JURNAL TEKNOLOGI REKAYASA INFORMASI DAN KOMPUTER
E-ISSN: 2797-1724, DOI: 10.30811/jtrik.v9i1.8865
Vol. 9 No. 1 2026 pp. 66 -73

Deep Learning–Based Phenotype Classification of Arabidopsis thaliana from Top-View Imagery

Fauzan Azima[1], Muhammad Farhan[2] , Zata Hilya Syauqina[3*]

[1,2] Program Studi Teknik Informatika, Jurusan Teknologi Informasi dan komputer, Politeknik Negeri Lhokseumawe, Indonesia
[3] Program Studi Sistem Informasi, Fakultas Sains dan Teknologi, UIN Syarif Hidayatullah Jakarta, Indonesia

*Corresponding Author: zata.hilya@mhs.uinjkt.ac.id

Article info: Received 10/01/2026, yyyy, Revised 20/02/2026, Accepted 28/02/2026

Abstract

The development of digital image processing and machine learning enables automated and objective plant phenotyping, reducing reliance on manual observations that are time-consuming and subjective. This study aims to classify Arabidopsis thaliana leaf conditions into three classes, namely Healthy, Senescent, and Anthocyanin-Rich, using a Convolutional Neural Network (CNN) based on top-view images from the public Quantitative Plant and Zenodo datasets. A total of 1,500 images were used, representing diverse variations in leaf color, pigmentation levels, and visual conditions. The images were processed through several preprocessing stages, including resizing, pixel normalization, data augmentation, and stratified dataset splitting to maintain class balance. A custom CNN model was developed and trained to automatically extract visual features from leaf images, and its performance was evaluated using accuracy, confusion matrix, precision, recall, and F1-score metrics. Experimental results indicate that the model achieved an overall accuracy of 82%, with the best performance observed in the Healthy and Senescent classes. However, the Anthocyanin-Rich class still exhibited classification errors due to visual similarities with other classes. These findings demonstrate the potential of CNN-based approaches to support automated plant phenotyping, although further improvements are required to enhance model generalization and classification accuracy for visually similar classes.

Keywords: Leaf, arabidopsis thaliana, classification, phenotyping, deep learning

1. Introduction

The development of digital image processing and machine learning technologies has significantly advanced plant phenotyping, the process of automatically quantifying plant morphological and physiological characteristics using digital images. This approach allows researchers to objectively, consistently, and repeatedly analyze various important plant parameters, such as leaf color, surface shape and texture, leaf area, growth patterns, and aging rates. Compared with manual observation methods, digital image-based analysis offers advantages in terms of time efficiency, data scalability, and reduction of subjective bias that often arises from differences in observer perception. One of the most widely used model plant species in phenotyping studies is Arabidopsis thaliana, due to its relatively small size, short life cycle, and well-mapped genome, making it ideal for research into plant genetics, physiology, and molecular biology. However, the rapid development of genomics and DNA sequencing technologies has not been matched by the speed of visual phenotypic data acquisition and analysis. This imbalance is known as the phenotyping bottleneck, a limitation in efficiently acquiring, processing, and analyzing plant visual data that ultimately can hinder a comprehensive understanding of the relationship between plant genotype and phenotype [1].

To overcome the phenotyping bottleneck, various studies have developed automated approaches based on image processing and deep learning, particularly by leveraging the Convolutional Neural Network (CNN) architecture. CNNs are known to have excellent capabilities for automatically extracting complex visual features from plant images, enabling them to detect, classify, and count plant leaves accurately from top-view images. One important study in this field is ARADEEPOPSIS, which introduced an automatic semantic segmentation pipeline to classify Arabidopsis thaliana leaves into three main physiological states: green (healthy), anthocyanin-rich, and senescent. The model developed in this study achieved a mean Intersection over Union (mIoU) of 83.1%, indicating a high ability to map leaf physiological conditions from top-view images precisely. In addition, another study by Yuan et al. proposed an Optimized Multi-Task Learning

(MTL) approach that enables multiple phenotype analysis tasks to be performed simultaneously, including genotype classification, leaf number prediction, and leaf area estimation, with very high accuracy. This approach shows that deep learning is not only effective for a single type of visual task but also capable of representing plant phenotypic characteristics in a multidimensional manner. On the other hand, the LC-Net model was developed using two inputs: original images and leaf segmentation results to improve leaf-counting accuracy. This model has been tested on the CVPPP and KOMATSUNA datasets, with evaluation results showing a Dice score of 95.04% and an Intersection over Union (IoU) of 90.58%, confirming the effectiveness of CNNs in plant phenotype analysis [2], [3].

