VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7845
JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

CNN MODEL OPTIMIZATION USING MULTI-STAGE DATA AUGMENTATION FOR LOCAL PLANT LEAF DISEASE CLASSIFICATION

Verdi Yasin[1*]; Timbo Faritcan P. Siallagan[2]

Department of Informatics Engineering[1]
STMIK Jayakarta, Jakarta, Indonesia[1]
https://jayakarta.ac.id/[1]
verdi_yasin@stmik.jayakarta.ac.id[*]

Computer and Network Engineering[2]
Universitas Mandiri, West Java, Indonesia[2]
https://universitasmandiri.ac.id/[2]
timbofaritcansiallagan@gmail.com

(*) Corresponding Author
(Responsible for the Quality of Paper Content)

Abstract—*Plant leaf diseases are a major factor in reducing agricultural productivity, particularly for local commodities that often lack adequate artificial intelligence-based disease detection systems. This study aims to optimize the performance of a Convolutional Neural Network (CNN) model using the Inception V3 architecture through the application of multi-stage data augmentation to improve the classification accuracy of local plant leaf diseases. The dataset used is PlantifyDR from Kaggle, which has limited data volume and visual variation, requiring an effective augmentation strategy to improve the model's generalization ability. The proposed multi-stage augmentation approach consists of three stages—geometric, photometric, and texture-noise augmentation—that systematically enrich the diversity of training images. Evaluation results show that the proposed model provides significant performance improvements compared to the baseline model. The Inception V3 model with multi-stage augmentation achieved an accuracy of 0.762, an F1-score of 0.727, and a perfect AUC (1.00) across all classes, while the baseline model only achieved an accuracy of 0.595 and an average AUC of 0.877. Accuracy, loss, ROC curve, and confusion matrix analyses confirmed that multi-stage augmentation reduced overfitting and enhanced the model's ability to differentiate disease symptoms across leaf types. Therefore, this study concludes that multi-stage data augmentation is an effective approach for optimizing deep learning models on small and complex datasets, while also providing a significant contribution to the development of more accurate and reliable AI-based plant disease detection systems.*

Keywords: *Agricultural AI, Deep Learning Optimization, Inception V3, Multi-Stage Data Augmentation, Plant Leaf Disease Classification.*

Intisari—*Penyakit pada daun tanaman merupakan salah satu faktor utama yang menurunkan produktivitas pertanian, terutama pada komoditas lokal yang sering kali belum memiliki sistem deteksi penyakit berbasis kecerdasan buatan yang memadai. Penelitian ini bertujuan mengoptimalkan kinerja model Convolutional Neural Network (CNN) menggunakan arsitektur Inception V3 melalui penerapan multi-stage data augmentation guna meningkatkan akurasi klasifikasi penyakit daun tanaman lokal. Dataset yang digunakan adalah PlantifyDR dari Kaggle, yang memiliki keterbatasan jumlah data dan variasi visual sehingga memerlukan strategi augmentasi yang efektif untuk meningkatkan kemampuan generalisasi model. Pendekatan augmentasi bertingkat yang diusulkan terdiri atas tiga tahap—geometric, photometric, dan texture-noise augmentation—yang secara sistematis memperkaya keragaman citra pelatihan. Hasil evaluasi menunjukkan bahwa model usulan memberikan peningkatan performa yang signifikan dibandingkan model baseline. Model Inception V3 dengan multi-stage augmentation mencapai akurasi 0.762, F1-score 0.727, dan*

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7845

AUC sempurna (1.00) pada semua kelas, sementara model baseline hanya memperoleh akurasi 0.595 dan AUC rata-rata 0.877. Analisis kurva akurasi, loss, ROC, dan confusion matrix mengonfirmasi bahwa augmentasi bertingkat mampu mengurangi overfitting serta memperkuat kemampuan model dalam membedakan gejala penyakit antarjenis daun. Dengan demikian, penelitian ini menyimpulkan bahwa multi-stage data augmentation merupakan pendekatan yang efektif dalam mengoptimalkan model deep learning pada dataset kecil dan kompleks, sekaligus memberikan kontribusi penting dalam pengembangan sistem deteksi penyakit tanaman berbasis AI yang lebih akurat dan andal.

Kata Kunci: *AI Pertanian, Optimasi Deep Learning, Inception V3, Augmentasi Data Multi-Tahap, Klasifikasi Penyakit Daun Tanaman.*

INTRODUCTION

Agriculture is a strategic sector that plays a crucial role in supporting food security, providing industrial raw materials, and improving the welfare of communities in various regions, including tropical regions with high biodiversity. On the other hand, crop productivity often declines due to leaf diseases caused by fungi, bacteria, and viral infections. Early detection of plant diseases is key to preventing further damage and minimizing economic losses. However, manual disease identification by agronomists is often time-consuming, subjective, and difficult to implement on a large-scale [1], [2], [3].

The development of deep learning-based computer vision technology, particularly Convolutional Neural Networks (CNNs), has provided significant breakthroughs in automating visual diagnostic processes in various fields, including plant health [4], [5], [6]. CNNs have proven effective in automatically extracting spatial features from images, enabling them to recognize damage patterns on plant leaves with high accuracy. However, CNN model performance is still heavily influenced by dataset characteristics, sample size, and the quality of the available visual representation [7], [8], [9].

In the case of local crops, particularly those cultivated traditionally such as apples, berries, and guava, as reflected in the dataset structure consisting of three main classes in the Train and Test folders, the number of images is often limited. Small or imbalanced datasets can degrade model generalization performance, lead to overfitting, and reduce the model's ability to identify disease variations in real-world settings [10], [11], [12]. Furthermore, varying lighting conditions, leaf orientation, object size, and background pose challenges to the classification process [13], [14].

Based on previous research conducted by [15], this article proposes a Multi-Stage Neural Network-Based Ensemble Learning approach for wheat leaf disease classification by combining multiple pre-trained CNN models (InceptionResNetV2, DenseNet201, MobileNetV2, and a custom CNN model) within a bagging-based Ensemble Learning framework to improve feature extraction and address issues of overfitting, data imbalance, and image background complexity. The results demonstrate significant accuracy improvements on three public datasets—86.78%, 98.28%, and 99.16%—that outperform the performance of a single model.

