INTERNATIONAL JOURNAL OF RESEARCH AND INNOVATION IN APPLIED SCIENCE (IJRIAS)
ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue XI November 2025

Deep Learning-Based Wheat Disease Detection and Classification System Using Convolutional Neural Networks

Ms. Drashti Shah[1], Mr. Dhruv Chauhan[2], Dr (Mrs). Mahasweta Joshi[3]

[1,2]B. Tech Computer Engineering, Birla Vishvakarma Mahavidyalaya (BVM) Engineering College, V. V. Nagar, India,

[3]Assistant Professor, Computer Department, Birla Vishvakarma Mahavidyalaya (BVM) Engineering College, V. V. Nagar, India

DOI: https://dx.doi.org/10.51584/IJRIAS.2025.101100090

Received: 04 December 2025; Accepted: 10 December 2025; Published: 19 December 2025

ABSTRACT

Wheat, one of the major crops in the world, is vulnerable to many diseases that cause tremendous yield and quality loss. This paper proposes a deep learning method for the automatic detection and classification of wheat diseases based on a Convolutional Neural Network (CNN). We respond to the imperative of early and precise identification of diseases in wheat crops in order to reduce agricultural losses.The system learned on a data set of more than 14,000 wheat leaf images corresponding to 15 classes of various rusts, blights, insects, and normal leaves. Our suggested CNN model reached a training accuracy of 97.02% and validation accuracy of 91.00%. The model design uses data augmentation strategies and dropout regularization to promote generalization as well as avoid overfitting

In addition, we created a friendly, web-based platform based on Streamlit that combines the trained model with a MySQL database for real-time disease detection and recommendation of corresponding treatments. The outcomes show the efficacy and practicality of deep learning in contemporary agricultural disease management, offering farmers and agricultural specialists a worthwhile resource.

Keywords: Wheat Disease Detection, Deep Learning, Convolutional Neural Network (CNN), Agricultural Technology, Image Classification, Plant Pathology.

INTRODUCTION

Wheat is one of the most cultivated cereal crops and is a food staple for more than one billion people worldwide. Yet wheat is constantly being threatened by a vast range of pathogens (fungi, bacteria, viruses) and insects, which can lower both yield (volume) and grain quality. Of those, fungal diseases (rusts, mildews, blotches, root rots) are particularly destructive. Worldwide economic losses due to wheat rusts alone amount to billions of dollars per year.[11].

There are also viral diseases. Wheat streak mosaic virus (WSMV) and wheat dwarf virus (WDV), for instance, can induce severe yield losses and stunting. Leaf blight in India, caused by the fungus Alternaria triticina, is critical: under heavy infection, yield loss of up to 60 % has been observed. [12].

The recent advancements in artificial intelligence, particularly in computer vision and deep learning, offer powerful solutions to these challenges. Convolutional Neural Networks (CNNs), a class of deep neural networks, are exceptionally well-suited for image analysis tasks and have demonstrated remarkable success in image classification. This research leverages CNNs to develop a comprehensive system for wheat disease detection that can accurately classify 15 different types of diseases and healthy plant conditions from leaf images. The system extends beyond simple classification by integrating the predictive model into a practical web application that provides users with suggested treatment measures from a connected database, creating a complete end-to-end diagnostic tool.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue XI November 2025

LITERATURE REVIEW

The use of computer vision in agriculture has increased considerably, with most researchers targeting self-diagnosis of crop diseases.

Mohanty et al. (2016) — "Using Deep Learning for Image-Based Plant Disease Detection" They train a deep convolutional neural network to recognize 14 crop species and 26 diseases (or lack thereof). The model after training gets an accuracy of 99.35 %. The public dataset employed contained 54,306 images of diseased and healthy plant leaves (controlled environment). They experimented with CNN architectures like AlexNet and GoogLeNet[6].

Sladojevic et al. (2016) — "Deep neural networks based recognition of plant diseases by leaf image classification". Sladojevic et al. (2016) used a comparable approach to detecting plant disease via leaves images … with fewer diseases (13) and plants (5).— as described in a summary review mentioning works such as Sladojevic et al.Success rates of their models were "between 91 % and 98 %" based on test data sets. [5].

