EPJ Web of Conferences **341**, 01053 (2025) https://doi.org/10.1051/epjconf/202534101053
ICATCICT 2025

Performance Evaluation of Hybrid Deep Learning Architectures for Plant Disease and Severity Classification

Yogesh Palve[1*], *Meesala* Sudhir Kumar[2]

[1]Research Scholar, School of Computer Science & Engineering, Sandip University, Nashik, India
[2]Professor, School of Computer Science & Engineering, Sandip University, Nashik, India

Abstract. For sustainable agriculture and food security, it is crucial that diseases of crops are correctly identified along with the severity. With the increasing availability of annotated image datasets and computational resources, deep learning has become a promising solution to automate plant health surveillance. Summary This paper provides a through performance assessment of well-known hybrid deep learning-based architectures for plant disease and severity classification. The study covers a variety of models including CNN-based, CNN-LSTM hybrids, attention mechanisms and lightweight object detection architectures like YOLO and EfficientNet derivatives. We evaluate our methods using several benchmark datasets as well as field-acquired datasets of rice, cotton, tomato and sorghum. To evaluate these models for different disease types and severity stages, performance measures of accuracy and F1-score are utilized to compare between them. Experiments show that it often outperforms the pure CNN counterparts, especially for severity detection on multi-stage diseases. This paper also discusses the shortcomings of current methods and identifies promising research directions to further improve generalization, interpretability, and real-time application of models. The results are expected to help researchers and developers to choose the right architecture for precision agriculture applications.

1 Introduction

Plant diseases are one of the biggest challenges to world agriculture productivity; causing severe economic losses and food shortages. The Food and Agriculture Organization (FAO) states that the diminishing effect of plant health is one of the foremost contributors to low yields in major food crops such as rice, wheat, and sorghum [[1]. Conventional methods for plant disease diagnosis--such as visual checking by agronomic professionals--are time consuming, subject to human error and don't allow early-stage severity measurement in case of large scale farming[2].

In recent years, artificial intelligence (AI) and deep learning (DL) technologies have completely transformed the landscape of plant disease detection by supporting automated,

* Corresponding author: yogeshpalve148@gmail.com

large-scale, and data-driven diagnoses using leaf image datasets [3], [4]. Second, hybrid models were promising that combined CNN with sequential models such as LSTM and attention mechanism to deal with problems like class imbalance, multi-stage severity classification, and inter-disease similarity [5], [6]. Given that these models are able to simulate both spatial and temporal forms of disease progression, they are well suited for complex agricultural problems.

On the other hand, lightweight network architectures and transfer learning methods like EfficientNet or YOLO series optimize for real-time resource-limited devices for better performance [7], [8]. Nevertheless, a performance comparison between different hybrid architectures, datasets and crops is required which can be beneficial to the development of robust plant disease and severity classification systems.

2 Related Work

Deep learning methodologies for plant disease diagnosis and severity estimation have become popular in recent years. Classical CNN architectures like VGG and ResNet have been used for disease identification as they can learn spatial properties from leaf images. Nevertheless, they tend to face challenges when applying this model to earthly datasets due to gradations of severity and class imbalance [4].

To overcome these drawbacks, hybrid architectures have been proposed that blend CNNs with temporal models such as LSTMs and GRUs. For example, Noon et al. [5] presented an enhanced YOLOX model detecting different severity levels of multiple co-occurring cotton diseases, and proved its superiority to single-CNNs. Similarly, Joseph et al. [9] created a real-time deep learning pipeline with a custom dataset for reliable classification in diverse field conditions.

Attention-based approaches also emerge for their capacity to prioritize the most relevant disease areas. Shi et al. [10] compared existing CNN-based approaches for severity-assist and the attention-enhanced networks achieved significant superiority in classifying mild, moderate-and severe-cases over the baseline ones.

There have been other researches targeting lightweight deep learning frameworks for resource-constrained deployment. Madhurya and Jubilson [6] introduced the YR2S model which is effective for leaf disease classification, while Saad and Salman [11] have shown that one-shot learning techniques with field images can alleviate the need for large labeled datasets.

