A Comparative Study of Convolutional Neural Networks (CNNs) and Traditional Machine Learning Methods for Multilabel Plant Pathology Classification

Mohammed *254735@student.pwr.edu.pl*
Robert Kanimba *276884@student.pwr.edu.pl*
Yusuf *277034@student.pwr.edu.pl*
Wrocław University of Science and Technology

***Abstract*—This study presents a comprehensive comparison between Convolutional Neural Networks (CNNs) and traditional machine learning algorithms, namely Random Forest, K-Nearest Neighbors, and Gaussian Naive Bayes, in the realm of multilabel plant pathology classification. Utilizing the Plant Pathology 2020 FGVC7 dataset from Kaggle [12], our research aims to evaluate the effectiveness and efficiency of deep learning and classical machine learning approaches in handling complex multilabel classification tasks. The dataset comprises high-resolution images of plant leaves, labeled with multiple disease states, providing a challenging platform for model evaluation. We employed a ResNet18-based CNN model and compared its performance against the aforementioned traditional algorithms, using metrics such as accuracy, F1 score, training time, and prediction time. Our results indicate a significant performance disparity between CNNs and traditional methods, highlighting the strengths and limitations of each approach in dealing with multilabel classification problems in agricultural domains. The findings of this study contribute valuable insights into the applicability and efficiency of different machine learning paradigms in plant pathology, offering guidance for future research and practical implementations in agricultural technology.**

***Index Terms*—Convolutional Neural Networks, Machine Learning, Plant Pathology, Image Classification, Multilabel Classification, Agricultural Technology, Random Forest, K-Nearest Neighbors, Naive Bayes, Plant Diseases.**

I. Introduction

Plant diseases pose a significant threat to global food security, affecting crop yield and quality. Early and accurate disease detection is crucial for effective plant disease management. In recent years, the advent of machine learning, particularly Convolutional Neural Networks (CNNs), has revolutionized image-based plant pathology. This study explores the efficacy of CNNs in comparison to traditional machine learning techniques - Random Forest, K-Nearest Neighbors, and Gaussian Naive Bayes - in multilabel classification tasks within the domain of plant pathology.

The application of machine learning in agriculture, especially for disease detection, offers potential for automation, accuracy, and efficiency in monitoring crop health. While CNNs have shown promise in image recognition and classification tasks, traditional machine learning methods remain widely used due to their simplicity and lower computational requirements. This research seeks to provide a comprehensive comparison between these approaches, using the Plant Pathology 2020 FGVC7 dataset from Kaggle, which contains high-resolution images of leaves annotated with various disease labels.

The objective is to assess not only the classification accuracy but also the practical aspects such as training time and prediction efficiency, which are crucial for real-world applications. The findings aim to contribute to the growing body of knowledge in agricultural technology and provide insights into the selection of appropriate machine learning techniques for plant disease classification.

II. Related Work

The automated classification of plant diseases has become increasingly reliant on advanced deep learning models due to their high accuracy and robust feature extraction capabilities. Atila et al. [1] explored the EfficientNet deep learning model for plant leaf disease classification, demonstrating its superior performance on the PlantVillage dataset when compared with other deep learning models. They highlighted the advantage of transfer learning and the significance of using augmented datasets to improve model accuracy and precision.

The review by Li et al. [2] provides comprehensive insights into the application of deep learning for plant disease recognition, emphasizing the benefits of automatic feature extraction that deep learning offers over traditional manual selection methods. This work underscores the potential for increasing research efficiency and speeding up the technology transfer process in agricultural plant protection.

Abade et al. [3] conducted a systematic review of CNNs in identifying and classifying plant diseases. Their work encapsulates the last decade's research trends and identifies gaps that require the scientific community's attention, highlighting the innovative use of CNNs for plant disease detection.

Nasser and Akhloufi [4] proposed a hybrid CNN-Transformer architecture, CTPlantNet, showcasing remarkable results in multi-classification of apple leaf diseases. Their model outperformed state-of-the-art models, providing evidence for the effectiveness of hybrid deep learning architectures in this domain.

Liu et al. [5] introduced a deep convolutional neural network approach to identify apple leaf diseases with a novel architecture based on AlexNet. Their model, trained on a large dataset, showed significant improvements in accuracy and model parameter efficiency.

The comprehensive work of MacHardy et al. [6], though not recent, provides valuable background on the use of fungicides for controlling apple scab, which remains relevant in understanding disease control's historical context.

