JASMINE
Journal of Intelligent Systems and Machine Learning

Articles in Press

CNN-Based Deep Learning Utilization Model for Identification of Crystal Guava Leaf Diseases

Ali Mustadji[1], Arie Qur'ania[1], and Asep Saepulrohman[1,*]
[1] Computer Science, Faculty of Mathematics and Natural Sciences, Pakuan University, Bogor, Indonesia

*Author to whom any correspondence should be addressed.
E-mail: asepspl@unpak.ac.id

Received: October 14, 2025
Accepted for publication: January 21, 2026

ABSTRACT
Identification of plant diseases is a crucial step in maintaining plant health and preventing economic losses due to decreased productivity. This research aims to develop an intelligent system capable of identifying diseases in Crystal Guava (Psidium guajava L.) plants using a Convolutional Neural Network (CNN) method. The method combines digital image processing techniques with machine learning to classify Crystal Guava leaf images into two categories: healthy and diseased. The implemented CNN architecture is based on the Xception model, known for its superior performance in image classification tasks. The dataset used consisted of 1,500 Crystal Guava leaf images, including both healthy and diseased leaves. Test results showed that the developed system achieved 94% accuracy in identifying diseases in Crystal Guava plants, surpassing the performance of other architectures such as VGG16 and InceptionV3. This high accuracy demonstrates the model's ability to recognize complex features in Crystal Guava leaf images. These findings contribute to the development of artificial intelligence-based diagnostic tools for early detection of plant diseases. The proposed system is expected to assist farmers, agricultural researchers, and policymakers in making informed decisions to improve the productivity and health of Crystal Guava plants.

Keywords: Crystal Guava, CNN, Xception, image processing, plant disease detection

I. Introduction

In large-scale guava plantations, identifying disease symptoms is crucial to protect plants from damaging disease attacks. Farmers must thoroughly inspect each guava plant to prevent the spread of disease to healthy plants. Infected plants are then treated by pruning infected branches and leaves. Although guava has a certain degree of disease resistance, it remains susceptible to various infections. Guavas (Psidium guajava L.) are tropical plants native to Indonesia, having been introduced through the Taiwanese Technical Mission in Indonesia in the 1990s. This plant has great potential for development and cultivation in Indonesia due to its high disease resistance, adaptability to local climate and soil conditions, and favorable weather patterns. [1]. However, one of the main problems faced by crystal guava plantations is pest and disease attacks that can hamper plant productivity and disrupt its normal physiological processes.

Diagnosis of plant diseases is usually done based on observations of symptoms and changes in plant morphology Diagnosis of plant diseases has traditionally been carried out through direct observation of visible symptoms and changes in plant morphology [2]. However, this approach is often time-consuming and highly dependent on human expertise. With recent advances in deep learning, artificial neural networks can now be utilized to improve the accuracy and efficiency of plant disease classification. Among these methods, Convolutional Neural Networks (CNNs) have shown outstanding performance in various image classification tasks due to their capability to automatically extract discriminative features from images. Consequently, the application of CNNs in plant disease diagnosis has attracted significant attention, as these models are able to detect subtle visual patterns that are difficult to identify with the human eye. By training CNN models on large-scale image datasets, diagnostic accuracy can be significantly improved while reducing the time required for manual inspection. Moreover, CNN-based systems enable early disease detection, allowing farmers to take timely preventive actions and minimize crop losses [3]-[5].

In this study, the image dataset was collected directly from crystal guava plantations located in the Bogor region, Indonesia. The dataset consists of a total of 1,804 leaf images, acquired during the period 2023–2024. The images are categorized into three classes, namely healthy leaves, red rust disease, and leaf spot disease. All images were captured using a smartphone camera at a distance of less than 30 cm, focusing on the surface of the leaves to ensure clear visualization of disease characteristics. The dataset was subsequently divided into training, validation, and testing subsets to support robust model evaluation. Representative sample images from each class are provided to illustrate visual differences among the categories.

However, one of the main challenges in developing a reliable plant disease diagnosis system is the availability of high-quality training data. The dataset must be large enough to cover a wide range of diseases, weather conditions, and soil types that can affect plant growth and development. Furthermore, the images used for training must be accurately labeled and annotated to accurately diagnose the disease. To address these challenges, researchers have proposed several solutions, including the use of transfer learning and data augmentation. Transfer learning involves initially training a CNN model on a large dataset, then fine-tuning it using a smaller, more specific dataset for plant disease diagnosis [6]-[12]. Meanwhile, data augmentation is performed by generating new images through transformations such as rotation, flipping, and cropping.

In this research, we aim to develop a disease identification system for guava plants based on image analysis using a Convolutional Neural Network (CNN) with the GoogleNet Xception architecture. Our goal is to achieve high accuracy in disease classification, which can provide valuable insights for farmers and researchers. We will use a large dataset of guava leaf images collected from various locations and weather conditions to train a CNN model. The developed system consists of two main components: image processing and disease diagnosis. The image processing component involves preprocessing by applying filters and transformations to improve image quality and reduce noise. The disease diagnosis component will then use the processed images to classify them into various disease categories.

We estimate that the developed system will improve the accuracy of disease diagnosis in crystal guava plants and provide significant benefits for both farmers and researchers. We also hope that this research can contribute to the development of more reliable and efficient plant disease diagnosis systems in the future. The impact of this research will be significant because it can help farmers detect diseases early and prevent their spread, thereby increasing crop productivity and reducing economic losses due to diseases. In addition, this system can also support sustainable agricultural practices by reducing the use of pesticides and chemical fertilizers. This study aims to develop a reliable system for identifying diseases in crystal guava plants based on image analysis using CNN with the GoogleNet Xception architecture [13]-[15]. We hope this system can improve the accuracy of disease diagnosis and provide valuable insights for farmers and researchers.