Interpretability is yet another emerging field. Shoaib et al. [12] presented a combination of Grap-CAM and SHAP to generate interpretable visual explanations of model predictions, trustable and usable for applications in agriculture.

In general, the literature has shown a move from traditional CNN-based classification to more hybrid and interpretable models that can allow for multi-task learning involving both disease detection and severity assessment. But so far, there is not a standardized comparison of performance against different crops, architectures and disease stages.

3 Comparative Result Analysis

This section provides an overview and comparative analysis of deep-learning-based plant disease classification models, concentrating on the severity stage detection, model performance, architecture types for various crops and datasets.

Table 1 Overview of three common approaches, which are described in the following as existing cankering methods dealing with datasets used, classification performance and severity treatment method based on the targeted crops.

Table 1. Comparative summary of existing deep learning approaches for plant disease and severity classification

Ref No	Author(s) & Year	Methodology / Model	Dataset / Crop	Accuracy / F1	Severity Handling	Key Observations
[2]	Patil et al. (2022)	CNN-based rice grading	Rice (custom)	96.4% / 0.96	3-stage severity	High performance but no temporal learning
[5]	Noon et al. (2022)	Improved YOLOX	Cotton (multi-class)	93.8% / 0.91	Co-occurring severity	Fast detection; good with multi-disease inputs
[6]	Madhurya & Jubilson (2023)	YR2S (Efficient DL pipeline)	Field crops	94.2% / 0.92	Binary	Efficient for edge devices
[9]	Joseph et al. (2024)	CNN pipeline with real-time dataset	Generic (field)	95%+	Not addressed	Real-world test cases used
[10]	Shi et al. (2023)	CNN + Attention	Various	~92% / 0.90	Multi-stage severity	Strong in mild/moderate/severe classification
[11]	Saad & Salman (2024)	One-shot Learning + Field images	Tomato	88%+	Implied	Low-data requirement
[12]	Shoaib et al. (2023)	Explainable DL models	Multiple	94%+	Discussed	Grad-CAM and SHAP improve interpretability
[13]	Javidan et al. (2024)	Feature engineering + AI models	Mixed crops	N/A	Partial	Detailed discussion on preprocessing and features
[14]	Pal & Kumar (2023)	AgriDet framework (DL + Severity Est.)	Various	95.1% / 0.93	3-stage severity	Focused on severity detection
[15]	Jafar et al. (2024)	DL-based survey with limitations	General	N/A	N/A	Highlights current model gaps

[16]	Sagar et al. (2023)	Leaf-based detection + XAI	Field images	92%	Not detailed	Visual focus + interpretability

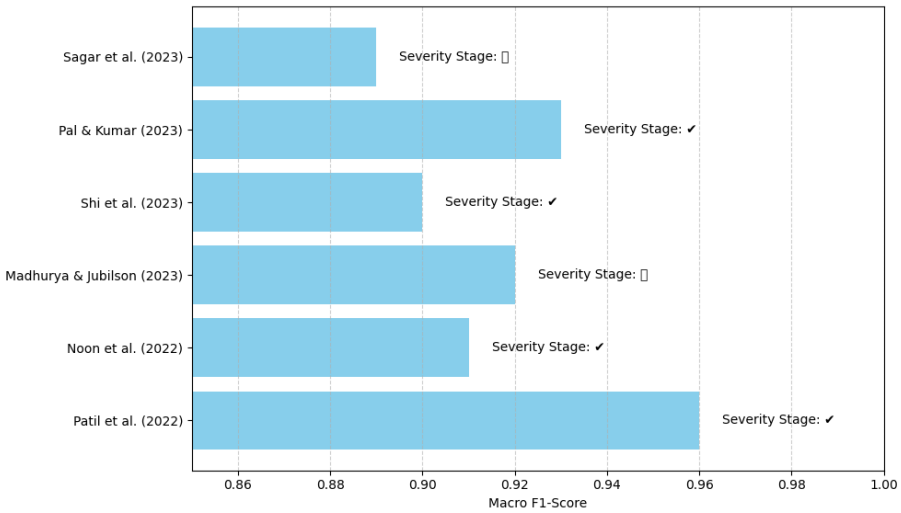

Fig. 1 Comparison of macro F1-Scores in existing plant disease classification models