Ferentinos (2018) — "Deep learning models for plant disease detection and diagnosis" Ferentinos compared/evaluated several deep learning architectures (VGG, AlexNet, GoogLeNet, etc.) through transfer learning in order to detect plant diseases. The paper describes high success/accuracy rates for those experiments. With regard to resource limitations and domain specificity, one proviso is that transfer learning models tend to need sensitive fine-tuning to adjust to domain-specific information, which can be nontrivial in reality (this is a general observation rather than a direct citation, but in line with criticism in numerous reviews) [7].

More specifically in the case of wheat, Khan et al. (2021) [10] employed the ResNet-50 architecture to classify wheat diseases with good performance on large datasets. Nevertheless, their effort failed to offer an end-to-end pipeline with a UI or a remedy suggestion system. In contrast, our method employs a low-weight, custom-developed CNN specially designed for the classification of wheat leaf disease. This tailorable architecture is computationally scalable and is implemented within an end-toend, usable solution having a Streamlit frontend and a MySQL-backed cure retrieval system, bringing it within reach and actionable by end-users.

METHODOLOGY

Our methodology revolves around three main elements: the dataset and data preprocessing, design and training of the CNN model, and the web application of the end-user.

Dataset Description

The study utilizes a publicly available dataset from Kaggle titled 'Wheat Plant Diseases'. The dataset contains over 14,000 labeled images of wheat leaves. It was partitioned as follows:

- Training Set: 13,105 images
- Validation Set: 300 images
- Test Set: 700 images

The data is split into 15 classes, for 14 prevalent diseases/pests and a 'Healthy' class and others are: Aphid, Black Rust, Blast, Brown Rust, Common Root Rot, Fusarium Head Blight, Leaf Blight, Mildew, Mite, Septoria, Smut, Stem Fly, Tan Spot, Yellow Rust..

Data Preprocessing and Augmentation

In order to preprocess the data for the model and make it more capable of generalizing, we employed the ImageDataGenerator class of TensorFlow/Keras.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue XI November 2025

- Normalization: Every image saw its pixel values being converted from the original range [0,255] to a normalized range [0,1] by dividing by 255.0. This is a common technique to ensure stable and efficient training.

- Data Augmentation: For defense against overfitting and for exposing the model to a greater number of image variations, the following random transformations were applied solely to the training set:

1. Shear transformation (range: 0.2)
2. Zoom transformation (range: 0.2)
3. Horizontal flipping

Validation and test datasets were not augmented but only normalized to enable the evaluation of the model's performance on unchanged data.

CNN Model Architecture

A sequential CNN was designed from scratch. The architecture consists of a convolutional base for feature extraction followed by a dense classifier head.

Input Layer: 128x128x3 RGB images

├ Conv2D(32,3x3) + ReLU + MaxPooling(2x2)

├ Conv2D(64,3x3) + ReLU + MaxPooling(2x2)

├Conv2D(128,3x3) + ReLU + MaxPooling(2x2)

├Conv2D(256,3x3) + ReLU + MaxPooling(2x2)

├ Dropout (0.25)

├ Flatten ()

├ Dense (1200) + ReLU

├ Dropout (0.40)

└ Dense (15) + Softmax

Training Configuration

The model was compiled and trained using th following configuration:

- **Optimizer**: Adam
- **Loss Function**: categorical_crossentropy, suitable for multi-class classification
- **Metrics**: Accuracy
- **Batch Size**: 32
- **Epochs**: 100
- **Hardware**: The training was accelerated using a GPU within the TensorFlow framework

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue XI November 2025

Web Application and System Implementation

A user-friendly web application was created with the Streamlit library in Python to offer a useful interface for the model.The application enables a user to upload an image of a wheat leaf. The preprocessed uploaded image is then input into the trained CNN model for real-time prediction. After classification, the name of the predicted disease is employed to search a locally stored MySQL database with a table that lists diseases against their corresponding cures or control measures. The program will then present the predicted disease and its proposed cure to the user

RESULTS AND ANALYSIS

4.1 Model Performance and Training Progress

The model was tested after training for 100 epochs. The ultimate performance metrics were as follows:

● **Training Accuracy**: 97.02%

● **Validation Accuracy**: 91.00%

As can be seen from Figure 3, the accuracy of the model on both the training and validation data increased steadily throughout the epochs. There was fast learning in the first 20-30 epochs and subsequent steady improvement until the accuracy levelled off, meaning that the model had converged well. Dropout layers were successfully used to maintain the difference between the training and validation curves minimal, which meant that overfitting was well-controlled.

Class-wise Performance Analysis

A detailed classification report was generated to evaluate the model's performance on each individual class.