II. Material and Methods

A. Crystal Guava Plant Disease

The highest risk of reducing Crystal Guava production includes pests and plant diseases. Pests commonly found in Crystal Guava plantations are Attacus atlas, Setora nitens, Trabala sp., hanging caterpillars, leaf-eating caterpillars, and twig borers (Lepidoptera) [16]-[20]. These insects also include Paracoccus marginatus, fruit flies, shield bugs (Hemiptera), grasshoppers (Orthoptera), Bactrocera carambolae (Diptera), and Carpophilus sp. (Coleoptera). In addition to pests, various plant diseases are also found in Crystal Guava plantations, including red rust (Cephaleuros sp.), fruit canker caused by Pestalotia (Pestalotia sp.), anthracnose (Colletotrichum sp.), and white downy mildew (Triposporium sp.). These diseases are a significant threat to Crystal Guava production and require careful management to prevent damage.

B. RGB Colors

RGB color is a color model used in feature extraction from color images using wavelet packet decomposition. This method aims to create unique and invariant features for each image to improve image recognition systems. Experimental results show that this method successfully generates unique features for each image, which can be used for image identification. This method is efficient and flexible, allowing various levels of decomposition to extract features. These results also show that the extracted features remain consistent even after image rotation.

C. *Convolution Neural Network (CNN)*

A Convolutional Neural Network (CNN) is a type of neural network that has gained popularity in areas such as image classification, object detection, and segmentation due to its ability to efficiently learn abstract features. CNNs have revolutionized the fields of computer vision and natural language processing, leading to the development of advanced models and applications. A key component of CNNs is the convolutional layer, which plays a crucial role in extracting features from input data. CNNs have been successful in tasks such as image analysis, text recognition, action recognition, and text-image generation. Various CNN architectures, including LeNet-5, AlexNet, VGGNet, ResNet, GoogLeNet, and DenseNet, have been developed, each with unique features and applications. To train CNNs effectively, proper data preprocessing, parameter initialization, regularization techniques, and optimization methods are crucial to avoid problems such as overfitting and slow convergence [21].

1) *GoogleNet*

The Inception-v1 architecture inherits the framework structure from LeNet and AlexNet, with some modifications to the network depth and width. The specific architecture of Inception-v1 is depicted as a neural network diagram, where type represents the type of layer, depth represents the number of layers, pooling represents the pooling operation, fc represents the fully connected layer, and softmax represents the output layer. The final output will be a probability. Since then, GoogLeNet has been continuously improved, leading to the development of the Inception-v2, Inception-v3, and Inception-v4 architectures successively. Among them, Inception-v2 mainly adds a batch normalization layer, Inception-v3 mainly replaces the two-dimensional convolutional kernel with a one-dimensional convolutional kernel, and Inception-v4 mainly incorporates the concept of residual networks [22].

2) *Xception*

Xception, which stands for extreme inception, is a CNN architecture based entirely on separable convolutional layers. It has 36 convolutional layers, divided into 14 modules. Except for the first and last modules, all other modules have linear residual connections. The Xception architecture relies heavily on depthwise separable convolutional layers [23].

D. *Dataset Description and Class Distribution*

The dataset employed in this study was derived from field data as documented in the research report by Ali Mustadji (2024). The image dataset consists of Crystal Guava leaf samples collected directly from guava plantations located in the Bogor region, Indonesia. A total of 1,804 leaf images were acquired over the period 2023–2024, ensuring coverage of multiple growing seasons and natural disease variations. All images were captured using a smartphone camera at a distance of less than 30 cm, with a focus on the leaf surface area to clearly represent disease symptoms. This acquisition method follows practical conditions commonly encountered by farmers and supports real-world applicability of the proposed system. Prior to model training, all images were reviewed and labeled according to observable disease characteristics as described in the reference dataset documentation. The dataset is categorized into three disease classes, namely: Healthy Leaves, Red Rust Disease, and Leaf Spot Disease. These categories were selected based on their prevalence in Crystal Guava plantations and their distinct visual features, including color changes, rust-like textures, and irregular spotting patterns. Sample images for each class are provided in the dataset to illustrate inter-class visual differences. To ensure reliable training and evaluation, the dataset was divided into training, validation, and testing subsets following standard deep learning practice. The training set was used to learn model parameters, the validation set to tune hyperparameters and monitor

Table 1. Dataset Composition and Class Distribution.

Class	Training Set	Validation Set	Testing Set	Total
Healthy Leaf	420	90	90	600
Red Rust Disease	421	91	90	602
Leaf Spot Disease	422	90	90	602
Total	1,263	271	270	1,804

overfitting, and the testing set to evaluate final model performance on unseen data. The class distribution across these subsets was maintained proportionally to preserve the original data characteristics and reduce class imbalance effects. The dataset composition is summarized in Table 1.

E. Transfer Learning

Transfer learning is a process in which a pre-trained model is first trained on a similar but different task and then used to tackle a new task. The concept of transfer learning involves utilizing most of the layers from the pre-trained model, while only a few final layers are retrained for a new, different task. This technique aims to facilitate the learning of better prediction functions by leveraging the knowledge and features learned during the initial training process [24].