Mahlein [7] discussed the use of various imaging sensors for plant disease detection, highlighting the specific requirements for precision agriculture and plant phenotyping. This study provides a foundation for understanding sensor-based methodologies alongside machine learning approaches.

Mwebaze et al. [8] described a dataset of cassava leaf images for a fine-grained visual categorization challenge aimed at improving semi-supervised learning algorithms for plant disease monitoring, reflecting the importance of high-quality datasets in advancing the field.

Sladojevic et al. [9] emphasized the potential of deep CNNs for recognizing plant diseases through leaf image classification, advocating for a novel approach to training and implementing disease recognition models that could be practically adopted.

Lastly, the hierarchical deep convolutional neural networks (HD-CNNs) introduced by Yan et al. [10], and the deep transfer learning approach with an enhanced lightweight network by Chen et al. [11], represent innovative directions in plant disease classification that leverage hierarchical structures and network modifications to improve recognition accuracy in large-scale datasets.

Our research aims to compare Convolutional Neural Networks and traditional algorithms on a multilabel plant pathology task. We focus on statistical analysis to evaluate the significance of their performance differences, guiding the selection of appropriate models for agricultural applications. This approach provides empirical evidence to support the use of specific machine learning techniques in plant disease detection.

III. Methodology

A. Dataset

The study utilizes the Plant Pathology 2020 FGVC7 dataset, sourced from Kaggle. This dataset comprises high-resolution images of plant leaves, annotated with multiple disease states such as healthy, multiple diseases, rust, and scab. Each image is labeled with one or more of these categories, making it a multilabel classification problem. The dataset is split into training and test sets, ensuring a representative distribution of each disease category. In figure 1, we present sample images from each disease category.

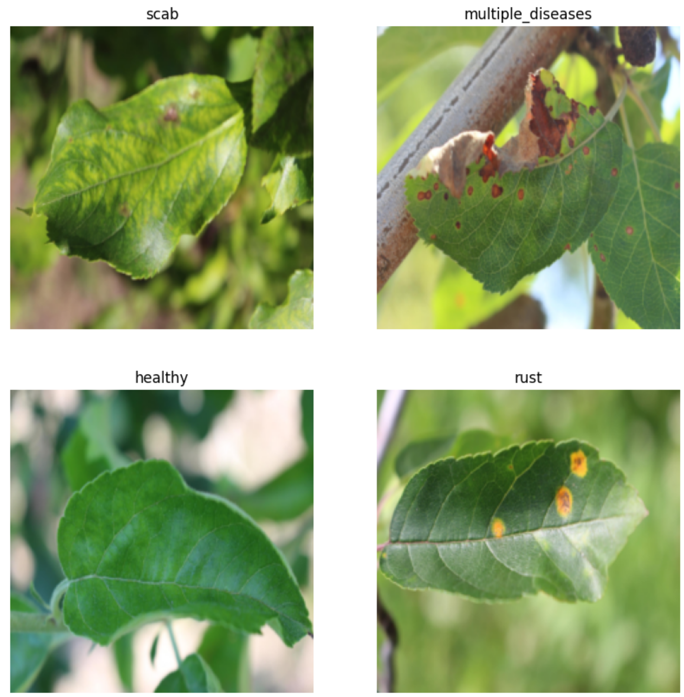

Fig. 1: Dataset Example Images

B. Machine Learning Models

1) Convolutional Neural Networks (CNNs): The deep learning model employed in this study is based on the ResNet18 architecture, a popular CNN model known for its efficiency and performance in image classification tasks. The model is pre-trained on ImageNet and fine-tuned on the plant pathology dataset. Adjustments to the final layer are made to accommodate the multilabel classification.

2) Traditional Machine Learning Methods: Three traditional machine learning algorithms are selected for comparison:

- **Random Forest Classifier:** An ensemble learning method known for its robustness and accuracy in classification tasks.
- **K-Nearest Neighbors (KNN) Classifier:** A non-parametric method used for its simplicity and effectiveness in classification problems.
- **Gaussian Naive Bayes:** A probabilistic classifier that applies Bayes' theorem with the assumption of independence among predictors.

Each model is trained and tested on the same dataset for a fair comparison.

C. Experimental Setup

The dataset is pre-processed before training. For the CNN, images are resized to 224x224 pixels, and standard image augmentations are applied to prevent overfitting. The traditional machine learning models use flattened image data as input. The dataset is divided into training and validation sets, with a stratified split to ensure each set is representative of the overall dataset.