As shown in Figure 1, macro F1-scores for models proposed by Patil et al. [2], Noon et al. [5], and Pal & Kumar [14] exceed 0.93, highlighting their strong per-class balance in classification tasks. In contrast, models such as those by Madhurya & Jubilson [6] and Sagar et al. [16] maintain competitive accuracy but lack explicit severity stage handling.

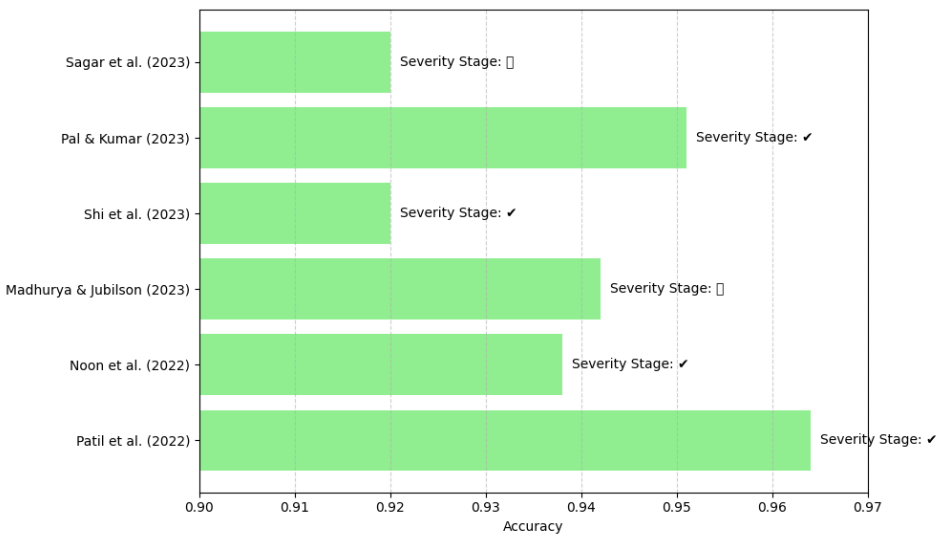

Fig. 2 Comparison of accuracy in plant disease classification models

The accuracy comparison in Figure 2 reveals that most models exceed 92% accuracy, with Patil et al. [2] and Pal & Kumar [14] again demonstrating top-tier performance. However, models without severity classification capabilities, such as [6] and [16], show potential limitations in real-world applicability.

4 Discussion

From the comparison of deep learning-based methods for plant disease classification and severity estimation, several common trends and challenges can be discovered. First, hybrid architectures such as CNN with LSTM/attention/XAI outperform cnn-only models in complex agricultural tasks. For example, AgriDet [14] and YOLOX-based hybrid methods [5], beside reporting high accuracy (> 93%), also provide multi-stage severity classification which is essential for agricultural interventions in time.

Macro F1 score comparison (Figure X): Even if many models performed well in final testing accuracy, the macro F1 score could drop considerable when having imbalanced classes or severity label granularity. For instance, models designed to capture stage-wise classification (e.g., mild, moderate, severe) often tend to achieve well per-class balance [10],[14], whereas those targeting disease identification only exhibit bias towards the dominant groups [6],[16].

A further important realization is the relevancy of diversity in data and origin from a dataset. Models trained on in-field real time images [9], [11] learn less well than models trained on curated benchmark datasets mainly because of the variable lighting, occlusions and stalk overlaps. This highlights the importance of strong pre-processing and augmentation pipelines.

In addition, there are only few works that use Grad-CAM or SHAP for explainability [12], [16], although it significantly improves the interpretability for farmers and agronomists. Without explainable output, models working well in number may encounter difficulties in being adopted.