The table 1 report displays a mixed performance by classes:

● **High-Performing Classes:** The model performed exceedingly well for a number of classes, with F1-scores of 0.95 or more for 'Aphid', 'Black Rust', 'Blast', 'Brown Rust', 'Common Root Rot', 'Fusarium Head Blight', 'Mildew', 'Mite', 'Septoria', 'Smut', and 'Stem fly'. This means that the visual features of these diseases are unique and well-learned by the model.

● **Challenging Classes**: The model had a very poor performance with the 'Healthy' class, which had a very low recall of 0.10 and an F1-score of a mere 0.18. Although its precision was 1.00 (i.e., when it predicted 'Healthy', it was always correct), it missed 90% of the actual healthy images and predicted them as diseased. 'Yellow Rust' also had a low precision (0.54), which means it was often predicted when the actual class was different.

Confusion Matrix Analysis

The confusion matrix provides a visual breakdown of correct and incorrect predictions for all classes.

The matrix in Figure 4 graphically validates the class-wise performance analysis:

● There is a clear diagonal pattern for all classes, which indicates a large number of correct predictions.

● The 'Healthy' class row indicates that a mere 2 out of 20 samples were correctly labeled, while the rest of 18 samples were mislabeled as other disease classes. This is the strongest flaw of the model.

● There is minimal overlap between visually alike diseases, like Fusarium Head Blight being classified as Healthy or Tan spot.

DISCUSSION

The overall validation accuracy of 91.00% indicates that the custom CNN model is very powerful at multi-class classification of wheat diseases from leaf images.The multi-layer convolutional architecture was able to learn and effectively extract discriminative visual features for the majority of the 14 diseases.The deployment of this model in a Streamlit web application using a MySQL backend for treatment recommendations offers a convenient and easy-to-use end-to-end solution for end-users such as farmers and agronomists.

But the model's main limitation is that it is weak on the 'Healthy' class. The very low recall indicates that the model leans toward disease prediction. This may be because there is a class imbalance in the data or because the visual difference between a healthy leaf and one with an incipient disease is fine. This is a problem for practical usage because a high level of false positives might result in the useless and expensive use of chemical measures. The ambiguity of 'Yellow Rust' also suggests its visual signs might be very similar to other types.

CONCLUSION AND FUTURE WORK

Conclusion

This study has effectively designed and tested a deep learning network for the computer-aided detection and classification of wheat diseases. The major contributions of this work are:

1. The design of a strong custom CNN model with 91% validation accuracy for 15 different wheat disease and health classes.

2. Deployment of a full, end-to-end system from image upload through diagnosis and treatment recommendation using an intuitive web application..

3. A full-scale performance check that affirms the strength of the model in distinguishing the majority of diseases while clearly defining its weakness, especially in differentiating healthy leaves.

Future Work

In order to overcome the existing constraints and improve the system further, the following directions are suggested for future work::

1. Dataset Augmentation: The immediate priority is to enhance the dataset. This means acquiring more healthy wheat leaf images in order to balance the classes and taking images under varying field conditions (e.g., lighting, background, and growth stages) to enhance model robustness.

2. Model Refinements: Trying out complex deep learning models using transfer learning (i.e., applying pre-trained models such as ResNet or EfficientNet) might yield better accuracy. Using ensemble techniques or attention could also cause the model to pay more attention to more important features and diminish ambiguity among similar classes.

3. System Improvements: For greater accessibility and utility, the system may be created as a realtime, in-field diagnostic mobile application. Integration with IoT sensors to gather environmental information (humidity, temperature) may also add contextual information to enhance the accuracy of diagnosis.

This work is an addition to the development of agricultural technology through giving a pragmatic, Albased approach that has the potential to become impactful in agricultural productivity by making early disease intervention possible and assisting in the provision of global food resources.