F. Evaluation

There are several evaluation metrics used in this study: Precision is the calculation of the classification accuracy ratio of the number of positively labeled data or minority class data that is actually the positive class, the mathematical formula is given in Equation 1; Recall (sensitivity) is a calculation of the ratio of classification accuracy to the number of positive data correctly identified as the positive class, as formulated In Equation 2; The F1-score is a metric that quantifies the harmonic mean of the precision and recall ratios, thereby representing the equilibrium of the classification performance for data belonging to the minority class, as shown in Equation 3; Accuracy is the ratio of correct positive and negative predictions to total data, the mathematical formula is given in Equation 4.

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

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

$$F1 - score = 2 \times \frac{Recall \times Precision}{Recall + Precision} \tag{3}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

III. Results

The results of this study demonstrate that the application of a Convolutional Neural Network (CNN) method to Crystal Guava leaf images is capable of identifying plant conditions with a high degree of accuracy. The developed CNN model, based on the Xception architecture, successfully learned patterns and visual features from leaf images of both healthy and diseased plants. Image preprocessing, including normalization, data augmentation, and image quality enhancement, significantly improved the model's classification performance. Test results indicate that the developed system achieved 94% accuracy, surpassing benchmark architectures such as VGG16 and InceptionV3. This success demonstrates the effectiveness of a deep learning-based approach in automatically diagnosing plant diseases. Furthermore, the resulting model can serve as the basis for developing a decision support system for farmers and researchers to more efficiently monitor Crystal Guava plant health.

The research process begins with the collection of a dataset consisting of Crystal Guava leaf images obtained from plantations in the Bogor region. The collected images are then subjected to data labeling, where each image is assigned to one of three predefined classes: healthy leaves, red rust disease, and leaf spot disease. After labeling, the dataset is divided into three subsets: training, validation, and testing data. The training and validation datasets undergo an image preprocessing stage, which includes image resizing to a uniform resolution, batch generation, and data augmentation. These preprocessing steps are applied to improve data quality, increase data diversity, and enhance the robustness of the model against overfitting. The preprocessed training data are then used in the model training stage, where a Convolutional Neural Network (CNN) with transfer learning based on the Xception architecture is employed. During training, the model learns discriminative features from the leaf images using pre-trained ImageNet weights. The validation dataset is used to monitor model performance and to determine the best model based on accuracy and loss values. Once the optimal model is selected, it is evaluated using the testing dataset, which contains unseen data. This evaluation provides an unbiased assessment of the model's generalization capability. The final outputs include classification results, along with performance metrics such as accuracy and loss. The selected model is proposed as the final system and is designed for future integration into a web-based application to support automatic disease diagnosis in Crystal Guava plants.

In this study, the Deep Learning-based Xception architecture is used for pre-training in transfer learning to extract features with ImageNet weights and CNN as classifiers. The Xception system is identical to Inception (GoogleNet), where Inception is replaced by depth-separated convolutional layers. Specifically, the Xception structure is developed based on a linear stack of depth-separated convolutional layers (36 convolutional layers) with residual linear connections. The following is the transfer learning model of the Xception Architecture in Figure 1.

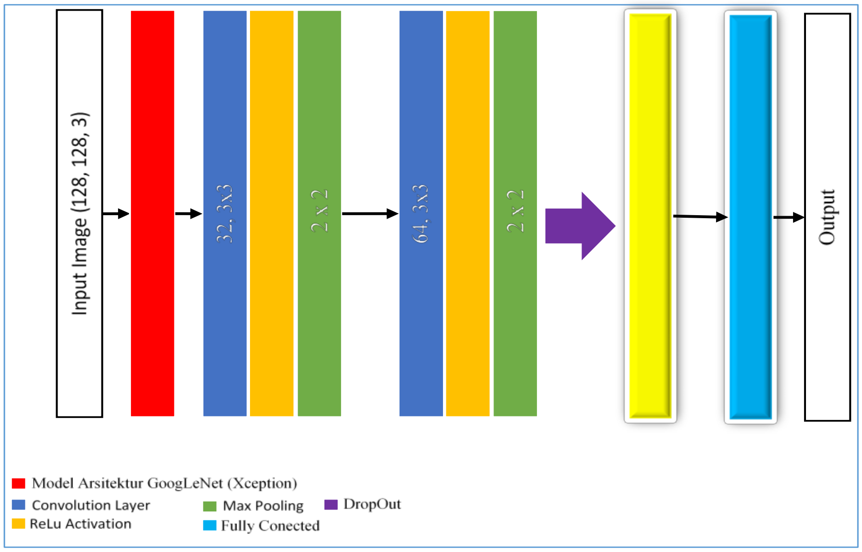

Figure 1. Transfer learning model of Xception-CNN Architecture

This study proposes a CNN architecture with two convolutional layers and a max-pooling filter, followed by image classification using transfer learning with the GoogleNet Xception model as pre-training with ImageNet weights. In the CNN architecture, the most significant component is the convolutional layer, as formulated in Equation 5.

$$s(t) = (x \times w) = \sum_{y=0}^{Row} \left(\sum_{x=0}^{column} x \times w \right) \tag{5}$$

This layer consists of a series of convolutional filters (also known as kernels). The input image, represented as an N-dimensional metric, is convolved with these filters to produce an output feature map, where the process of this convolution layer is depicted in Figure 2.