Based on these studies, digital image processing combined with deep learning methods plays a crucial role in the analysis and classification of modern plant phenotypes. The integration of leaf segmentation techniques, color and texture feature extraction, and the automatic learning capabilities of CNNs opens up great opportunities for accurately and efficiently detecting plant physiological conditions, including senescence and anthocyanin pigment accumulation, which often have overlapping visual characteristics. Although various approaches have shown promising results, distinguishing classes with high visual similarity remains a challenge that requires further study. Therefore, this study discusses the application of CNN-based digital image processing techniques to the public Arabidopsis thaliana dataset from Quantitative Plant, focusing on three main physiological categories: Healthy, Senescent, and Anthocyanin-Rich. This approach is expected to contribute to the development of a more efficient, objective, and reliable plant phenotype analysis system and to support further research in plant biology and precision agriculture.

Furthermore, the use of public datasets in plant phenotype research is crucial to ensure the reproducibility and validity of research results. Datasets such as Quantitative Plant and Zenodo provide top-view images of Arabidopsis thaliana across a wide variety of physiological, lighting, and environmental conditions, thereby better reflecting the plant's actual condition. However, this visual diversity also poses challenges for classification, especially for classes with overlapping visual characteristics, such as Senescent and Anthocyanin-Rich. Therefore, a classification approach capable of extracting visual features in-depth and adaptively is needed. The use of a CNN (custom model) specifically trained on top-view images is expected to capture visual patterns relevant to the leaf's physiological condition. Thus, this study not only contributes to the application of deep learning methods in plant phenotyping but also provides an overview of the challenges and potential of developing a digital image-based Arabidopsis thaliana leaf condition classification system to support automatic and sustainable plant phenotype analysis [4].

2. Method

The dataset used in this study consists of top-view images of Arabidopsis thaliana obtained from the Quantitative Plant dataset and the Zenodo repository. The dataset includes plant images representing three distinct physiological conditions: Healthy, Senescent, and Anthocyanin-rich. All images are utilized as input for training the convolutional neural network (CNN) model through a structured pipeline comprising data preprocessing, dataset partitioning, and model training stages [5], [6].