D. Evaluation Metrics

Model performance is assessed using accuracy, F1 score, training time, and prediction time. Accuracy measures the proportion of correctly identified cases, while the F1 score provides a balance between precision and recall, crucial for multilabel classification. Training time and prediction time are recorded to evaluate the computational efficiency of each model.

The models' robustness is further validated using a Repeated Stratified K-Fold cross-validation strategy, with the dataset divided into 5 folds, repeated twice. This approach ensures a comprehensive evaluation across different subsets of the data.

IV. Results

The performance of the classifiers was evaluated based on four key metrics: F1 score, accuracy, training time, and prediction time. The F1 score and accuracy reflect the models' ability to correctly classify the multilabel images, while training and prediction times indicate the computational efficiency.

A. F1 Score

The Convolutional Neural Network-based PlantModel significantly outperformed all traditional machine learning models with an F1 score of 0.7802. The RandomForestClassifier achieved the second-highest mean F1 score of 0.3745, followed by GaussianNB and KNeighborsClassifier with scores of 0.3185 and 0.2986, respectively. A comparative analysis based on F1 scores is summarized in Table I.

B. Accuracy

Consistent with the F1 score results, the PlantModel achieved the highest accuracy of 0.9308. The RandomForestClassifier was the best among the traditional models with an accuracy of 0.4898, outperforming both KNeighborsClassifier (0.3655) and GaussianNB (0.3402).

C. Training Time

The training times varied significantly across models. The PlantModel required the most time to train, averaging 4867.18 seconds. In contrast, the traditional models were much faster, with KNeighborsClassifier being the quickest at 0.1593 seconds, followed by GaussianNB and RandomForestClassifier at 2.5569 and 100.5662 seconds, respectively.

D. Prediction Time

Regarding prediction time, the RandomForestClassifier was the fastest at 0.0608 seconds, whereas the PlantModel was the slowest, taking an average of 78.6216 seconds. The GaussianNB and KNeighborsClassifier had prediction times of 3.9236 and 5.0391 seconds, respectively.

E. Precision Scores and Matthews correlation coefficients

In figure 2, we present the precision scores for each classifier. The PlantModel (CNN) demonstrates superior precision scores compared to the traditional models around 0.8. RandomForestClassifier and GaussianNB have similar precision scores 0.4, while KNeighborsClassifier has the lowest precision score 0.3. In figure 3, we present the Matthews correlation coefficients for each classifier. The PlantModel (CNN) demonstrates superior Matthews correlation coefficients compared to the traditional models around 1.0. RandomForestClassifier has Matthews correlation coefficients 0.4, while KNeighborsClassifier and GaussianNB has the lowest Matthews correlation coefficients 0.1.

These results suggest that while the PlantModel delivers superior classification performance, it does so at the cost of increased computational resources and time. Conversely, traditional machine learning models, though not as accurate, are more computationally efficient.

TABLE I: Classifier Comparison Based on F1 Scores

Classifier	F1 Score	Accuracy	Prediction Time (s)
RandomForestClassifier	0.3745	0.4898	0.0608
KNeighborsClassifier	0.2986	0.3655	5.0391
GaussianNB	0.3185	0.3402	3.9236
PlantModel (CNN)	0.7802	0.9308	78.6216

Note: This table presents the mean F1 scores, accuracy, and prediction times for each classifier used in the study. The F1 score and accuracy provide a measure of the model's classification performance, while the prediction time indicates the model's speed during inference. The PlantModel, based on CNN, demonstrates superior F1 score and accuracy but requires longer prediction time compared to traditional models.

TABLE II: F1 Score ttest results

-	RFC	KNN	GNB	CNN
RFC	N/A	Better	Better	Worse
KNN	Worse	N/A	Worse	Worse
GNB	Worse	Worse	N/A	Worse
CNN	Better	Better	Better	N/A

Note: The table displays the pairwise comparison of classifiers based on their F1 scores using t-tests. "Better" indicates that the row classifier has a significantly higher F1 score than the column classifier, "Worse" indicates the opposite, and "NA" denotes that comparison is not applicable. The CNN-based PlantModel consistently outperforms the traditional classifiers, as indicated by the "Better" results in its row.