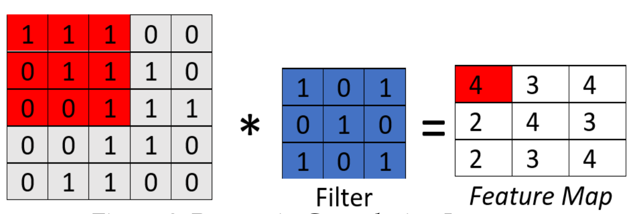

Figure 2. Process in Convolution Layer

A simple and widely used max-pooling method in CNNs, it does not require any parameter tuning. Max-pooling, as formulated in Equation 6, is a mechanism that optimizes the spatial size of feature maps while providing translational invariance to the network. This is achieved by selecting the largest value in the feature map, typically in a kxk neighborhood. The max-pooling technique identifies the largest element in each pooling region. A pooling layer is used to reduce the spatial dimension of the feature map, thereby reducing model complexity and parameters.

$$MaxPooling(F)_{ij} = max(f_{2i,2j}, f_{2i,2j+1}, f_{2i+1,2j}, f_{2i+1,2j+1}) \quad (6)$$

The Rectified Linear Unit (ReLU) layer is the most commonly used function in CNNs. All inputs are transformed to positive values. The computational cost of ReLU is relatively low compared to other functions. Mathematically, the ReLU layer maps all negative values from the previous layer to 0, as formulated in Equation 7. This helps maintain the mathematical stability of CNNs by preventing learned weights from getting stuck near 0 or growing unbounded.

$$ReLu\ (x) = max(0, x) \quad (7)$$

The Fully Connected (FC) layer in a network is a crucial component where each neuron is connected to all neurons in the previous layer, known as a fully connected approach. This layer serves as a classifier for CNNs. It follows the basic architecture of a multi-layer perceptron, which is a type of feed-forward artificial neural network (ANN). The input to the FC layer comes from the last convolutional or pooling layer. This input is a vector, created by flattening the feature maps obtained after applying the flattening operation. The mathematical formula is given in Equation 8.

$$\sum_{i=1}^{n} I_i \times x_{ij} = J_i \quad (8)$$

Table 2. In-depth Analysis of CNN Model Performance for Crystal Guava Leaf Disease Classification.

Analysis Aspect	**Xception (Proposed Model)**	**VGG16**	**InceptionV3**	**Impact on Performance**
Core Architecture	Depthwise separable convolution with residual connections	Standard stacked convolution layers	Inception modules with multi-scale convolutions	Xception separates spatial and channel-wise feature extraction, improving representational efficiency
Parameter Efficiency	Fewer effective parameters with high feature representation capability	Very large number of parameters (~138M)	Moderate number of parameters (~23M)	Reduced overfitting risk on medium-sized datasets
Fine-Grained Feature Extraction	Excellent at capturing subtle texture patterns and leaf surface details	Limited ability to capture fine local features	Good multi-scale feature extraction but complex	Leaf diseases often exhibit subtle visual patterns best captured by Xception
Suitability for Leaf Image Characteristics	Highly adaptive to organic textures and color variations	More suitable for large, high-contrast objects	Suitable for complex multi-scale objects	Xception better models natural variations on leaf surfaces
Training Stability	Highly stable with transfer learning using ImageNet weights	Prone to overfitting without strong regularization	Stable but requires extensive hyperparameter tuning	Faster and more stable convergence during training
Sensitivity to Data Augmentation	High and effective	Relatively low	Moderate	Augmentation significantly improves Xception generalization
Generalization Capability	High performance on unseen test data	Limited for small–medium datasets	Good but parameter-sensitive	Consistent performance in real-world scenarios
Test Accuracy	94%	Lower than Xception	Lower than Xception	Demonstrates clear superiority of the proposed model
Computational Complexity	Computationally efficient	High computational cost	Moderate to high	Suitable for real-time and web-based deployment
Practical Deployment Readiness	Highly suitable for real-time applications	Less suitable due to complexity	Moderately suitable	Supports implementation in decision support systems

The Softmax function is a commonly used function for calculating probabilities, typically used for multiclass classification tasks. The output of the Softmax function ranges from 0 to 1 and sums to 1 when all elements are summed. This function is typically applied at the end of a fully-connected layer in a CNN to generate probability values for an object belonging to each class. The Softmax formula can be represented by calculating the exponential of the input value i, divided by the sum of all exponential values, as shown in Equation 9.

$$softmax(t) = \frac{\exp^{netinput_i}}{\sum_{j=i}^{n} \exp^{netinput_j}} \tag{9}$$

Table 2 presents a more in-depth analytical table and discussion, clearly explaining why the Xception model outperforms the other compared models (VGG16 and InceptionV3).

IV. Discussion

The Transfer Learning-based CNN model proposed in this study was compared with several other popular CNN architectures to assess its effectiveness and performance in classifying diseases in Crystal Guava plants. The training process was conducted using fixed hyperparameters: 50 epochs, a batch size of 32, the Adam optimizer with a learning rate of 0.0001, and a categorical crossentropy loss function. These parameters were selected to optimize the model's learning process and ensure convergence to a stable solution without overfitting.

The training results showed that the Transfer Learning model with the Xception architecture performed better than other CNN models such as VGG16 and InceptionV3. The Xception architecture is capable of extracting complex features from leaf images thanks to its use of depthwise separable convolution, which efficiently separates the spatial and channel feature extraction processes. This approach results in less computational effort while maintaining robust feature representation. Furthermore, the use of pre-trained weights from ImageNet during the transfer learning stage has been shown to accelerate the training process and improve model accuracy. This is because the model already has a prior understanding of the general features of natural images, which are then fine-tuned to the specific characteristics of Crystal Guava leaf images.

Figure 3. Results of image data pre-processing.

The proposed model achieved 94% accuracy, which is higher than a conventional CNN model without transfer learning. This performance demonstrates that the transfer learning approach can leverage knowledge gained from a large dataset (ImageNet) for application to new domains with relatively smaller datasets. Overall, the results of this study confirm that the Xception architecture with transfer learning is an effective approach for image-based plant disease detection. The advantages of this model lie not only in its high accuracy but also in its computational efficiency and training stability. These findings demonstrate the significant potential of utilizing deep learning in digital agriculture, particularly for automated and rapid plant disease diagnosis. Data augmentation aims to enable the machine to learn and recognize from a variety of different images while simultaneously leveraging data to achieve optimal system model performance. The results after the preprocessing stage are shown in Figure 3.