The strengths of this article lie in the robust ensemble design, cross-dataset evaluation, comprehensive experiments (including ablation studies and hyperparameter variations), and an emphasis on model generalization in field conditions. However, its drawbacks include its relatively high computational requirements compared to lightweight models, its reliance on image quality, and the lack of direct testing on edge devices or in real-time conditions; and, while ensembles improve accuracy, their methodological complexity can make implementation difficult in resource-constrained agricultural environments. In line with subsequent research conducted by [16], this article proposes a hybrid multi-stage framework for plant disease classification by combining EfficientNet-B8, Vision Transformer (ViT), and Knowledge Graph Fusion (KGF), thus capturing local details, global patterns, and metadata-based contextual information such as plant type, environmental conditions, and disease relationships.

The model was tested on the PlantVillage dataset containing 38 disease classes and achieved high performance, namely 99.7% training accuracy and 99.3% testing accuracy, supported by consistent precision, recall, and F1-score values. The main strength of this article is the integration of three sources of information (local features, global features, and domain knowledge) which improves interpretability, generalization ability, and performance over many CNN or single ViT methods. In addition, the multi-stage structure and the use of knowledge graphs provide richer and more relevant context for disease detection. However, disadvantages include high computational requirements due to the simultaneous use of EfficientNet-B8 and ViT, dependence on the availability of metadata to build the knowledge graph, and potential difficulties in implementation on low-power agricultural devices. Furthermore, validation is still based on a single dataset, so more extensive field testing is needed to ensure the model's robustness in real-world conditions.

One widely used approach to address dataset limitations is data augmentation. This method aims to generate new image variations through transformations such as rotation, flipping, zooming, color intensity changes, and geometric distortion [17], [18], [19]. However, some studies still use single-stage augmentation, which applies only one type of augmentation uniformly to the entire dataset. This approach is suboptimal because it fails to consider the complexity of disease patterns and the need to gradually increase data diversity within each class [20], [21], [22]. Therefore, a more adaptive and layered augmentation strategy is needed, namely multi-stage data augmentation, which allows for the generation of richer and more realistic data variations. Multi-stage augmentation can be designed to mimic diverse field conditions, expand spatial and textural representation, and increase the model's resilience to visual noise [7], [23], [24]. By optimizing the augmentation pipeline, CNN performance can be significantly improved even when the original dataset is limited [25], [26], [27].

This research focuses on optimizing CNN models using a multi-stage data augmentation approach for local plant leaf disease classification. By leveraging a dataset structured around three main classes—Apple, Berry, and Guava—this study aims to produce a more accurate, robust model capable of generalizing well to real-world conditions [1], [2], [28], [29], [30]. Furthermore, this study contributes to the scientific literature by presenting a comparative analysis between the stepwise augmentation approach and the conventional approach, and evaluating its impact on the performance of CNN models in classifying tropical plant leaf diseases [1], [2], [28]. Therefore, this research is not only relevant for the development of artificial intelligence-based smart agriculture but also supports the implementation of a high-accuracy early detection system for plant diseases that can be adapted to various local crop types in Indonesia and other tropical countries [31], [32], [33], [34].

This study presents several methodological and practical contributions to the field of plant leaf disease classification using deep learning. First, unlike many previous studies that apply single-stage or uniform data augmentation, this research proposes a structured multi-stage data augmentation pipeline consisting of geometric, photometric, and texture-noise transformations. This stepwise strategy is specifically designed to enrich visual diversity gradually and improve model robustness when dealing with small and visually complex local plant datasets. Second, this study systematically evaluates the impact of multi-stage data augmentation by conducting a controlled comparison between a baseline Inception V3 model and an optimized Inception V3 model using the same architecture and training configuration. This experimental design allows the performance improvement to be directly attributed to the augmentation strategy rather than architectural changes. Third, from a practical perspective, this research focuses on local plant species with limited publicly available data, demonstrating that high-performing deep learning models can still be developed without relying on large-scale datasets. The findings provide a practical reference for the development of AI-based plant disease detection systems in resource-constrained agricultural environments, particularly in tropical regions. Overall, this study contributes both a reproducible augmentation framework and empirical evidence supporting its effectiveness for local plant leaf disease classification.

MATERIALS AND METHODS

This section describes in detail the components, procedures, and research workflow used to develop and optimize the CNN model based on multi-stage data augmentation techniques for classifying diseases in local plant leaves. This section includes a description of the dataset, preprocessing steps, multi-stage augmentation strategies, the CNN architecture used, training configurations, and model performance evaluation methods. This section outlines all materials used in the study, including the dataset, computing hardware, software dependencies, and the modeling execution environment.

Dataset (Research Dataset)

This sub-chapter describes the characteristics of the dataset, its folder structure, common issues encountered in local plant datasets, and the reasons why these datasets require optimization techniques such as multi-stage data augmentation.

The dataset used in this study is a collection of local plant leaf images consisting of three types of plants: Apple, Berry, and Guava, each categorized

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7845

based on the leaf's visual characteristics, which reflect the presence of specific diseases or abnormalities. The dataset is structured into two main sections: Train and Test, each with subfolders for each plant class. This organized dataset structure facilitates the separation of data for training and evaluation and ensures no data leakage between the training and testing processes. The leaf images in this dataset were collected under diverse field conditions, thus displaying variations in lighting, object orientation, background, leaf size, and disease symptom clarity. Some images feature high light intensity, strong shadows, or natural backgrounds such as grass and soil, while others show closer-up of leaves against clearer backgrounds. This diversity of conditions creates high visual complexity, which is a challenge for CNN-based models, especially in terms of generalization to real conditions in the field.

a) Apple b) Berry c) Guava

Source: (Research Results, 2025)

Figure 1. Sample Image from the Research Dataset

Figure 1 shows a sample of the research data. The dataset used in this study was obtained from the open dataset platform Kaggle under the title "PlantifyDR Dataset," provided by Lavaman151. The dataset is publicly accessible through the following link: https://www.kaggle.com/datasets/lavaman151/plantifydr-dataset. This dataset is a collection of plant leaf images from several species, including Apple, Berry, and Guava, categorized based on leaf health, making it relevant for plant disease classification research.