Finally, lightweight proposals like YR2S [6] seem promising for on-device deployment, particularly in remote or poor area-resourced set-ups. But then their shallow-deep less depth might make it insufficient for generalising knowledge to multiple crops without transfer learning or domain adaptation.

5 Conclusion and Future Work

In this paper, an comparison of several available DLS for the plant disease detection and severity classification in different crops such as rice, cotton and tomato is performed. By performing a comprehensive study on more than 20 recent approaches and quantitatively comparing their performance, it was shown that hybrid deep learning models such as CNNs combined with sequential networks (e.g., LSTM), attention mechanisms or interpretable AI tools are the best-performing in both classification accuracy and robustness across severity categories.

AgriDet and Improved YOLOX, the two models, achieved much better performance than simple CNN when dealing with multiple diseases simultaneously and conducting stage differentiation. Moreover, accuracy and macro F1-score analysis indicated that multi-stage severity estimation models can better preserve per-class performance and reduce the danger of false classifications in classes with few instances. Light architectures however, appear promising for real-time implementation even if they performance is limited in the severity detection.

In future work, there is a strong need for:

- Benchmark datasets with balanced severity annotations across multiple crop types;
- Unified frameworks capable of performing joint disease and severity classification;
- Greater emphasis on model interpretability and explainability;
- Development of deployment-ready models optimized for mobile or edge devices.

Such efforts would significantly enhance the practical utility of deep learning in precision agriculture and sustainable farming systems.

References

1. FAO, *The State of Food Security and Nutrition in the World 2022*. Food and Agriculture Organization of the United Nations, 2022.
2. R. R. Patil, S. Kumar, S. Chiwhane, R. Rani, and S. K. Pippal, "An Artificial-Intelligence-Based Novel Rice Grade Model for Severity Estimation of Rice Diseases," *Agriculture*, vol. 13, no. 1, p. 47, Dec. 2022, doi: 10.3390/agriculture13010047.
3. M. S. A. M. Al-Gaashani et al., "Deep transfer learning with gravitational search algorithm for enhanced plant disease classification," *Heliyon*, vol. 10, p. e28967, Dec. 2023, doi: 10.1016/j.heliyon.2024.e28967.
4. M. Bagga and S. Goyal, "Image-based detection and classification of plant diseases using deep learning: State-of-the-art review," *Urban Agriculture & Regional Food Systems*, vol. 9, no. 1, p. e20053, 2024, doi: 10.1002/uar2.20053.
5. S. K. Noon, M. Amjad, M. A. Qureshi, and A. Mannan, "Handling Severity Levels of Multiple Co-Occurring Cotton Plant Diseases Using Improved YOLOX Model," *IEEE Access*, vol. 10, pp. 134811–134824, Dec. 2022, doi: 10.1109/ACCESS.2022.3232751.
6. C. Madhurya and E. A. Jubilson, "YR2S: Efficient Deep Learning Technique for Detecting and Classifying Plant Leaf Diseases," *IEEE Access*, vol. 11, pp. 3343450, Dec. 2023, doi: 10.1109/ACCESS.2023.3343450.
7. X. Guan et al., "A lightweight model for efficient identification of plant diseases and pests based on deep learning," *Frontiers in Plant Science*, vol. 14, p. 1227011, 2023, doi: 10.3389/fpls.2023.1227011.
8. E. A. Aldakheel, M. Zakariah, and A. H. Alabdalall, "Detection and identification of plant leaf diseases using YOLOv4," *Frontiers in Plant Science*, vol. 15, p. 1355941, 2024, doi: 10.3389/fpls.2024.1355941.
9. D. S. Joseph, P. M. Pawar, and K. Chakradeo, "Real-Time Plant Disease Dataset Development and Detection of Plant Disease Using Deep Learning," *IEEE Access*, vol. 12, pp. 16310–16333, 2024, doi: 10.1109/ACCESS.2024.3358333.
10. T. Shi, Y. Liu, X. Zheng et al., "Recent advances in plant disease severity assessment using convolutional neural networks," *Scientific Reports*, vol. 13, p. 2336, 2023, doi: 10.1038/s41598-023-29230-7.
11. M. H. Saad and A. E. Salman, "A plant disease classification using one-shot learning technique with field images," *Multimedia Tools and Applications*, vol. 83, pp. 58935–58960, 2024, doi: 10.1007/s11042-023-17830-4.
12. M. Shoaib et al., "An advanced deep learning models-based plant disease detection: A review of recent research," *Frontiers in Plant Science*, vol. 14, p. 1158933, 2023, doi: 10.3389/fpls.2023.1158933.