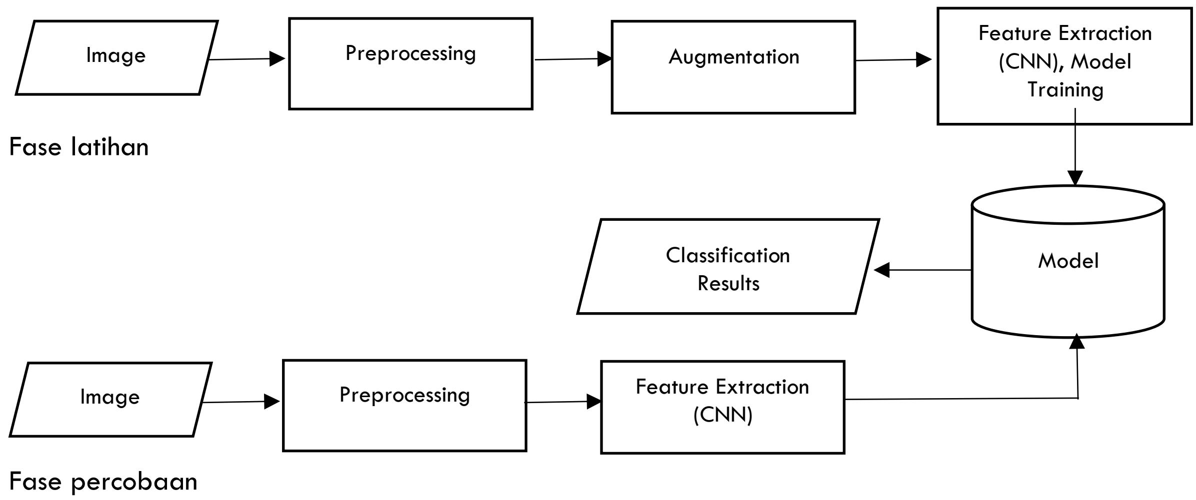


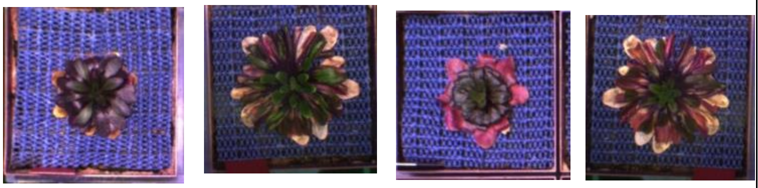
Figure 1. Proposed Method

2.1 Datasets

This study employs an Arabidopsis thaliana image dataset obtained from the public repositories Quantitative Plant and Zenodo. The dataset consists of top-view images representing three physiological leaf conditions: Healthy, Senescent, and Anthocyanin-Rich. The images exhibit substantial variability in lighting conditions, plant size, pigmentation levels, plant age, and other visual leaf characteristics. Such diversity makes the dataset well suited for convolutional neural network (CNN)-based classification tasks. In total, the dataset comprises 1,500 Arabidopsis thaliana leaf images, all captured from a top-view perspective. The images reflect a wide range of environmental conditions, including variations in light intensity and growing medium texture, which further enhance the visual heterogeneity of the dataset. This variability is essential for training a robust CNN model capable of generalizing across diverse real-world scenarios [7].

Each image was manually labeled according to the plant's physiological leaf condition: Healthy, Senescent, or Anthocyanin-Rich. The class distribution includes 800 Healthy images, 600 Senescent images, and 100 Anthocyanin-Rich images, resulting in an imbalanced dataset that reflects natural occurrence patterns. Prior to model training, all images underwent a preprocessing pipeline consisting of resizing to 256 × 256 pixels, pixel value normalization, and label verification based on visual criteria documented in the Arabidopsis thaliana phenotype literature. Table 1 presents representative visual examples from each physiological class in the dataset. The first row (No. 1) illustrates a Healthy plant, characterized by bright green leaves and an overall vigorous appearance. The second row (No. 2) shows Senescent plants, which typically exhibit yellowing or reddish leaf coloration due to chlorophyll degradation, along with signs of aging or wilting. The third row (No. 3) depicts Anthocyanin-Rich plants, distinguished by pronounced purple or dark red pigmentation resulting from anthocyanin accumulation in response to specific environmental stressors. The visual diversity across examples—spanning differences in lighting, background, and plant size—highlights the complexity of the dataset and its suitability for training a CNN model for robust physiological classification.

Table 1 Example of image on each class in Aradopsis thaliana datasets

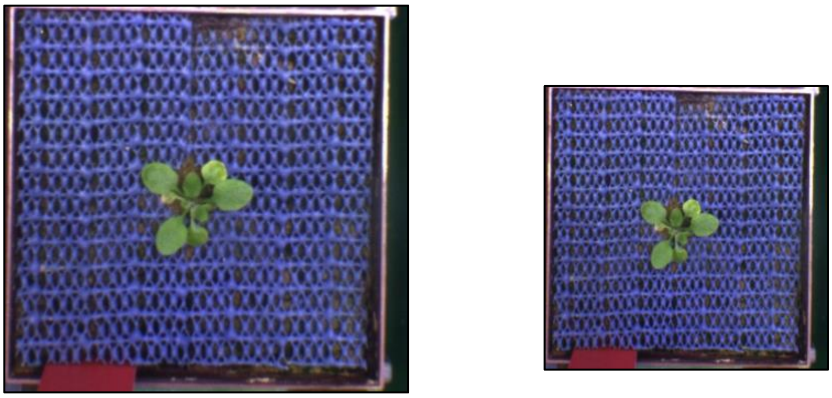
No	Images	Class
1		*Healthy*
2		*Senescent*
3		*Anthocyanin-Rich*