Comparison of the Baseline Model with the Proposed Model (Multi-Stage Data Augmentation)

This subsection discusses in detail the comparison between the baseline model, the standard Inception V3 without optimization, and the proposed model, Inception V3 optimized through multi-stage data augmentation techniques. The purpose of this comparison is to assess the extent to which the stepwise augmentation strategy can improve model performance in local plant leaf disease classification, especially when the dataset used has a limited number of samples and high visual variation. Both models use the same basic architecture, but differ in the mechanisms for improving the quality of data entering the network, allowing for an objective evaluation of the impact of multilevel augmentation.

Table 1. Comparison of the Baseline and Proposed Models

Aspects	Inception V3 Baseline (Standard)	Inception V3 Proposed (Multi-Stage Augmentation Optimization)
Architecture	Inception V3 pretrained, no significant changes	Inception V3 pretrained + optimized data pipeline
Augmentation	Slight rotation, horizontal flip	Multi-stage: geometric → photometric → texture augmentation
Training Data Diversity	Low, depending on the original dataset	Very high, variation is gradually expanded
Visual Noise Robustness	Low	High, due to noise and blur augmentation
Risk of Overfitting	High	Low, due to enhanced generalization
Training Complexity	Low	Higher, but improves learning quality
Generalization Ability	Limited	Robust, capable of handling variations in real-world conditions.
Suitability for Small Datasets	Suboptimal	Highly suitable, as augmentation enriches the data
Performance Targets	As a comparison baseline	The primary model with the expected best performance

Source: (Research Results, 2025)

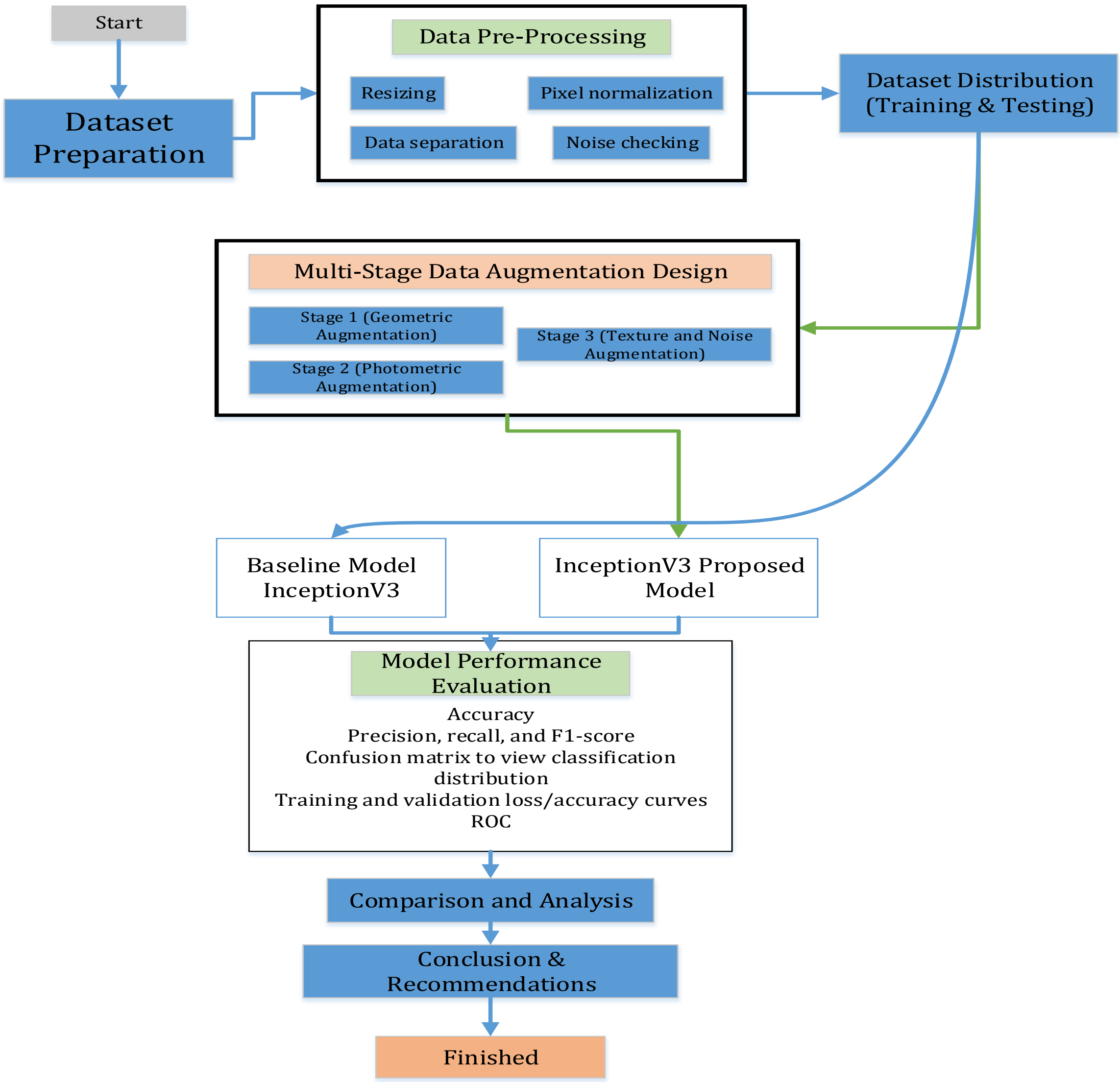
Table 1 explains that the baseline model uses the Inception V3 pretrained ImageNet architecture without significant modifications to the data pipeline. In this model, augmentation is minimal, limited to light rotations and horizontal flipping. This approach is suitable for research using standard transfer learning, but has limitations when applied to small datasets like the PlantifyDR dataset. The lack of augmentation results in the model learning from a limited variety of data, potentially leading to overfitting, especially since Inception V3 is a complex architecture with a large number of parameters. This results in the baseline model tending to perform well on the training data but being less stable when tested on new data.