The evaluation process presents the results of previously trained models using training data and then interprets their performance using the confusion matrix method. The visualized evaluation results of the training, validation, and testing models can be seen in Figure 4. Figure 5 illustrates the training and validation performance of the proposed Xception-based CNN model over 50 epochs, evaluated using loss and accuracy metrics.

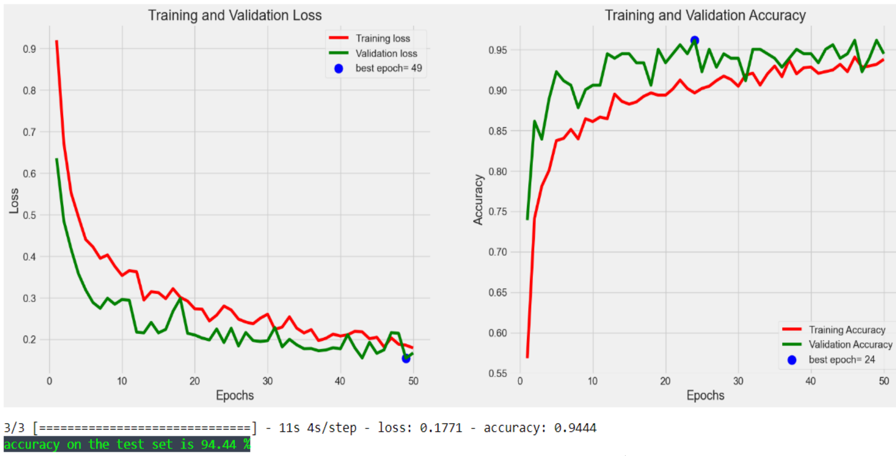

Figure 4. CNN+Exception Model.

1) Training and Validation Loss

The left figure shows the trend of training and validation loss during the learning process. At the initial epochs, both training and validation loss values decrease sharply, indicating that the model is able to quickly learn fundamental visual features from the Crystal Guava leaf images. As training progresses, the loss curves gradually stabilize and converge toward lower values. Importantly, the validation loss consistently follows the training loss without significant divergence. This behavior indicates that the model does not suffer from severe overfitting and demonstrates good generalization capability. The lowest validation loss is achieved around epoch 49, which is marked as the best epoch, suggesting that the model reaches optimal learning near the end of training.

2) Training and Validation Accuracy

The right figure presents the training and validation accuracy curves. Both curves show a rapid increase in accuracy during the early epochs, reflecting effective feature extraction and learning by the CNN. The validation accuracy remains consistently high and, in several epochs, slightly exceeds the training accuracy. This phenomenon commonly occurs when regularization techniques and data augmentation are applied, as they encourage better generalization on unseen data. The model achieves stable performance after approximately 20 epochs, with accuracy values fluctuating around a high plateau. The best validation accuracy is achieved around epoch 24, indicating that the model is already well-optimized before completing all training epochs.

3) Overall Performance Evaluation

The final evaluation on the test dataset reports an accuracy of 94.44%, confirming the robustness and effectiveness of the proposed Xception-based model. The close alignment between training and validation curves, combined with low loss values and high accuracy, demonstrates that the model successfully learns discriminative features while maintaining strong generalization performance. Overall, these results validate the suitability of the Xception architecture for Crystal Guava leaf disease classification and support its potential application in real-world decision support systems for plant disease monitoring.

The Figure 5 illustrates the training and validation performance of the CNN model over 50 epochs in terms of loss and accuracy. The training loss decreases consistently as the number of epochs increases, indicating that the model effectively learns discriminative features from Crystal Guava leaf images. In contrast, the validation loss shows noticeable fluctuations but remains within a stable range, suggesting variability in the validation data without severe overfitting. The accuracy curves demonstrate a rapid improvement during the early epochs, followed by a stable trend, where training accuracy remains slightly higher than validation accuracy. The best validation accuracy is achieved at approximately epoch 32, and the final evaluation on the test dataset yields an accuracy of 93.89%, confirming that the model exhibits good generalization capability and reliable performance for Crystal Guava leaf disease classification.

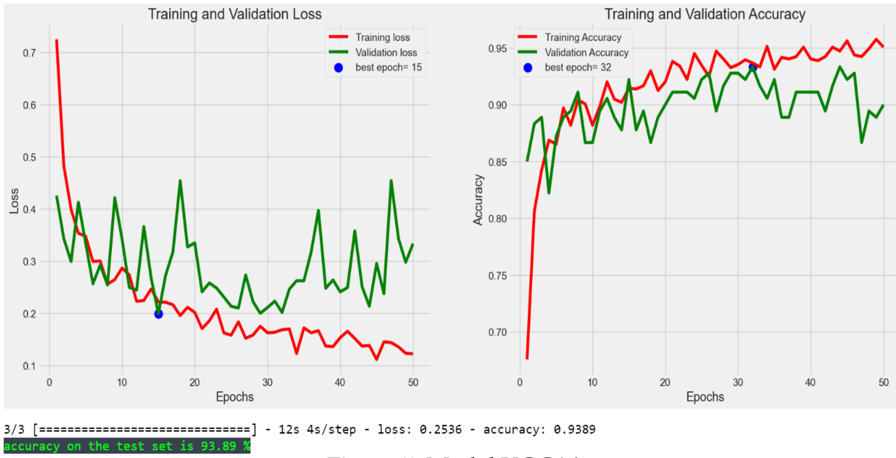

Figure 5. Model VGG16.