2.2 Preprocessing

The preprocessing stage plays a crucial role in preparing image data prior to the deep learning–based model training process. This stage aims to generate cleaner and more consistent input data that conform to the requirements of convolutional neural network (CNN) architectures. In contemporary research, effective preprocessing is widely recognized as a key factor influencing model stability, convergence, and generalization performance. The Arabidopsis thaliana dataset used in this study comprises three physiological leaf categories: Healthy, Senescent, and Anthocyanin-Rich. To ensure high data quality and robust model performance, five preprocessing steps were applied: image resizing, pixel value normalization, data augmentation, dataset partitioning, and label encoding [8].

- Resize

Image resizing is performed to standardize the dimensions of all images so that they conform to the input size requirements of the CNN model. Image classification architectures typically require fixed input resolutions, such as 224 × 224 or 256 × 256 pixels, depending on the specific network design. Recent studies indicate that resizing images to a uniform resolution not only accelerates the training process but also preserves consistent spatial and visual structures across samples, thereby contributing to more stable and effective model learning [9].



(a) Before resize (b) After Resize

Figure 2. Exampe of Arabidopsis image before and after resize

- Pixel Normalization

Normalization is applied to adjust the scale of pixel values in order to enhance the stability of the backpropagation process during model training. Raw RGB pixel intensities in the range of 0–255 can lead to unstable gradients if used directly. To address this issue, pixel values are normalized to a standardized range of 0–1. Intensity-scale–based normalization is widely adopted in contemporary computer vision research, as it facilitates faster convergence and contributes to more stable and efficient training of deep learning models [10].

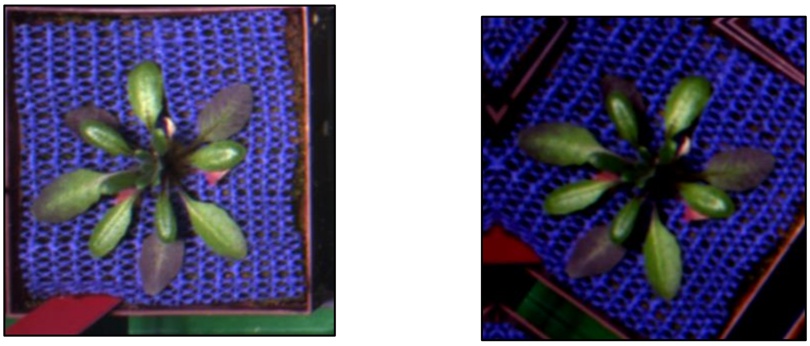
- Image Augmentation

Data augmentation is employed to enhance the visual diversity of the training dataset, particularly in situations where class distributions are imbalanced. Transformations such as rotation, horizontal and vertical flipping, and variations in light intensity enable the model to learn more robust and invariant feature representations under diverse environmental conditions. Previous studies in plant image analysis have demonstrated that data augmentation can significantly improve model performance, especially when dealing with limited datasets or underrepresented classes [11], [12].

Augmentation techniques applied include:

a. Random rotation ($\pm 20°$)
b. Zoom in/out
c. Horizontal and vertical flip
d. Shifting (shifting object position)
e. Light brightness adjustment

These augmentation techniques were selected because Arabidopsis thaliana leaves do not exhibit a fixed orientation; therefore, the applied transformations do not alter the underlying class labels.



(a) Before Augmentation (b) After Augmentation

Figure 3. Exampe of Arabidopsis image before and after augmentation

2.3 Dataset Splitting

The dataset was partitioned into three subsets: training, validation, and testing. This partitioning strategy ensures a clear separation between data used for model learning, performance monitoring during training, and final evaluation. A stratified splitting approach was employed to preserve the original class distribution across all subsets. This method is widely recommended in classification research, as it helps reduce class imbalance bias and provides a more reliable assessment of model performance [13].