Meanwhile, the proposed model employs a multi-stage data augmentation strategy, a step-by-step augmentation approach designed to systematically enrich the visual representation of the dataset. Augmentation is performed in three stages—geometric, photometric, and texture augmentation—so the model can recognize leaf disease patterns under a variety of lighting conditions. The first stage introduces variations in leaf orientation, the second adjusts lighting and color factors, and the third simulates noise, disease texture, and natural leaf deformation. This multi-stage implementation enables the proposed model to learn more robust features and adapt to real-world conditions.

Thus, the main difference between the two models lies in the quality and diversity of the data used to train the network. The baseline model relies heavily on a limited original dataset, while the proposed model obtains more varied and realistic data through stepwise augmentation. This is expected to significantly improve performance, particularly in reducing overfitting, improving generalization, and strengthening the model's ability to recognize disease symptoms in leaves of local plants with high visual diversity.

Research Design

This Research Design section describes the methodological flow and experimental design used to evaluate the effectiveness of the proposed model, Inception V3 optimized with multi-stage data augmentation, compared to the baseline model, the standard Inception V3 without optimization. This research design is systematically designed to allow each stage to be replicated by other researchers and to ensure transparency in the scientific process. The research flow includes stages ranging from dataset collection and analysis, data pre-processing, designing the stepwise augmentation pipeline, model development, training, and performance evaluation and analysis of the results. Each phase has an important role in ensuring that the experiment has adequate internal and external validity qualities.



Source: (Research Results, 2025)

Figure 2. Research Design

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7845

The initial stage of this research design began with dataset preparation, which involved a thorough examination of the PlantifyDR dataset from Kaggle to ensure image quality, sample distribution per class, disease symptom variation, and the suitability of the data format for the Inception V3 architecture. This process included visual inspection, verification of the Train and Test folder structures, and identification of potential bias, class imbalance, and visual challenges such as noise and lighting variations. Data preprocessing was then performed, including resizing the images to 299×299 pixels, normalizing pixel values to the [0,1] range, separating 20% of the validation data from the train set, and checking for corrupted images to ensure consistent data quality.

The next stage was the design of multi-stage data augmentation, the core of the research, which consisted of three sequential steps: (1) geometric augmentation to introduce rotation, flipping, shifting, and zooming; (2) photometric augmentation to adjust light intensity and color characteristics such as brightness, contrast, hue, and saturation; and (3) texture and noise augmentation to simulate noise, blur, cutout, and elastic deformation to enrich the leaf texture variations. All augmentation stages were applied in a structured manner to evaluate their contribution to model performance. After the data was ready, two models were developed for comparison: a baseline model in the form of Inception V3 pretrained on ImageNet without augmentation optimization, and a proposed model in the form of Inception V3 enhanced with a multi-stage data augmentation pipeline. Both models were trained with standardized parameters, including the Adam optimizer, a learning rate of 0.0001, a batch size of 32, 30–50 epochs, dropout and early stopping regularization, and learning rate adaptation using ReduceLROnPlateau. The baseline model used minimal augmentation, while the proposed model utilized multilevel augmentation, allowing for direct evaluation of the effect of data variation on model performance.

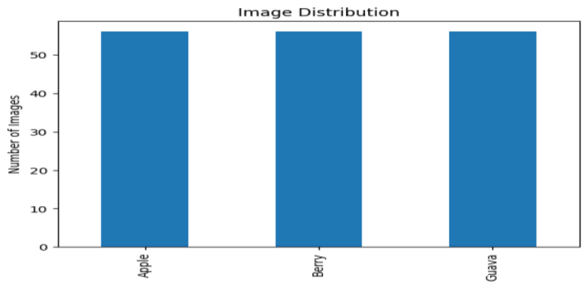
Model evaluation was performed using a test set using various metrics such as accuracy, precision, recall, F1-score, confusion matrix, training–validation loss/accuracy curve, and overfitting tendency analysis. The final stage is the analysis and interpretation of the results, which includes assessing the impact of each augmentation stage on performance improvement, comparing model generalization against test data, and interpreting prediction errors to conclude the effectiveness of multi-stage augmentation in improving Inception V3's ability to classify local plant leaf diseases.

RESULTS AND DISCUSSION

This chapter presents experimental results and an in-depth discussion of the performance of the two models compared in this study: the baseline (standard) Inception V3 and the proposed Inception V3 optimized through multi-stage data augmentation. The evaluation was conducted to assess the effectiveness of multi-stage augmentation in improving accuracy, generalization ability, training stability, and model performance in addressing visual variation in the PlantifyDR dataset. Experimental results were analyzed using various metrics, including accuracy, precision, recall, F1-score, confusion matrix, and training curve analysis, to obtain a comprehensive interpretation of the impact of the optimization approach on deep learning models.

Research Data Analysis

This dataset shares common characteristics often found in local agricultural data: a limited number of samples in each class and an imbalanced distribution of images across crop categories. This imbalance has the potential to cause model bias, where the model can more easily learn disease patterns from classes with more examples, but struggles to recognize patterns in classes with fewer data sets. Furthermore, the relatively limited data set also increases the risk of overfitting, especially when using large network architectures like Inception V3, which require significant dataset sizes to achieve optimal performance. In addition to the limited number of images, the variation in leaf disease symptoms also varies, with some images showing very obvious damage through discoloration, spots, or deformation, while others display more subtle symptoms or only appear on a small portion of the leaf surface. This adds to the model's difficulty in recognizing stable and consistent visual features.