REFERENCES

1. kushagra3204, "Wheat Plant Diseases Dataset," Kaggle, 2024. [Online]. Available:https://www.kaggle.com/datasets/kushagra3204/wheat-plant-diseases

2. Chollet, F. (2017). Deep Learning with Python. Manning Publications.
3. TensorFlow Developers. (2025). TensorFlow (Version X.X). Retrieved from https://www.tensorflow.org/
4. Streamlit Inc. (2025). Streamlit: The fastest way to build and share data apps. Retrieved from https://docs.streamlit.io/
5. Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., & Stefanovic, D. (2016).
6. Deep neural networks-based recognition of plant diseases by leaf image classification. Computational intelligence and neuroscience, 2016.
7. Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using deep learning for imagebased plant disease detection. Frontiers in plant science, 7, 1419
8. Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis.
9. Computers and Electronics in Agriculture, 145, 311–318.
10. Barbedo, J. G. A. (2019). Plant disease identification from individual lesions and spots using deep learning. Biosystems Engineering, 180, 96–107.
11. Too, E. C., Yujian, L., Njuki, S., & Yingchun, L. (2019). A comparative study of finetuning deep learning models for plant disease identification. Computers and Electronics in Agriculture, 161, 272–279.
12. Khan, M. A., Akram, T., Sharif, M., & Javed, M. Y. (2021). Wheat disease classification using hybrid deep learning techniques. IEEE Access, 9, 39563–39577
13. National library of medicine https://pmc.ncbi.nlm.nih.gov/articles/PMC6638159/
14. Institute of agriculture and natural resources https://cropwatch.unl.edu/early-disease-detection-highlights-importance-scouting-nebras kawheat-field

Figure and Table Legends:

Table 1: Classification report showing precision, recall, and F1-score for each of the 15 classes.

Figure 1: Model Architecture.

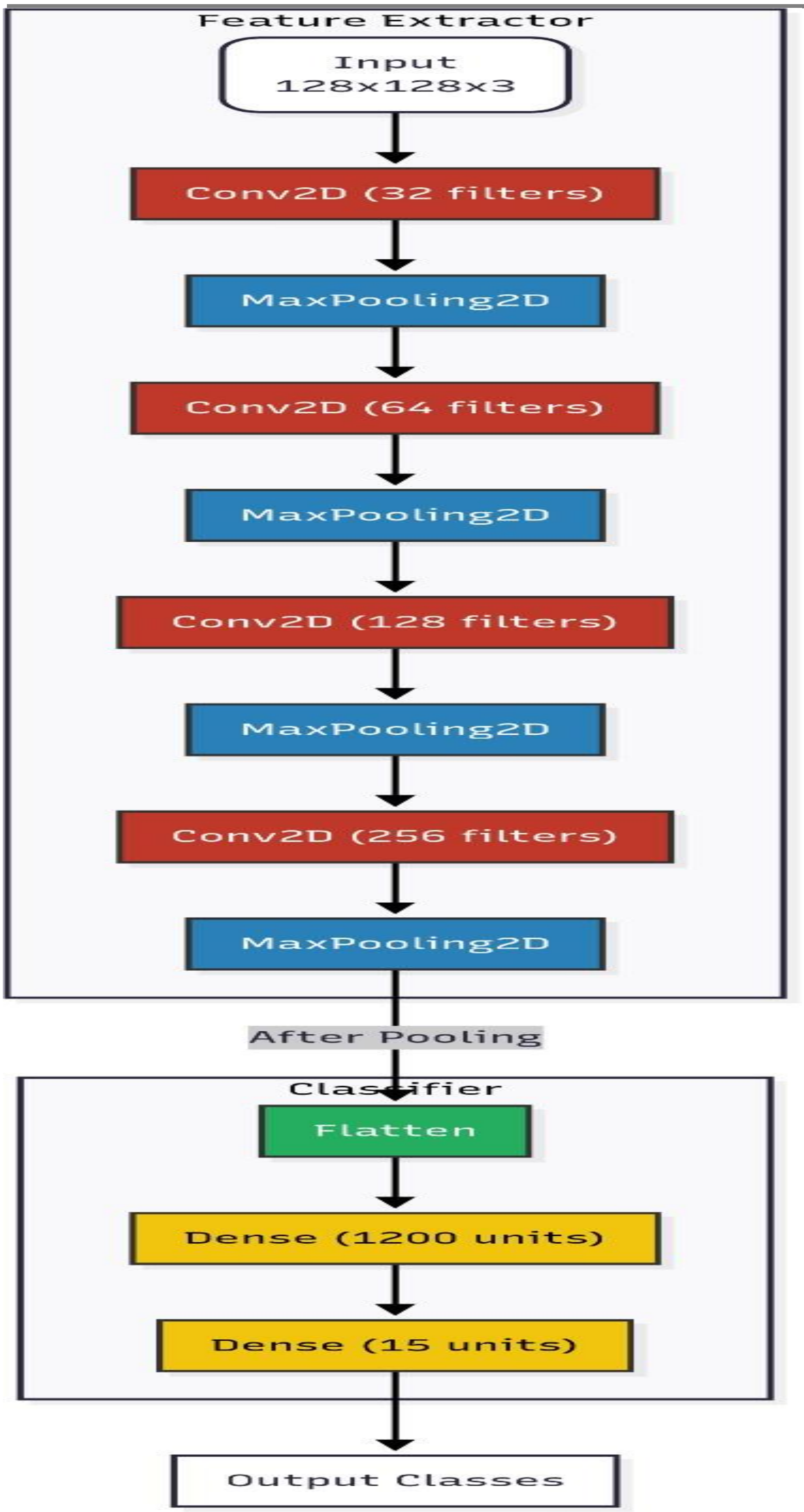
Figure 2: Image used on the home page of the Streamlit application, welcoming users to the system.