In this study, the distribution proportions are:

a. 70% training data
b. 20% validation data
c. 10% testing data

2.3 Convolutional Neural Network (CNN)

Following the preprocessing stage, this study employed a Convolutional Neural Network (CNN) to classify the physiological conditions of Arabidopsis thaliana leaves. CNN was selected due to its proven capability in learning complex visual patterns from plant imagery and its widespread application in Arabidopsis phenotype research. For instance, CNN-based approaches have been successfully utilized for silique detection in the DeepPod model and in high-throughput phenotyping systems that analyze top-view images to measure plant traits. Furthermore, deep learning studies involving fluorescence imaging have demonstrated that CNN architectures are highly effective in extracting intricate physiological features. Therefore, the use of CNN in this study is both methodologically appropriate and consistent with advancements in contemporary Arabidopsis research [14], [15].

The model architecture employed in this study was custom designed and trained from scratch. This approach is consistent with contemporary plant phenotyping pipelines that utilize lightweight CNN architectures to analyze Arabidopsis thaliana leaf characteristics. The network begins with a Conv2D layer using a 3×3 kernel and ReLU activation, which enables the extraction of low-level leaf features such as surface texture and physiological color variations. Similar CNN-based feature extraction strategies have been adopted in multi-task learning studies aimed at analyzing leaf count and rosette area in Arabidopsis. Subsequently, a MaxPooling2D layer is applied to reduce spatial dimensionality while retaining the most salient features. The resulting feature maps are then flattened using a Flatten layer and passed to a fully connected Dense layer with 128 neurons for higher-level feature representation. To enhance model generalization and mitigate overfitting, a Dropout layer with a rate of 0.4 is incorporated. This regularization strategy aligns with approaches used in Mask R-CNN–based leaf tracking and segmentation models. Finally, the output layer employs a Softmax activation function with three neurons to generate class probability distributions corresponding to the Healthy, Senescent, and Anthocyanin-Rich leaf conditions. The use of Softmax for multi-class classification is consistent with its application in high-throughput Arabidopsis trait estimation systems, such as APTES [15], [16]. The model compilation result show in Figure 4.

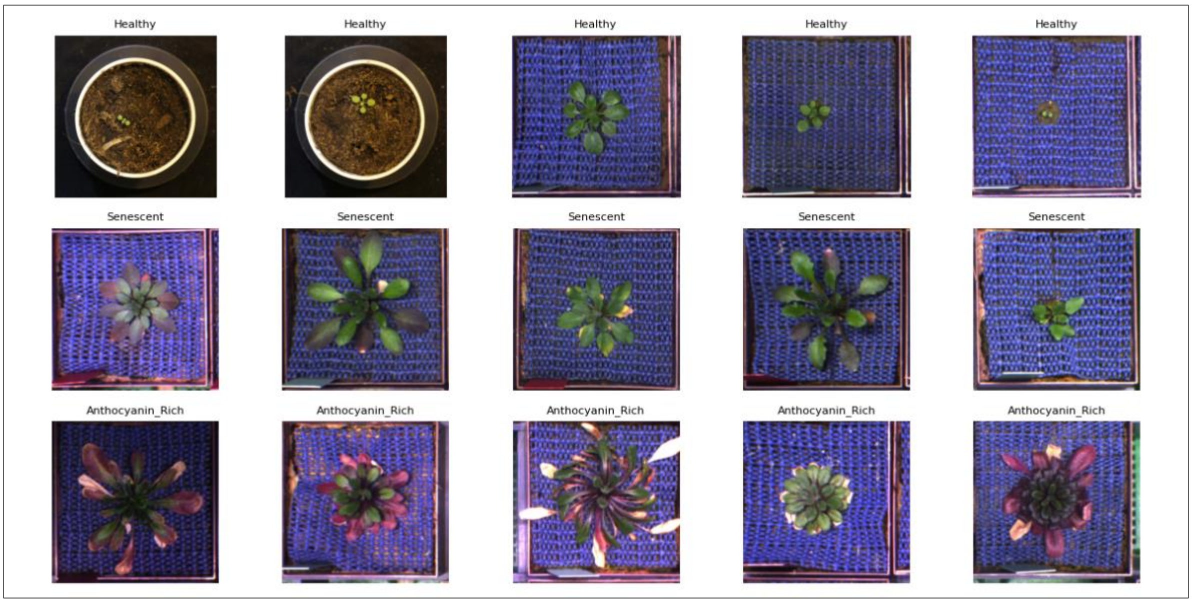
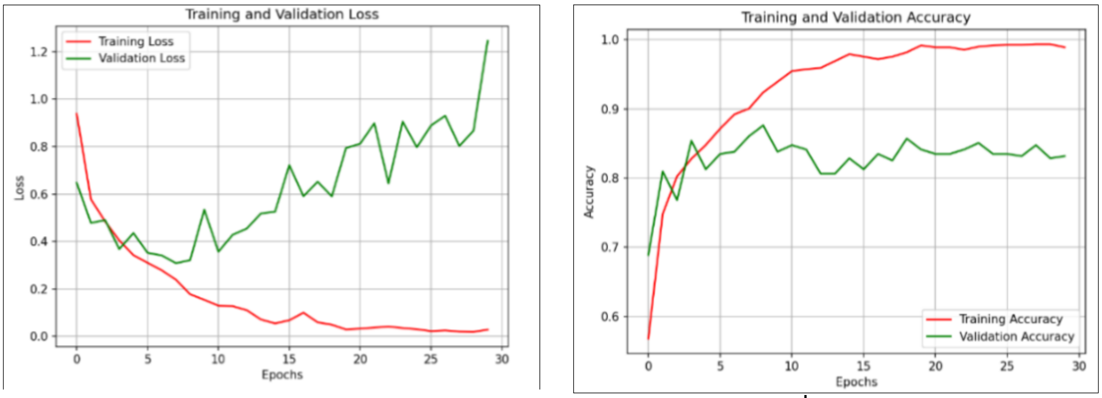