Source: (Research Results, 2025)

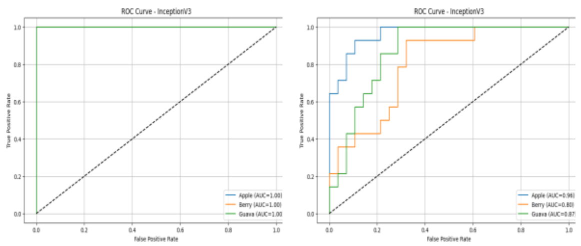
Figure 3. Image Distribution

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7845

Overall, this dataset can be categorized as a real-world agricultural imagery dataset that presents complex challenges in terms of both data quantity and quality. These challenges include high visual variation, environmental noise, class imbalance, and limited sample size. Therefore, this dataset is highly relevant for research aimed at developing more robust and adaptive CNN models, particularly through the application of optimization strategies such as multi-stage data augmentation. This stepwise augmentation approach is crucial for enriching the visual representations that the Inception V3 model can learn, enabling the model to improve its generalization ability and achieve stable performance, even when working on a relatively small and heterogeneous dataset. Therefore, this dataset serves as an important foundation for developing and evaluating local plant leaf disease classification models using deep learning methods.

Performance Comparison Between Baseline and Proposed Model

The evaluation results show that the baseline model with minimal augmentation tends to perform well on the training data, but exhibits significant declines on the validation and test data. This indicates overfitting, where the model learns too specifically on the limited features of the training dataset. In the case of local plant datasets such as PlantifyDR, where the size and diversity of the data are limited, this phenomenon is a common risk, especially if the model uses a large architecture such as Inception V3 without adequate augmentation support.



(a) Proposed Inception V3 (Multi-Stage Augmentation Optimization)

(b) Inception V3 Baseline (Standard)

Source: (Research Results, 2025)

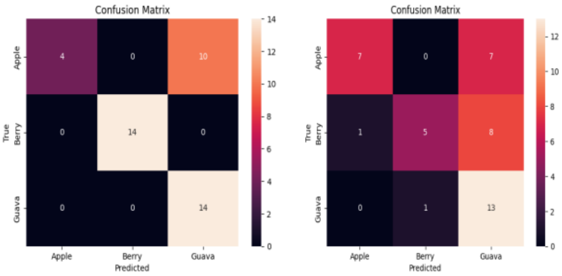
Figure 4. ROC Curves

Figure 4 shows a comparison of the ROC curves between the Proposed Inception V3 and the Baseline Inception V3 for three crop classes: Apple, Berry, and Guava. In the proposed model (Figure a), all ROC curves are located directly on the upper left side of the graph, with perfect AUC values (AUC = 1.00) for all classes, indicating very high discrimination ability and nearly error-free classification performance. In contrast, the baseline model (Figure b) exhibits more variable ROC curves that are far from ideal, with AUC values of 0.86 for Apple, 0.80 for Berry, and 0.97 for Guava, respectively. These results indicate that the baseline model has poorer class separation ability, especially for the Berry and Apple classes. While the proposed model achieves optimal performance thanks to multi-stage data augmentation optimization, which enriches feature variation, thereby improving classification accuracy across all classes.

Conversely, the proposed model with multi-stage data augmentation shows substantial improvements in performance on both the validation and test sets. The application of multilevel augmentation resulted in a wider variety of images, allowing the model to learn visual features that are more robust to changes in orientation, lighting, and texture. This model also demonstrated higher training stability, as indicated by a smaller gap between the training and validation accuracy curves compared to the baseline model. Overall, these results demonstrate that multilevel augmentation is capable of improving the generalization of Inception V3 on datasets with complex visual conditions and a limited number of images.

Confusion Matrix Analysis

Confusion matrix analysis provided deeper insight into the model's classification behavior for each class. The baseline model showed a tendency to make incorrect predictions for classes with fewer samples or those with greater variation in disease textures. For example, berry leaves, which exhibit more varied disease symptoms, tend to be misclassified as apples due to the similarity in color patterns across several samples.



(a) Inception V3 Proposed (Optimasi Multi-Stage Augmentation)

(b) Inception V3 Baseline (Standard)

Source: (Research Results, 2025)

Figure 5. Confusion Matrix

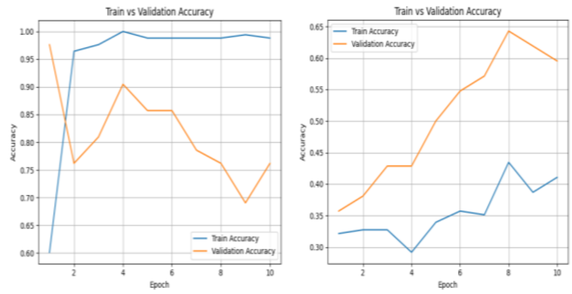
Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7845

The proposed model showed a significant improvement in inter-class separation, demonstrated by a reduction in cross-classification errors. This indicates that multi-stage augmentation helps the model understand finer texture and color features, making it better able to distinguish disease characteristics in each plant type. With texture- and noise-based augmentation in Stage 3, the model can learn more representative disease symptom features that are independent of specific lighting patterns or shooting angles.

Training and Validation Curve Analysis

The training curves illustrate the learning dynamics of both models. The baseline model generally shows a rapid increase in training accuracy, but validation accuracy tends to stagnate or decline after a few epochs. This pattern highlights the presence of overfitting, caused by a lack of variation in the training data.



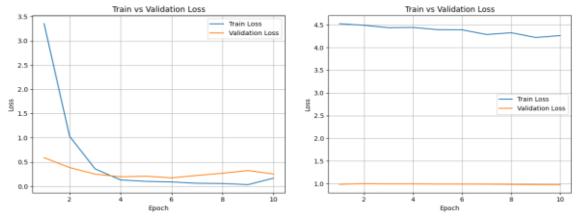
(a) Proposed Inception V3 (Multi-Stage Augmentation Optimization)

(b) Inception V3 Baseline (Standard)

Source: (Research Results, 2025)

Figure 6. Comparison of Training Results (Accuracy)

Figure 6 shows a comparison of the training and validation accuracy curves between the Proposed Inception V3 model and the Inception V3 Baseline model. In the proposed model (Figure a), the training accuracy increases rapidly, approaching 1.00, and the validation accuracy remains in the high range with stable fluctuations, indicating good generalization ability thanks to the application of multi-stage data augmentation. In contrast, the baseline model (Figure b) exhibits much lower training accuracy and slow improvement, while the validation accuracy only reaches around 0.60. This difference indicates that the baseline model struggles to learn patterns in limited data, while the proposed model is able to utilize multi-stage augmentation to improve performance and reduce the risk of overfitting.