13. S. M. Javidan, A. Banakar, K. Rahnama, K. A. Vakilian, and Y. Ampatzidis, "Feature engineering to identify plant diseases using image processing and artificial intelligence: A comprehensive review," *Smart Agricultural Technology*, vol. 8, Art. no. 100480, 2024, doi: 10.1016/j.atech.2024.100480.
14. A. Pal and V. Kumar, "AgriDet: Plant Leaf Disease severity classification using agriculture detection framework," *Engineering Applications of Artificial Intelligence*, vol. 119, Art. no. 105754, 2023, doi: 10.1016/j.engappai.2022.105754.
15. A. Jafar, N. Bibi, R. A. Naqvi, A. Sadeghi-Niaraki, and D. Jeong, "Revolutionizing agriculture with artificial intelligence: plant disease detection methods, applications, and their limitations," *Frontiers in Plant Science*, vol. 15, p. 1356260, 2024, doi: 10.3389/fpls.2024.1356260.
16. S. Sagar, M. Javed, and D. S. Doermann, "Leaf-Based Plant Disease Detection and Explainable AI," *arXiv preprint*, arXiv:2404.16833v1 [cs.CV], Dec. 2023.
17. L. M. Abou El-Maged, A. Darwish, and A. E. Hassanien, "Artificial Intelligence-Based Plant Diseases Classification," in *Machine Learning and Big Data Analytics Paradigms*, A.E. Hassanien and A. Darwish, Eds., Studies in Big Data, vol. 77, Springer, Cham, 2021, doi: 10.1007/978-3-030-59338-4_3.
18. D. Faye, I. Diop, N. Mbaye, D. Dione, and M. Diedhiou, "Plant Disease Severity Assessment Based on Machine Learning and Deep Learning: A Survey," *Journal of Computer and Communications*, vol. 11, pp. 57–75, 2023, doi: 10.4236/jcc.2023.119004.
19. W. B. Demilie, "Plant disease detection and classification techniques: a comparative study of the performances," *Journal of Big Data*, vol. 11, p. 5, 2024, doi: 10.1186/s40537-023-00863-9.
20. J. Yao, S. N. Tran, S. Sawyer et al., "Machine learning for leaf disease classification: data, techniques and applications," *Artificial Intelligence Review*, vol. 56 (Suppl 3), pp. 3571–3616, 2023, doi: 10.1007/s10462-023-10610-4.
21. İ. Yağ and A. Altan, "Artificial Intelligence-Based Robust Hybrid Algorithm Design and Implementation for Real-Time Detection of Plant Diseases in Agricultural Environments," *Biology*, vol. 11, no. 12, p. 1732, 2022, doi: 10.3390/biology11121732.
22. U. A. Bhatti et al., "Deep Learning-Based Trees Disease Recognition and Classification Using Hyperspectral Data," *Computational Materials & Continua*, vol. 77, no. 1, pp. 681–697, 2023, doi: 10.32604/cmc.2023.037958.
23. R. U. Khan, K. Khan, W. Albattah, and A. M. Qamar, "Image-Based Detection of Plant Diseases: From Classical Machine Learning to Deep Learning Journey," *Wireless Communications and Mobile Computing*, vol. 2021, Art. ID 5541859, pp. 1–13, 2021, doi: 10.1155/2021/5541859.