Figure 4. Compilation Result

3. Results

3.1 Model Performance

The performance of the CNN model was evaluated by analyzing the trends in loss and accuracy values throughout the training process. Overall, the model exhibited strong learning capability on the training data, as indicated by a steady decline in training loss from a high initial value to nearly zero by the final epochs. In contrast, a different trend was observed for the validation loss, which decreased during the early training stages but subsequently fluctuated and increased markedly in the later epochs. This divergence suggests the presence of overfitting, a condition in which the model becomes excessively tailored to the training data, leading to diminished generalization performance on unseen validation samples.



a. Loss b. accuracy

Figure 5. loss dan Accuracy Results

Figure 5 illustrates the trends of loss and accuracy values observed during the training process. As shown in Figure 5(a), the training loss decreases sharply over the epochs, whereas the validation loss begins to rise toward the final epochs, providing a clear indication of overfitting. Similarly, Figure 5(b) demonstrates that training accuracy increases consistently, while validation accuracy shows stagnation and noticeable fluctuations. Despite this behavior, the model achieved a test accuracy of 0.82, indicating that it is still capable of classifying the physiological conditions of Arabidopsis thaliana leaves with a reasonably good level of accuracy. A comparable pattern is evident in the accuracy curves, where training accuracy steadily increases to nearly 0.99, while validation accuracy remains within the range of 0.82–0.87 and does not exhibit significant improvement. These findings suggest that although the model effectively learns patterns from the training data, its ability to generalize to unseen data remains limited and requires further improvement.