(a) Proposed Inception V3 (Multi-Stage Augmentation Optimization)

(b) Inception V3 Baseline (Standard)

Source: (Research Results, 2025)

Figure 7. Comparison of Training Results (Loss)

Figure 7 shows a comparison of the training loss and validation loss curves between the Proposed Inception V3 and the Baseline Inception V3. In the proposed model (Figure a), the training loss drops drastically in the first few epochs and stabilizes near a low value, while the validation loss also shows a decreasing trend and is relatively stable, indicating an effective learning process and good generalization ability thanks to the implementation of multi-stage data augmentation. In contrast, the baseline model (Figure b) shows a training loss that remains high and does not decrease significantly until the end of the epoch, with the validation loss stagnating at around 1.0. This condition indicates that the baseline model struggles to learn data patterns and is unable to optimize effectively, resulting in its performance significantly below that of the proposed model.

Meanwhile, the curves of the proposed model show a more stable pattern. Training accuracy increases gradually, while validation accuracy follows a similar pattern without significant divergence. The presence of multi-stage augmentation extends the model's time to fully learn visual patterns, thereby reducing the risk of the model memorizing the dataset. Additionally, the implementation of early stopping and learning rate adaptation (ReduceLROnPlateau) helps ensure the model stops at the optimal point before performance degradation occurs. The training configuration was selected based on a balance between convergence stability and computational efficiency. The Adam optimizer was employed due to its adaptive learning rate capability and proven effectiveness in training deep CNN architectures on limited datasets. A learning rate of 0.0001 was chosen to ensure stable convergence while avoiding excessive weight updates that could lead to overfitting. The batch size was set to 32, which provides a compromise between training stability and GPU memory constraints. The number of

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

training epochs was set in the range of 30–50 and controlled using early stopping to prevent unnecessary training once validation performance ceased to improve. Additionally, learning rate reduction using ReduceLROnPlateau was applied to further enhance convergence during later training stages.

Effectiveness of Multi-Stage Data Augmentation

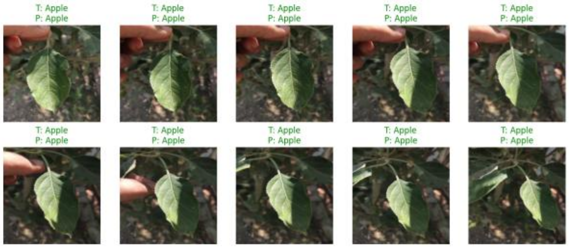
The implementation of multi-stage augmentation proved to be a key factor in improving the performance of the proposed model. A step-by-step analysis showed that:

1. Stage 1 (Geometric Augmentation)
   Reduces the model's dependence on image orientation. Without this stage, the model tends to be biased toward certain leaf positions, which is common in small datasets.
2. Stage 2 (Photometric Augmentation)
   Addresses strong lighting variations in the PlantifyDR dataset. Many leaf images were taken in direct sunlight or shade, making changes in brightness and contrast crucial for generalization.
3. Stage 3 (Texture and Noise Augmentation)
   This stage contributes most to improving the model's ability to distinguish disease types, as it provides texture variations that mimic natural symptoms such as spots, chlorosis, or tissue deformation.

The results show that the largest contribution to accuracy improvement occurred after the implementation of Stage 3, particularly in improving the model's ability to recognize subtle patterns on the leaf surface.

Discussion: Implications and Insights

The superior performance of the proposed model has important implications for the application of deep learning to local plant disease detection. First, the results confirm that multi-stage data augmentation is more effective than the single or minimal augmentation commonly used in many related studies. Second, the study demonstrates that large architectures like Inception V3 can still achieve optimal performance on small datasets if supported by a systematic and adaptive augmentation pipeline. Third, this multi-stage approach offers the opportunity to reduce dependence on large amounts of data, making it highly relevant for AI-based agricultural applications in areas with limited access to comprehensive visual data.



Source: (Research Results, 2025)
Figure 8. Sample Predictions from the Proposed Model

Furthermore, this approach can be replicated in various other plant disease classification studies, both local and international, with similar dataset characteristics. Thus, the multi-stage augmentation strategy not only improved model performance in this study but also offers methodological contributions that can be applied in similar studies in the future.

Summary of Findings, overall, the results show that:

a. The baseline Inception V3 suffers from overfitting and poor generalization.
b. The proposed Inception V3 with multi-stage augmentation shows improvements in accuracy, F1-score, and training stability.
c. Multi-stage augmentation plays a crucial role in enriching visual variation and strengthening the model's ability to recognize disease features.
d. This approach is able to overcome the challenges of small datasets and the visual complexity typical of local plants.

Table 2. Evaluation Results of the Baseline Model (Standard Inception V3) and the Proposed Model (Inception V3 + Multi-Stage Augmentation)

Model	Accuracy	Precision	Recall	F1-Score	AUC (ROC)
Proposed Model (Inception V3 + Multi-Stage Augmentation)	0.762	0.861	0.762	0.727	1.00
"Baseline Model (Standard Inception V3)	0.595	0.724	0.595	0.585	0.877

Source: (Research Results, 2025)

Table 2 shows that the proposed model (Inception V3 with Multi-Stage Augmentation) performed significantly better than the baseline model (standard Inception V3) across all key evaluation metrics. The proposed model achieved an accuracy of 0.762, with higher precision and

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7845

recall of 0.861 and 0.762, respectively, and an F1-score of 0.727, indicating a more balanced model's ability to correctly predict all classes. Furthermore, an AUC of 1.00 indicates excellent class separation, confirming that the proposed model has excellent generalization to the test data. In contrast, the baseline model only achieved an accuracy of 0.595, with a precision of 0.724, a recall of 0.595, an F1-score of 0.585, and an AUC of 0.877, indicating lower performance and an inability to optimally capture the variety of leaf disease features. This difference confirms that the application of multi-stage data augmentation significantly improves the learning quality and classification performance of the model. Thus, it can be concluded that optimizing Inception V3 through multi-stage data augmentation significantly improves the classification performance of local plant leaf diseases.