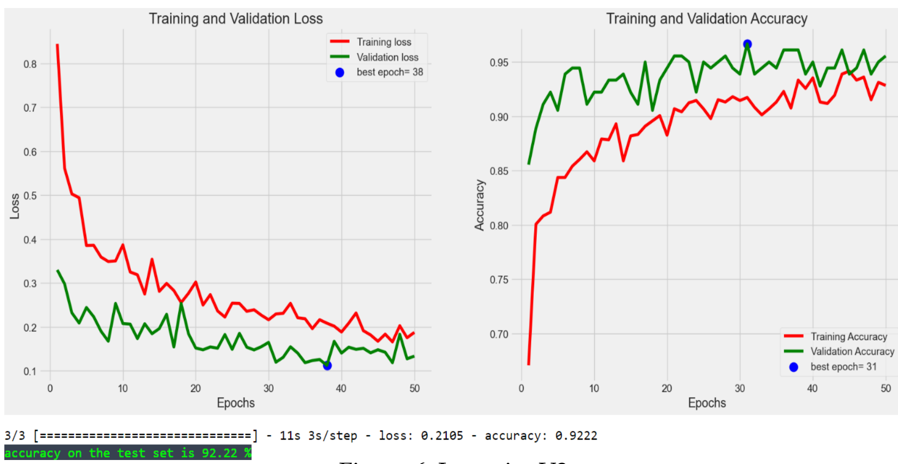

Figure 6. InceptionV3.

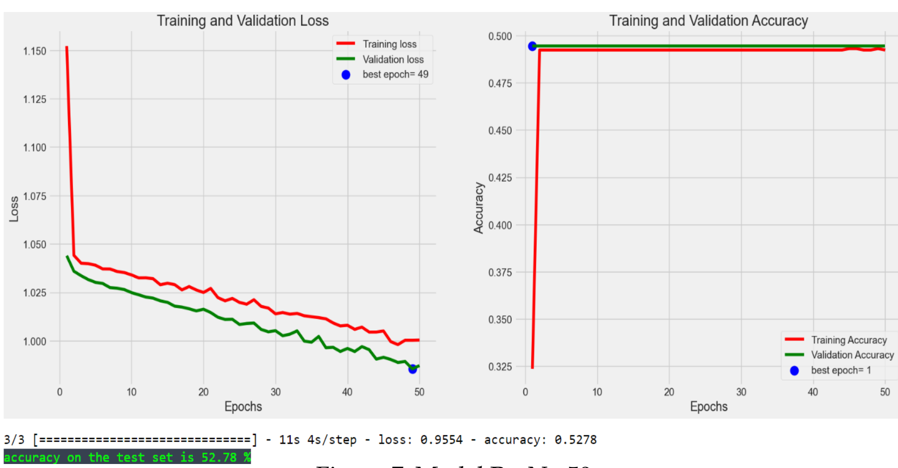

Figure 7. Model ResNet50.

The Figure 6 shows that both training and validation loss decrease consistently throughout the training process, indicating that the model effectively learns important patterns from the leaf images. The accuracy curves increase rapidly in the early epochs and then stabilize, with validation accuracy remaining comparable to or slightly higher than training accuracy, suggesting good generalization without significant overfitting. The best validation performance is achieved around epoch 38, and the final evaluation on the test set reaches an accuracy of 92.22%, demonstrating that the model performs reliably for Crystal Guava leaf disease classification. Meanwhile, as illustrated by Figure 7, the ResNet50 model demonstrated the highest level of accuracy with a percentage of 52.78%.

To evaluate model performance, four evaluation parameters are used: training and testing accuracy, F1-score, and model size. The goal is to identify the best model, characterized by the highest accuracy and F1-score, while also having a smaller model size. The model with the highest accuracy and F1-score, combined with a smaller model size, will be considered the best performing model. These performance metrics are important in determining the effectiveness of each model and identifying areas for improvement. The results will provide valuable insights into the strengths and weaknesses of each model, ultimately leading to the selection of the most appropriate model for a given application.

Table 3. Comparison of CNN Models Results.

Model	Training Accuracy	Macro F1	Testing Accuracy	Model Size
VGG16	95.7%	94%	93.89%	8.13MB
ResNet50	49,3%	23%	52.78%	721.42 KB
InceptionV3	94.1	92%	92.22%	16.03MB
Xception-CNN	94%	94%	94.44%	2.32MB

Based on Table 3, a comparison of model performance can be concluded that the Transfer Learning Xception-CNN model performed best. Although it had lower training accuracy compared to VGG16 (Figure 5) and InceptionV3 (Figure 6), it had higher testing accuracy and a smaller model size, indicating that Xception-CNN had better classification prediction performance than the other two architectures. Overall, the Transfer Learning model had higher accuracy and more stable performance compared to the model without transfer learning. This is consistent with the graph shown in Figure 5, which shows that the highest training result showed an accuracy of 94% and a loss of 0.17 after 50 epochs. The error converged to a minimum value, indicating convergence and stability through repeated iterations during the training phase. Consequently, the network was able to accurately classify data.

Based on this model, the next step was to validate it using a confusion matrix to determine whether the test data matched the model's predictions, as shown in Figure 8. The results of the validation test using the confusion matrix method are as follows: The confusion matrix shows that the model has accurately predicted the correct class for most of the test data. This indicates that the model is able to generalize well and make accurate predictions on new data that has not been seen before. The metrics shown in the confusion matrix, such as precision, recall, and F1-score, indicate that the model performs well in terms of classification accuracy. High values for these metrics indicate that the model is reliable and can be trusted to make accurate predictions.

