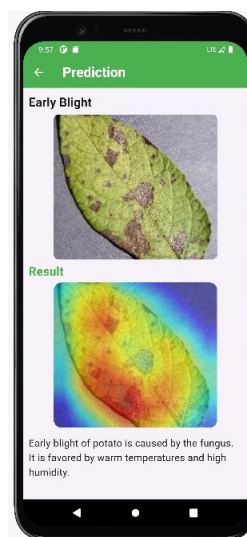


Figure 11: Example of application execution.

portance of adapting our methodology to account for such conditions and possibly incorporating additional processes, such as image segmentation, to enhance performance in real-world scenarios.

7 CONCLUSION

The objective of this study was to propose a solution for potato disease detection using deep learning and transfer learning techniques. Using a rigorous methodology, we have built three convolutional neural network models, integrating transfer learning and the attention mechanism using hyperparameter tuning to select the optimal hyperparameters. In particular, three models, Att-MobileNetV2, Att-DenseNet and Att-InceptionV3 were evaluated based on the PlantVillage reference database and images collected from Google to support different image quality, lighting conditions and the presence of overlapping symptoms. The Att-MobileNetV2 model was selected for the prediction of potato leaf diseases thanks to its superior performance and its ability to generalize on unseen data, reinforcing our confidence in its practical use.

Despite the promising results, this work requires improvement and extension. Indeed, it is important to consider another dataset that contains differences in leaf shapes, sizes, colours, lighting conditions, and photo backgrounds to enhance the model's performance.

REFERENCES

- Flutter - build apps for any screen. <https://flutter.dev/>. (Accessed on 09/04/2024).
- Afzaal, H., Farooque, A. A., Schumann, A. W., Hussain, N., McKenzie-Gopsill, A., Esau, T., Abbas, F., and Acharya, B. (2021). Detection of a potato disease (early blight) using artificial intelligence. *Remote Sensing*, 13(3).
- Bahar, N. H., Lo, M., Sanjaya, M., Van Vianen, J., Alexander, P., Ickowitz, A., and Sunderland, T. (2020). Meeting the food security challenge for nine billion people in 2050: What impact on forests? *Global Environmental Change*, 62:102056.
- Baker, N. and Capel, P. (2011). Environmental factors that influence the location of crop agriculture in the conterminous united states. Technical report, US Department of the Interior, US Geological Survey, Reston, VA, USA.
- Barman, U., Sahu, D., Barman, G., and Das, J. (2020). Comparative assessment of deep learning to detect the leaf diseases of potato based on data augmentation. In *2020 International Conference on Computational Performance Evaluation (ComPE)*, pages 682–687.
- Chattopadhyay, A., Sarkar, A., Howlader, P., and Balasubramanian, V. N. (2018). Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks. In *2018 IEEE winter conference on applications of computer vision (WACV)*, pages 839–847. IEEE.
- Chen, J., Chen, J., Zhang, D., Sun, Y., and Nanekaran, Y. A. (2020). Using deep transfer learning for image-based plant disease identification. *Computers and electronics in agriculture*, 173:105393.
- Geetharamani, G. and Pandian, A. (2019). Identification of plant leaf diseases using a nine-layer deep convolutional neural network. *Computers & Electrical Engineering*, 76:323–338.
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.
- Howard, A., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*.
- Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. (2017). Densely connected convolutional networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708.
- Islam, M., Dinh, A., Wahid, K., and Bhowmik, P. (2017). Detection of potato diseases using image segmentation and multiclass support vector machine. In *2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*, pages 1–4.
- Kamal, K., Yin, Z., Wu, M., and Wu, Z. (2019). Depthwise separable convolution architectures for plant disease classification. *Computers and Electronics in Agriculture*, 165:104948.
- Khalifa, N., Taha, M., Abou El-Maged, L., and Hassanien, A. (2021). Artificial intelligence in potato leaf disease classification: A deep learning approach. *Machine Learning and Big Data Analytics Paradigms: Analysis, Applications and Challenges*, pages 63–79.
- Koné, B. A. T., Bouaziz, B., Grati, R., and Boukadi, K. (2023a). Boruta-attlstm: A novel deep learning architecture for soil moisture prediction. In *International Conference on Intelligent Systems and Pattern Recognition*, pages 234–246. Springer.
- Koné, B. A. T., Grati, R., Bouaziz, B., and Boukadi, K. (2023b). A new long short-term memory based approach for soil moisture prediction. *Journal of Ambient Intelligence and Smart Environments*, (Preprint):1–14.
- Liang, Q., Xiang, S., Hu, Y., Coppola, G., Zhang, D., and Sun, W. (2019). Pd2se-net: Computer-assisted plant disease diagnosis and severity estimation network. *Computers and Electronics in Agriculture*, 157:518–529.
- Mahum, R., Munir, H., Mughal, Z.-U.-N., Awais, M., Sher Khan, F., Saqlain, M., Mahamad, S., and Tlili, I. (2023). A novel framework for potato leaf disease detection using an efficient deep learning model. *Human and Ecological Risk Assessment: An International Journal*, 29(2):303–326.
- Rozaqi, A. and Sunyoto, A. (2020). Identification of disease in potato leaves using convolutional neural network (cnn) algorithm. In *2020 3rd International Conference on Information and Communications Technology (ICOIACT)*, pages 72–76.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., and Chen, L.-C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4510–4520.
- Sanjeev, K., Gupta, N., Jeberson, W., and Paswan, S. (2020). Early prediction of potato leaf diseases using ann classifier. *Orient. J. Comput. Sci. Technol.*, 13:2–4.
- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 618–626.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826.
- Tiwari, D., Ashish, M., Gangwar, N., Sharma, A., Patel, S., and Bhardwaj, S. (2020). Potato leaf diseases detection using deep learning. In *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, pages 461–466.
- Wirth, R. and Hipp, J. (2000). Crisp-dm: Towards a standard process model for data mining. In *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining*, volume 1, pages 29–39. Manchester.
- Woo, S., Park, J., Lee, J.-Y., and Kweon, I. S. (2018). Cbam: Convolutional block attention module. In *Proceedings of the European conference on computer vision (ECCV)*, pages 3–19.