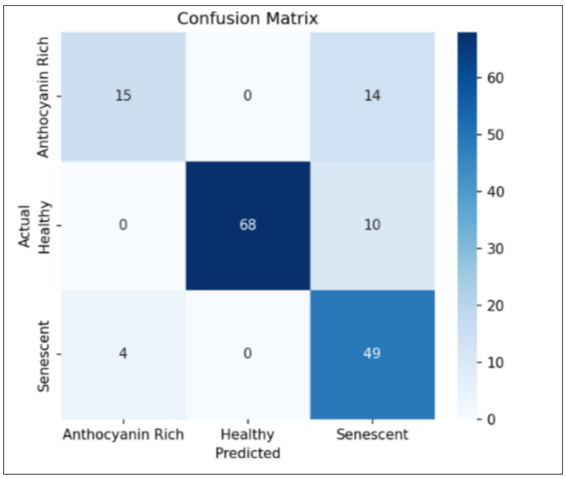


Figure 6. Confusion Matrix

Based on the confusion matrix shown in Figure 6, the model demonstrates strong classification performance across the three Arabidopsis thaliana leaf classes: Anthocyanin-Rich, Healthy, and Senescent. The Healthy class achieved the highest accuracy, with 68 samples correctly classified, indicating that the visual characteristics of healthy leaves are more readily distinguishable than those of the other two classes. For the Senescent class, the model also performed well, correctly identifying 49 samples, although 4 samples were misclassified as Anthocyanin-Rich. In contrast, the Anthocyanin-Rich class proved to be the most challenging to differentiate, as 14 samples were incorrectly predicted as Senescent. These misclassifications suggest overlapping visual features—particularly in terms of color and texture—between leaves undergoing anthocyanin accumulation and those entering the senescence phase. Overall, the confusion matrix indicates that the model is capable of accurately recognizing all three classes; however, distinguishing between classes with similar visual characteristics remains a challenge and warrants further refinement.

Table 2. Classification Report

Class	Precision	Recall	F1-Score	Support
Anthocyanin Rich	0.79	0.52	0.62	29
Healthy	1.00	0.87	0.93	78
Senescent	0.67	0.92	0.78	53
Accuracy	-	-	0.82	160

As reported in Table 2, the CNN model demonstrates class-dependent performance variations. For the Anthocyanin-Rich class, a precision of 0.79 indicates a relatively high proportion of correct positive predictions, whereas the lower recall of 0.52 reveals limited sensitivity in detecting all anthocyanin-rich samples, leading to a substantial number of false negatives. The Healthy class achieves the strongest performance, with a precision of 1.00 and a recall of 0.87, reflecting highly reliable and consistent classification. For the Senescent class, the model attains a precision of 0.67 and a recall of 0.92, suggesting that although misclassifications remain in the form of false positives, the model effectively captures the majority of true senescent samples. Overall, these findings indicate robust performance for the Healthy and Senescent classes, while highlighting the need for further refinement to improve discrimination of the Anthocyanin-Rich class, which exhibits overlapping visual features with other physiological conditions.

4. Conclusions

A Convolutional Neural Network (CNN) applied to top-view images of Arabidopsis thaliana successfully classified leaf physiological conditions into three categories: Healthy, Senescent, and Anthocyanin-Rich, achieving a satisfactory level of performance. The preprocessing pipeline—including image resizing, pixel normalization, data augmentation, and stratified dataset partitioning—played a critical role in ensuring data consistency and supporting effective model learning. The model exhibited strong classification capability for the Healthy and Senescent classes, as reflected by high precision and recall values, indicating that the visual features of these classes were well captured by the network. However, the evaluation results also revealed limitations in distinguishing the Anthocyanin-Rich class from the Senescent class. This limitation was evidenced by a lower recall value and notable misclassification in the confusion matrix, primarily due to overlapping visual characteristics related to leaf color changes. Additionally, the training loss and accuracy trends indicated the presence of overfitting, as the model achieved substantially higher performance on the training data compared to the validation data. Overall, while CNNs are effective for extracting visual features from Arabidopsis thaliana leaf images, further improvements—such as addressing dataset imbalance, optimizing network architecture, and incorporating stronger regularization strategies—are required to enhance generalization performance on more diverse and complex datasets.

REFERENCE

[1] P. Huther, N. Schandry, and others, "ARADEEPOPSIS An automated workflow for top-view plant phenomics using semantic segmentation of leaf states," *Plant Cell*, vol. 32, no. 12, pp. 3674–3688, 2020, doi: 10.1093plcellkoaa255.

[2] P. Yuan, S. Xu, Z. Zhai, and H. Xu, "Research of intelligent reasoning system of Arabidopsis thaliana phenotype

based on automated multi-task machine learning," *Front. Plant Sci.*, vol. 14, 2023, doi: 10.3389/fpls.2023.1048016.