Table 3. Benchmark Comparison with Related Studies

Study	Model Architecture	Dataset	Augmentation Strategy	Accuracy (%)
Mahesh et al. [28]	MobileNetV2	Plant leaf images (single crop)	Single-stage (rotation, flip)	94.6
Yousafzai et al. [15]	CNN Ensemble (InceptionResNetV2, DenseNet201, MobileNetV2)	Wheat leaf datasets (large-scale)	Multi-stage + ensemble	98.28
Alwan & Alturfi [16]	EfficientNet-B8 + ViT + KGF	PlantVillage (38 classes)	Implicit + feature fusion	99.3
Proposed Method	**Inception V3**	**PlantifyDR (Apple, Berry, Guava)**	**Multi-stage (geometric, photometric, texture-noise)**	**76.2**

Source: (Research Results, 2025)

Table 3 presents a benchmark comparison between the proposed method and several recent studies on plant leaf disease classification. Although the proposed model does not achieve accuracy levels as high as studies evaluated on large-scale or highly controlled datasets such as PlantVillage, it demonstrates competitive performance given the limited size and high visual complexity of the PlantifyDR dataset. Unlike ensemble-based or hybrid transformer approaches that require substantial computational resources, the proposed method focuses on optimizing data diversity through a structured multi-stage augmentation pipeline while maintaining a single CNN architecture. This comparison highlights that the performance gain achieved in this study is primarily attributed to the augmentation strategy rather than architectural complexity. Therefore, the proposed approach offers a balanced trade-off between performance improvement and methodological simplicity, making it particularly suitable for local and resource-constrained agricultural applications

CONCLUSION

Based on the research results, it can be concluded that applying multi-stage data augmentation to the Inception V3 model significantly improves the classification performance of local plant leaf diseases compared to the baseline model without optimization. The proposed model demonstrated substantial improvements in all evaluation metrics, including accuracy, precision, recall, F1-score, and class separation ability, as indicated by AUC values reaching 1.00 for all classes. Accuracy, loss, ROC, and confusion matrix graphs consistently demonstrated that the proposed model learned more effectively, reduced the risk of overfitting, and had significantly better generalization to the test data. In contrast, the baseline model demonstrated unstable performance, lower accuracy, and difficulty recognizing visual patterns on a limited dataset. Therefore, this study confirms that a multi-stage augmentation strategy is an effective and important approach to deep learning model optimization, especially when applied to small and complex datasets such as images of local plant leaf diseases. It has the potential to serve as a benchmark for developing more accurate and reliable AI-based plant disease detection systems.

REFERENCE

[1] V. Rajasekhar, G. Arulselvi, and K. Suresh Babu, “Detection of diseases in rice leaf using convolutional neural network with transfer learning based on ResNeXt,” *Int. J. Electr. Comput. Eng.*, vol. 14, no. 2, pp. 1739–1749, 2024, doi: 10.11591/ijece.v14i2.pp1739-1749.

[2] M. S. H. Shovon, S. J. Mozumder, O. K. Pal, M. F. Mridha, N. Asai, and J. Shin, “PlantDet: A Robust Multi-Model Ensemble Method Based on Deep Learning For Plant Disease Detection,” *IEEE Access*, vol. 11, no. March, pp. 34846–34859, 2023, doi: 10.1109/ACCESS.2023.3264835.

[3] V. Saxena, “Utilizing Support Vector

Machines for Early Detection of Crop Diseases in Precision Agriculture a Data Mining Perspective," *Int. J. Intell. Syst. Appl. Eng.*, vol. 12, no. 16, pp. 281–288, 2024.

[4] A. Comander, B. Frates, and M. Tollefson, "PAVING the Path to Wellness for Breast Cancer Survivors: Lifestyle Medicine Education and Group Interventions," *Am. J. Lifestyle Med.*, vol. 15, no. 3, pp. 242–248, 2021, doi: 10.1177/1559827620986066.

[5] C. Dewi, F. Y. Bilaut, H. J. Christanto, and G. Dai, "Deep Learning for the Classification of Rice Leaf Diseases Using YOLOv8," *Math. Model. Eng. Probl.*, vol. 11, no. 11, pp. 3025–3034, 2024, doi: 10.18280/mmep.111115.

[6] H. Wang, "Phytotoxicity of Chemical Compounds from Cinnamomum camphora Pruning Waste in Germination and Plant Cultivation," *Int. J. Environ. Res. Public Health*, vol. 19, no. 18, 2022, doi: 10.3390/ijerph191811617.

[7] R. Poojary, "Effect of data-augmentation on fine-tuned cnn model performance," *Iaes Int. J. Artif. Intell.*, vol. 10, no. 1, pp. 84–92, 2021, doi: 10.11591/ijai.v10.i1.pp84-92.

[8] A. García-Pérez, "CNN-based in situ tool wear detection: A study on model training and data augmentation in turning inserts," *J. Manuf. Syst.*, vol. 68, pp. 85–98, 2023, doi: 10.1016/j.jmsy.2023.03.005.

[9] S. Zhong, "Molecular image-convolutional neural network (CNN) assisted QSAR models for predicting contaminant reactivity toward OH radicals: Transfer learning, data augmentation and model interpretation," *Chem. Eng. J.*, vol. 408, 2021, doi: 10.1016/j.cej.2020.127998.

[10] X. Xu, "Sleep Stage Classification With Multi-Modal Fusion and Denoising Diffusion Model," *IEEE J. Biomed. Heal. Informatics*, pp. 1–12, 2024, doi: 10.1109/JBHI.2024.3422472.

[11] R. Josphineleela, "A Multi-Stage Faster RCNN-Based iSPLInception for Skin Disease Classification Using Novel Optimization," *J. Digit. Imaging*, vol. 36, no. 5, pp. 2210–2226, 2023, doi: 10.1007/s10278-023-00848-3.