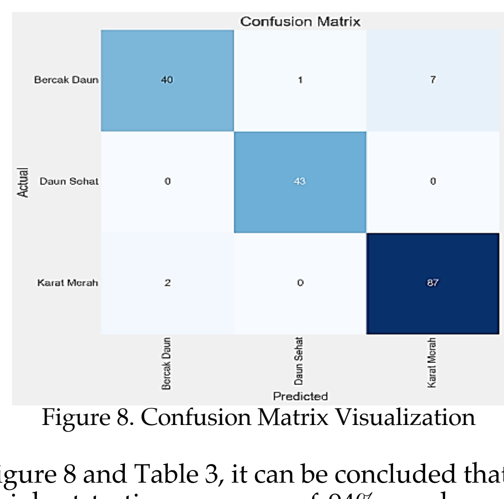

Figure 8. Confusion Matrix Visualization

Based on the results in Figure 8 and Table 3, it can be concluded that the Transfer Learning Xception model with CNN has the highest testing accuracy of 94%, as shown in Table 4 below. The highest precision is found in the Healthy Leaf class at 98% and the lowest in the Red Rust class at 93%. Recall is also highest in the Healthy Leaf class at 100% and the lowest in the Leaf Spot class at 83%. The average F1-score is 94%, which indicates that the model still has limitations in classifying diseases in several categories, such as Leaf Spot with an F1-score of 89% and Red Rust with an F1-score of 95%. This indicates that the model is still confused by the patterns of some test data. The problem that occurs in all models is the limited collection of training data, which consists of only a few perspectives taken repeatedly and not uniformly for each class. Collecting data from too few angles will make the model less able to learn image characteristics, while repeated data collection from one angle will cause overfitting at the beginning. In addition to data collection, long training times will increase the error value during testing, which can reduce recall and precision values.

Table 4. Results of Confusion Matrix

	Precision	***Recall***	***F1-score***	***Number of Sample***
Bercak Daun	0.95	0.83	0.89	48
Daun Sehat	0.98	1.00	0.99	43
Karat Merah	0.93	0.98	0.95	89
Accuracy			0.94	180
Macro avg.	0.95	0.94	0.94	180
Weighted avg.	0.95	0.94	0.94	180

V. Conclusion

This study was conducted to identify diseases in Crystal Guava plants using the Convolutional Neural Network (CNN) method. The results indicate that CNN-based models are capable of effectively detecting and classifying various types of diseases in Crystal Guava leaves, thereby improving the efficiency of plant health monitoring and increasing awareness of the importance of scientific approaches in agriculture. In this study, several CNN architectures, including VGG16, InceptionV3, and ResNet50, were employed for performance comparison. Although the Xception-CNN model achieved slightly lower accuracy during the training phase compared to VGG16 and InceptionV3, it demonstrated superior generalization capability during testing, achieving the highest classification accuracy of 94.44%. In comparison, VGG16 and InceptionV3 achieved testing accuracies of 93.89% and 92.22%, respectively. These findings confirm that the CNN approach is an effective and reliable alternative for plant disease identification.

Future research can further enhance this work by expanding the dataset to include a larger variety of disease types, different growth stages, and images captured under diverse environmental conditions, such as varying lighting and background complexity. Additionally, integrating advanced techniques such as attention mechanisms, ensemble learning, or hybrid CNN–Vision Transformer (ViT) models may improve feature discrimination and robustness. Future studies may also explore real-time deployment through mobile or Internet of Things (IoT)-based applications, enabling farmers to perform early disease detection directly in the field. Moreover, incorporating disease severity estimation and decision-support recommendations could transform the system into a comprehensive smart agriculture tool. These research directions highlight the strong potential of this study as a foundation for developing scalable, intelligent, and practical plant disease detection systems.

References

[1] Kundrat, Umar, L. S., & Sumarna. (2022). (Psidium guajava L) di Desa Ciwaringin Kecamatan Lemahabang Kabupaten Karawang. 4, 23–31.

[2] Hanik, N. R., Fitriani, R. D. A., Cahyanti, F. A., Oktavianingtyas, D., & Wahyuni, T. (2023). Identification of Pests and Diseases Crystal Guava (Psidium guajava L.) in Ngargoyoso District, Karanganyar Regency. Jurnal Biologi Tropis, 23(3), 127–135. https://doi.org/10.29303/jbt.v23i3.5021.

[3] Al-Adhaileh, M. H., Verma, A., Aldhyani, T. H. H., & Koundal, D. (2023). Potato Blight Detection Using Fine-Tuned CNN Architecture. Mathematics, 11(6). https://doi.org/10.3390/math11061516

[4] Alqadi, Z., Khrisat, M., Hindi, A., Omar Dwairi, M., & Khrisat, M. S. (2020). Features Analysis of RGB Color Image based on Wavelet Packet Information. International Journal of Computer Science and Mobile Computing, 9(3), 149–156. https://www.researchgate.net/publication/340249851

[5] Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. In Journal of Big Data (Vol. 8, Issue 1). Springer International Publishing. https://doi.org/10.1186/s40537-021-00444-8

[6] Brodzicki, A., Piekarski, M., Kucharski, D., Jaworek-Korjakowska, J., & Gorgon, M. (2020). Transfer

Learning Methods as a New Approach in Computer Vision Tasks with Small Datasets. Foundations of Computing and Decision Sciences, 45(3), 179–193. https://doi.org/10.2478/fcds-2020-0010