[3] M. Deb, K. G. Dhal, A. Das, A. E. G. Hussien, L. Abualigah, and A. Garai, "A CNN-based model to count the leaves of rosette plants (LC-Net)," *Sci. Rep.*, vol. 14, p. 1496, 2024, doi: 10.1038s41598-024-51983-y.

[4] A. S. Yuksel, A. H. Zaim, and M. A. Aydin, "A Comprehensive Analysis of Android Security and Proposed Solutions," *Int. J. Comput. Netw. Inf. Secur.*, vol. 6, no. 12, pp. 9–20, 2014, doi: 10.5815/ijcnis.2014.12.02.

[5] J. A. Cruz, D. P. D. Yin, J. D. Sandhu, and J. Chen, "Arabidopsis thaliana top-view annotated dataset for phenotyping." Zenodo, 2021. doi: 10.5281/zenodo.5141337.

[6] M. A. Dobrescu, M. V Giuffrida, and S. A. Tsaftaris, "Understanding deep neural networks for plant phenotyping with saliency maps," *Plant Methods*, vol. 16, pp. 1–13, 2020.

[7] B. Jiang, H. Fan, and X. Ma, "Plant Phenotyping Based on Deep Learning: A Review of Recent Advances," *Front. Plant Sci.*, vol. 11, p. 1122, 2020, doi: 10.3389/fpls.2020.01122.

[8] R. Barth, J. IJsselmuiden, J. Hemming, and E. J. van Henten, "Data pre-processing methods for high-performance deep learning in plant phenotyping," *Plant Phenomics*, vol. 2021, pp. 1–12, 2021, doi: 10.34133/2021/9846325.

[9] S. Bhugra and P. Kaur, "Deep learning-based leaf classification: Effectiveness of resizing and normalization strategies," *Ecol. Inform.*, vol. 70, p. 101732, 2022, doi: 10.1016/j.ecoinf.2022.101732.

[10] L. Zhang, Y. Chen, and H. Wang, "Pixel normalization strategies for improving CNN training on biological images," *Pattern Recognit. Lett.*, vol. 150, pp. 90–97, 2021, doi: 10.1016/j.patrec.2021.06.019.

[11] M. A. M. Haque, M. Nejad, and F. Hossain, "Improved plant classification using enhanced data augmentation strategies," *Appl. Sci.*, vol. 13, no. 4, p. 2114, 2023, doi: 10.3390/app13042114.

[12] A. Alharbi, S. Alqahtani, and A. Alanazi, "Deep learning-based plant disease detection using image augmentation techniques," *Comput. Electron. Agric.*, vol. 198, p. 107065, 2022, doi: 10.1016/j.compag.2022.107065.

[13] M. Rahnemoonfar, S. Gragg, and J. Li, "Deep learning-based image classification in plant phenotyping under data imbalance," *Front. Plant Sci.*, vol. 12, p. 662578, 2021, doi: 10.3389/fpls.2021.662578.

[14] A. Hamidinekoo and others, "DeepPod: A convolutional neural network based quantification of fruit number in Arabidopsis," *Gigascience*, vol. 9, no. 3, 2020, doi: 10.1093/gigascience/giaa012.

[15] J. Li, J. Peng, X. Jiang, A. C. Rea, and J. Hu, "DeepLearnMOR: A deep-learning framework for fluorescence image-based classification of organelle morphology," *Plant Physiol.*, vol. 186, no. 4, pp. 1786–1799, 2021, doi: 10.1093/plphys/kiab223.

[16] A. González-Muñoz and others, "A high-throughput pipeline for phenotyping, object detection and quantification of leaf trichomes," *Theor. Appl. Genet.*, vol. 138, no. 8, 2025, doi: 10.1007/s00122-025-04967-z.

[17] F. Jurado-Ruiz, T. P. Nguyen, J. Peller, M. J. Aranzana, G. Polder, and M. G. M. Aarts, "LeTra: a leaf tracking workflow based on convolutional neural networks and intersection over union," *Plant Methods*, vol. 20, no. 1, 2024, doi: 10.1186/s13007-024-01138-x.