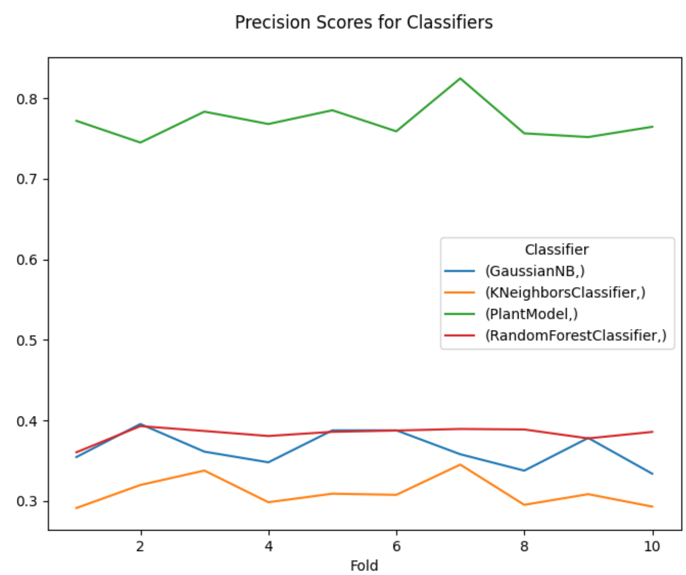

Fig. 2: precision scores
Note: The figure displays the precision scores for each classifier. The PlantModel (CNN) demonstrates superior precision scores compared to the traditional models.

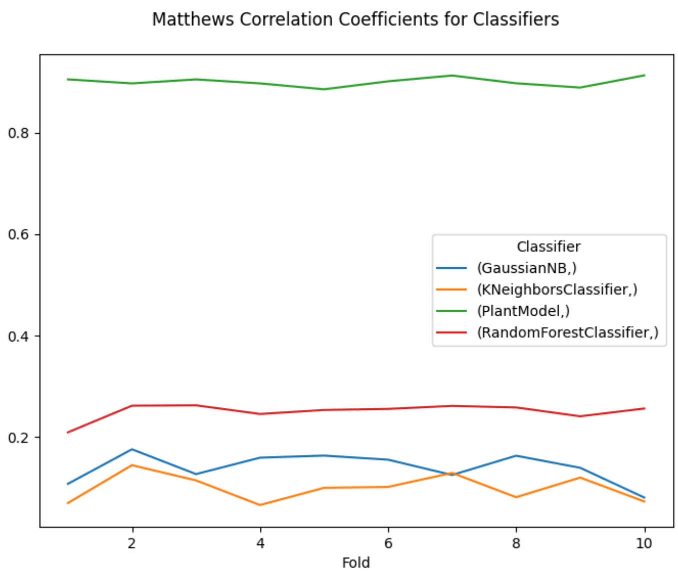

Fig. 3: Matthews correlation coefficients
Note: The figure displays the Matthews correlation coefficients for each classifier. The PlantModel (CNN) demonstrates superior Matthews correlation coefficients compared to the traditional models.

V. Discussion

The results of the comparative study reveal several key insights into the performance of convolutional neural networks (CNNs) relative to traditional machine learning methods in the context of plant pathology classification. The CNN-based PlantModel significantly outperformed traditional models in terms of F1 score and accuracy, corroborating the hypothesis that deep learning models are better suited for complex image-based classification tasks due to their ability to learn hierarchical feature representations.

The superior performance of the PlantModel can be attributed to its architecture, which is specifically designed to capture spatial hierarchies and patterns within images. These capabilities are crucial in plant pathology, where subtle visual cues can be indicative of disease presence or absence. In contrast, traditional models such as RandomForest, KNN, and GaussianNB lack these sophisticated feature extraction capabilities, leading to lower performance scores.

However, this high performance comes at the cost of computational efficiency. The PlantModel requires significantly more time for both training and prediction than the traditional algorithms. This trade-off highlights an important consideration for practical applications: the choice between accuracy and speed is context-dependent. In scenarios where real-time predictions are crucial, such as in automated systems for monitoring crop health, the increased prediction time of CNNs may be a limiting factor. On the other hand, for applications where accuracy is paramount, and time is less of a concern, the CNN's higher resource demand may be justifiable.

The RandomForestClassifier, while less accurate than the PlantModel, presents a good balance between performance and speed, marking it as a potential candidate for applications that require faster predictions but can tolerate a moderate reduction in accuracy.

It is also noteworthy that the performance of KNN and GaussianNB was lower than that of the other classifiers. This outcome may be due to the high dimensionality of the image data, which can negatively impact the performance of such algorithms due to the curse of dimensionality.

One limitation of this study is the focus on a single dataset. Future work should explore the generalizability of these findings across different datasets and plant diseases. Additionally, further research could investigate hybrid models that combine the strengths of CNNs and traditional algorithms to improve efficiency without substantially compromising accuracy.