Figure 3: Training and validation accuracy over 100 epochs. The plot shows steady learning and convergence.

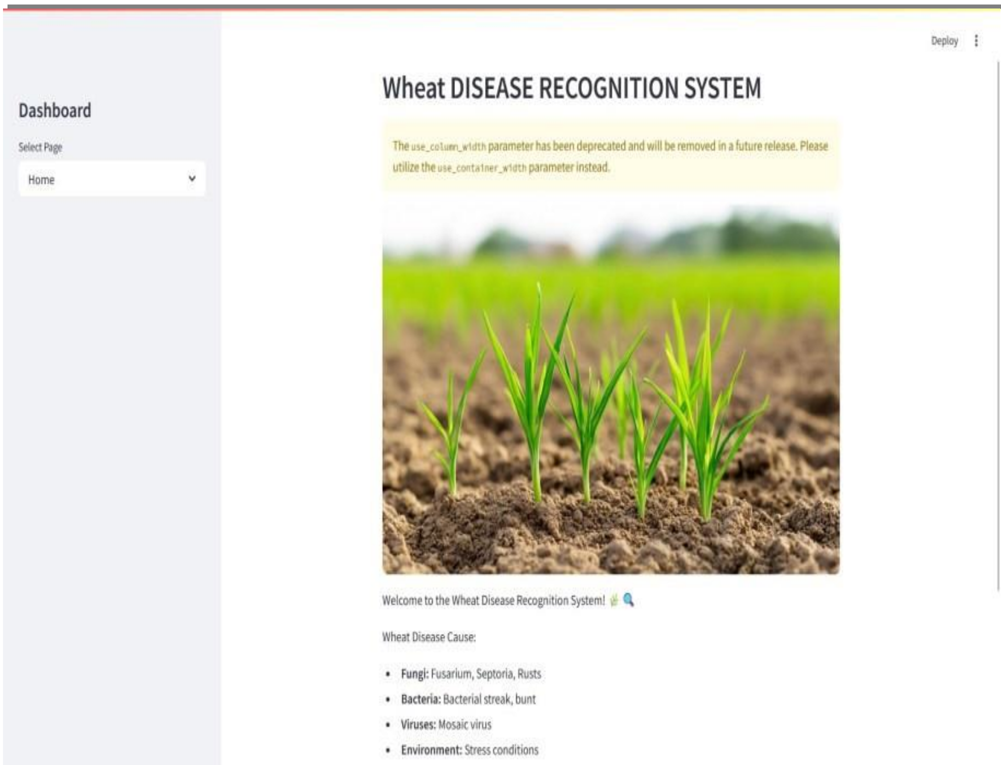
Figure 4: Confusion matrix for the 15 classes. The diagonal represents correctly classified instances.