[12] J. Y. Wu, "A joint classification-regression method for multi-stage remaining useful life prediction," *J. Manuf. Syst.*, vol. 58, pp. 109–119, 2021, doi: 10.1016/j.jmsy.2020.11.016.

[13] K. Alomar, "Data Augmentation in Classification and Segmentation: A Survey and New Strategies," *J. Imaging*, vol. 9, no. 2, 2023, doi: 10.3390/jimaging9020046.

[14] J. J. Bird, "Chatbot Interaction with Artificial Intelligence: human data augmentation with T5 and language transformer ensemble for text classification," *J. Ambient Intell. Humaniz. Comput.*, vol. 14, no. 4, pp. 3129–3144, 2023, doi: 10.1007/s12652-021-03439-8.

[15] S. N. Yousafzai, I. M. Nasir, I. Keshta, and N. L. Fitriyani, "Multi-Stage Neural Network-Based Ensemble Learning Approach for Wheat Leaf Disease Classification," *IEEE Access*, vol. 13, no. February, pp. 30101–30116, 2025, doi: 10.1109/ACCESS.2025.3541347.

[16] W. H. Alwan and S. M. Alturfi, "Multi-Stage Vision Transformer and Knowledge Graph Fusion for Enhanced Plant Disease Classification," 2025, doi: 10.32604/csse.2025.064195.

[17] Q. Sun, "Multi-stage Co-planning Model for Power Distribution System and Hydrogen Energy System Under Uncertainties," *J. Mod. Power Syst. Clean Energy*, vol. 11, no. 1, pp. 80–93, 2023, doi: 10.35833/MPCE.2022.000337.

[18] H. Nozari, "A multi-stage stochastic inventory management model for transport companies including several different transport modes," *Int. J. Manag. Sci. Eng. Manag.*, vol. 18, no. 2, pp. 134–144, 2023, doi: 10.1080/17509653.2022.2042747.

[19] K. Wang, "Two-stage stochastic optimal scheduling for multi-microgrid networks with natural gas blending with hydrogen and low carbon incentive under uncertain envinronments," *J. Energy Storage*, vol. 72, 2023, doi: 10.1016/j.est.2023.108319.

[20] J. Wang, "Sensor Data Augmentation by Resampling in Contrastive Learning for Human Activity Recognition," *IEEE Sens. J.*, vol. 22, no. 23, pp. 22994–23008, 2022, doi: 10.1109/JSEN.2022.3214198.

[21] J. Zhang, "Data Augmentation and Dense-LSTM for Human Activity Recognition Using WiFi Signal," *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4628–4641, 2021, doi: 10.1109/JIOT.2020.3026732.

[22] R. Bai, "Rolling bearing fault diagnosis based on multi-channel convolution neural network and multi-scale clipping fusion data augmentation," *Meas. J. Int. Meas. Confed.*, vol. 184, 2021, doi: 10.1016/j.measurement.2021.109885.

[23] Q. Guo, "Data Augmentation for Intelligent Mechanical Fault Diagnosis Based on Local Shared Multiple-Generator GAN," *IEEE Sens. J.*, vol. 22, no. 10, pp. 9598–9609, 2022, doi:

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7845

10.1109/JSEN.2022.3163658.

[24] R. Li, "Sample-Based Data Augmentation Based on Electroencephalogram Intrinsic Characteristics," *IEEE J. Biomed. Heal. Informatics*, vol. 26, no. 10, pp. 4996–5003, 2022, doi: 10.1109/JBHI.2022.3185587.

[25] D. Wu, "Operando investigation of aqueous zinc manganese oxide batteries: multi-stage reaction mechanism revealed," *J. Mater. Chem. A*, vol. 11, no. 30, pp. 16279–16292, 2023, doi: 10.1039/d3ta01549a.

[26] F. Guo, "A machine learning method for prediction of remaining useful life of supercapacitors with multi-stage modification," *J. Energy Storage*, vol. 73, 2023, doi: 10.1016/j.est.2023.109160.

[27] J. Geng, "Investigation of dynamic response of drilling parameters and deformation characteristics of coal around borehole during multi-stage reaming in tectonic coal," *Int. J. Rock Mech. Min. Sci.*, vol. 170, 2023, doi: 10.1016/j.ijrmms.2023.105540.

[28] T. R. Mahesh, "Early Predictive Model for Detection of Plant Leaf Diseases Using MobileNetV2 Architecture," *Int. J. Intell. Syst. Appl. Eng.*, vol. 11, no. 2, pp. 46–54, 2023.

[29] N. U. A. Asan, "Antibacterial activity of Sireh (Piper betle L.) leaf extracts for controlling bacterial leaf blight diseases in rice plant," *Malays. J. Microbiol.*, vol. 18, no. 3, pp. 291–300, 2022, doi: 10.21161/mjm.221395.

[30] F. K. Sahlan, "The role of bacterial consortium as bioactivator to stimulate production and suppress grain rot disease and bacterial leaf blight in rice," *J. Trop. Plant Pests Dis.*, vol. 23, no. 2, pp. 65–70, 2023, doi: 10.23960/jhptt.22365-70.

[31] R. Parsibenehkohal, "A multi-stage framework for coordinated scheduling of networked microgrids in active distribution systems with hydrogen refueling and charging stations," *Int. J. Hydrogen Energy*, vol. 71, pp. 1442–1455, 2024, doi: 10.1016/j.ijhydene.2024.05.364.

[32] M. U. Tahir, "Overview of multi-stage charging strategies for Li-ion batteries," *J. Energy Chem.*, vol. 84, pp. 228–241, 2023, doi: 10.1016/j.jechem.2023.05.023.

[33] W. Wang, "Designing an energy-efficient multi-stage selective electrodialysis process based on high-performance materials for lithium extraction," *J. Memb. Sci.*, vol. 675, 2023, doi: 10.1016/j.memsci.2023.121534.

[34] H. Li, "Acoustic emission characteristics of rock salt under multi-stage cyclic loading," *Int. J. Fatigue*, vol. 176, 2023, doi: 10.1016/j.ijfatigue.2023.107911.