[7] Doutoum, A. S., Eryigit, R., & Tugrul, B. (2023). Classification of Guava Leaf Disease using Deep Learning. WSEAS Transactions on Information Science and Applications, 20(October), 356–363. https://doi.org/10.37394/23209.2023.20.38

[8] Ghosh, A., Sufian, A., Sultana, F., Chakrabarti, A., & De, D. (2019). Fundamental concepts of convolutional neural network. In Intelligent Systems Reference Library (Vol. 172, Issue January). https://doi.org/10.1007/978-3-030-32644-9_36

[9] Hartati, S., Puspasari, L. T., Suminar, E., & Meliansyah, R. (2023). Pengenalan Penyakit pada Tanaman Jambu Kristal (Psidium guajava L.) dan Pengendaliannya di Desa Mekar Jaya, Kecamatan Banjaran, Kabupaten Bandung. Agrikultura Masyarakat Tani, 1(1), 28. https://doi.org/10.24198/agrimasta.v1i1.51096

[10] Hasan, M. A., Riyanto, Y., & Riana, D. (2021). Grape leaf image disease classification using CNN-VGG16 model. Jurnal Teknologi Dan Sistem Komputer, 9(4), 218–223. https://doi.org/10.14710/jtsiskom.2021.14013

[11] Krisnandi, D., Pardede, H. F., Yuwana, R. S., Zilvan, V., Heryana, A., Fauziah, F., & Rahadi, V. P. (2019). Diseases Classification for Tea Plant Using Concatenated Convolution Neural Network. CommIT (Communication and Information Technology) Journal, 13(2), 67–77. https://doi.org/10.21512/commit.v13i2.5886

[12] Mahmud, K. H., Adiwijaya, & Faraby, S. Al. (2019). Klasifikasi Citra Multi-Kelas Menggunakan Convolutional Neural Network Studi Terkait Residual Neural Network. E-Proceeding of Engineering, 6(1), 2127–2136.

[13] Muslem R, I., Johan, T. M. J., & Luthfi. (2023). Klasifikasi Citra Ikan Menggunakan Algoritma Convolutional Neural Network dengan Arsitektur VGG-16. 4(2), 978–985. https://doi.org/10.30865/klik.v4i2.1209

[14] Muslikh, A. R., Setiadi, D. R. I. M., & Ojugo, A. A. (2023). Rice Disease Recognition Using Transfer Learning Xception Convolutional Neural Network. Jurnal Teknik Informatika (Jutif), 4(6), 1535–1540. https://doi.org/10.52436/1.jutif.2023.4.6.1529

[15] Nugraha, G. S., Darmawan, M. I., & Dwiyansaputra, R. (2023). Comparison of CNN's Architecture GoogleNet, AlexNet, VGG-16, Lenet -5, Resnet-50 in Arabic Handwriting Pattern Recognition. Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control, 4(2). https://doi.org/10.22219/kinetik.v8i2.1667

[16] Oktaviani, D., Yudianto, A., Maimunah, Sukmasetya, P., & Hasani, R. A. (2024). Klasifikasi Citra Candi Berdasarkan Tekstur Bentuk Menggunakan Convolutional Neural Network. Jurnal Informatika Polinema, 10(2), 181–188. https://doi.org/10.33795/jip.v10i2.4999

[17] Pramudhita, D. A., Azzahra, F., Arfat, K., Magdalena, R., & Saidah, S. (2023). Strawberry Plant Diseases Classification Using CNN Based on MobileNetV3-Large and EfficientNet-B0 Architecture ARTICLE INFO ABSTRACT. Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika (JITEKI), 9(3), 522–534. https://doi.org/10.26555/jiteki.v9i3.26341

[18] Purwono, Ma'arif, A., Rahmaniar, W., Fathurrahman, H. I. K., Frisky, A. Z. K., & Haq, Q. M. U. (2022). Understanding of Convolutional Neural Network (CNN): A Review. International Journal of Robotics and Control Systems, 2(4), 739–748. https://doi.org/10.31763/ijrcs.v2i4.888

[19] Rohim, A., Sari, Y. A., & Tibyani. (2019). Convolution Neural Network (CNN) untuk Pengklasifikasian Citra Makanan Tradisional. Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer, 3(7), 7038–7042.

[20] Safran, M., Alrajhi, W., & Alfarhood, S. (2023). DPXception: a lightweight CNN for image-based date palm species classification. Frontiers in Plant Science, 14(January), 1–14. https://doi.org/10.3389/fpls.2023.1281724

[21] Sharma, P., Berwal, Y. P. S., & Ghai, W. (2020). Performance analysis of deep learning CNN models for disease detection in plants using image segmentation. Information Processing in Agriculture, 7(4), 566–574. https://doi.org/10.1016/j.inpa.2019.11.001

[22] Yu, Z., Dong, Y., Cheng, J., Sun, M., & Su, F. (2022). Research on Face Recognition Classification Based on Improved GoogleNet. Security and Communication Networks, 2022. https://doi.org/10.1155/2022/7192306

[23] Yuesheng, F., Jian, S., Fuxiang, X., Yang, B., Xiang, Z., Peng, G., Zhengtao, W., & Shengqiao, X. (2021). Circular Fruit and Vegetable Classification Based on Optimized GoogLeNet. IEEE Access, 9, 113599–113611. https://doi.org/10.1109/ACCESS.2021.3105112.

[24] Zafar, A., Aamir, M., Mohd Nawi, N., Arshad, A., Riaz, S., Alruban, A., Dutta, A. K., & Almotairi, S. (2022). A Comparison of Pooling Methods for Convolutional Neural Networks. Applied Sciences (Switzerland), 12(17), 1–21. https://doi.org/10.3390/app12178643.