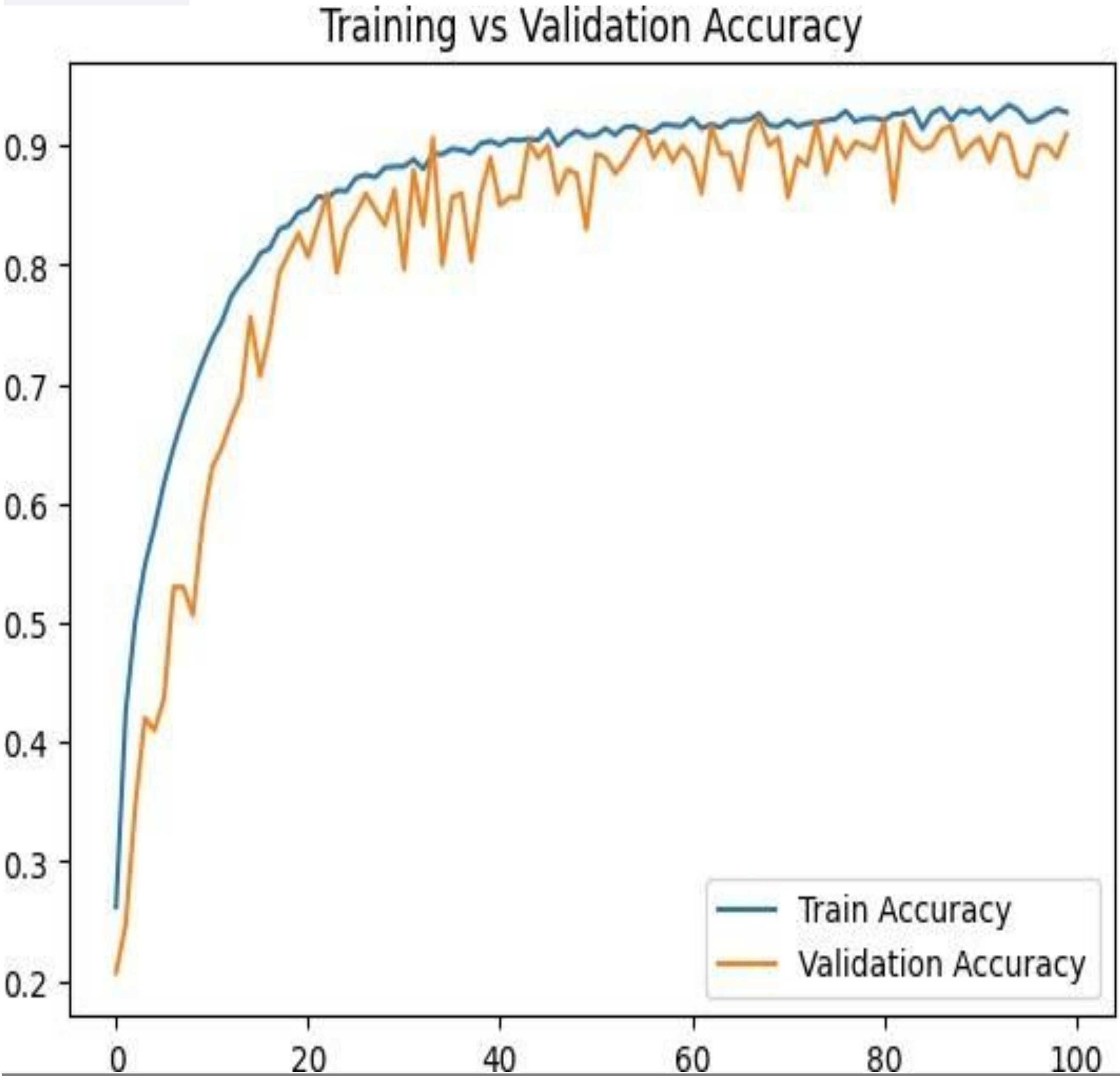
	precision	recall	f1-score	support
Aphid	0.91	1.00	0.95	20
Black Rust	1.00	0.95	0.97	20
Blast	1.00	1.00	1.00	20
Brown Rust	1.00	0.95	0.97	20
Common Root Rot	0.91	1.00	0.95	20
Fusarium Head Blight	1.00	0.90	0.95	20
Healthy	1.00	0.10	0.18	20
Leaf Blight	0.90	0.95	0.93	20
Mildew	1.00	1.00	1.00	20
Mite	1.00	0.95	0.97	20
Septoria	1.00	1.00	1.00	20
Smut	0.95	1.00	0.98	20
Stem fly	1.00	0.95	0.97	20
Tan spot	0.86	0.90	0.88	20
Yellow Rust	0.54	1.00	0.70	20
accuracy			0.91	300
macro avg	0.94	0.91	0.89	300
weighted avg	0.94	0.91	0.89	300

INTERNATIONAL JOURNAL OF RESEARCH AND INNOVATION IN APPLIED SCIENCE (IJRIAS)
ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue XI November 2025



INTERNATIONAL JOURNAL OF RESEARCH AND INNOVATION IN APPLIED SCIENCE (IJRIAS)
ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue XI November 2025





ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue XI November 2025