Overall, this study provides valuable insights that can guide the selection of machine learning models for plant pathology classification. It highlights the need for a nuanced approach to model selection that considers the specific requirements and constraints of each application scenario.

VI. Conclusion

This study conducted a comparative analysis of a CNN-based model and traditional machine learning classifiers for the task of multilabel plant pathology classification. The analysis was grounded on the Plant Pathology 2020 FGVC7 dataset, which provided a challenging set of high-resolution leaf images with various disease labels.

The CNN-based PlantModel demonstrated a clear advantage over traditional machine learning methods in terms of F1 score and accuracy. This underscores the potential of deep learning techniques in handling complex pattern recognition tasks in image-based datasets. While the traditional classifiers—Random Forest, KNN, and GaussianNB—were outperformed by the CNN, they offered significantly lower prediction times, which may be valuable in time-sensitive applications.

The findings of this study advocate for a strategic selection of machine learning models based on the specific needs of the application. For instance, in real-time monitoring systems where quick predictions are essential, a trade-off might be made in favor of more computationally efficient models. Conversely, for research purposes where accuracy is the priority, the higher computational cost of CNNs may be deemed acceptable.

However, the study is not without limitations. The focus on a single dataset and the intrinsic differences in model complexity and computational demands highlight the necessity for further research. Future studies could explore the generalizability of these models across various datasets and examine the integration of CNN features into traditional models to create hybrid approaches that could potentially offer a balance between accuracy and efficiency.

In conclusion, the research contributes to the body of knowledge in agricultural technology by providing a clearer understanding of the trade-offs involved in using CNNs versus traditional machine learning methods for plant disease classification. It paves the way for future explorations into optimizing these models for practical deployment in agricultural settings, with the ultimate goal of enhancing crop management and food security.

References

[1] Ü. Atila, M. Uçar, K. Akyol, and E. Uçar, ”Plant leaf disease classification using EfficientNet deep learning model,” *Ecological Informatics*, vol. 61, pp. 101182, 2021, Elsevier.

[2] L. Li, S. Zhang, and B. Wang, ”Plant disease detection and classification by deep learning—a review,” *IEEE Access*, vol. 9, pp. 56683–56698, 2021, IEEE.

[3] A. Abade, P. A. Ferreira, and F. de Barros Vidal, ”Plant diseases recognition on images using convolutional neural networks: A systematic review,” *Computers and Electronics in Agriculture*, vol. 185, pp. 106125, 2021, Elsevier.

[4] A. A. Nasser and M. A. Akhloufi, ”CTPlantNet: A Hybrid CNN-Transformer Architecture for Plant Disease Classification,” in *2022 International Conference on Microelectronics (ICM)*, pp. 156–159, 2022, IEEE.

[5] B. Liu, Y. Zhang, D. He, and Y. Li, ”Identification of apple leaf diseases based on deep convolutional neural networks,” *Symmetry*, vol. 10, no. 1, pp. 11, 2017, MDPI.

[6] W. E. MacHardy, D. M. Gadoury, and D. A. Rosenberger, ”Delaying the onset of fungicide programs for control of apple scab in orchards with low potential ascospore dose of Venturia inaequalis,” *Plant disease*, vol. 77, no. 4, pp. 372–375, 1993.

[7] A.-K. Mahlein, ”Plant disease detection by imaging sensors–parallels and specific demands for precision agriculture and plant phenotyping,” *Plant disease*, vol. 100, no. 2, pp. 241–251, 2016, Am Phytopath Society.

[8] E. Mwebaze et al., ”iCassava 2019 fine-grained visual categorization challenge,” *arXiv preprint arXiv:1908.02900*, 2019.

[9] S. Sladojevic et al., ”Deep neural networks based recognition of plant diseases by leaf image classification,” *Computational intelligence and neuroscience*, vol. 2016, 2016, Hindawi.

[10] Z. Yan et al., ”HD-CNN: hierarchical deep convolutional neural networks for large scale visual recognition,” in *Proceedings of the IEEE international conference on computer vision*, pp. 2740–2748, 2015.

[11] J. Chen, D. Zhang, and Y. A. Nanehkaran, ”Identifying plant diseases using deep transfer learning and enhanced lightweight network,” *Multimedia tools and applications*, vol. 79, pp. 31497–31515, 2020, Springer.

[12] Plant Pathology 2020 - FGVC7, https://www.kaggle.com/competitions/plant-pathology-2020-fgvc7/data