IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

A Comprehensive Research And Implementation Of Rice Leaf Disease Detection Using Machine Learning And Deep Learning

[1]Swati Tiwari, [2]Anuja Tipare, [3]Himani Salunkhe,[4]Priti Rathod
[4]Assistant Professor, [1,2,3]UG Student,
[1,2,3,4]Department of Computer Engineering,
[1,2,3,4]Ajeenkya D Y Patil School Of Engineering, Pune, India

Abstract - Rice is a cereal grain that is dominantly consumed by more than half of the world's population and is very significant in the Indian agricultural economy. Rice is the number one most eaten food globally, followed closely by other staples like wheat (in bread) and maize(corn). Agriculture and allied sectors contribute about 16 to 18% of India’s GDP; despite its lower GDP share, the sector employs more than 40% of the total population. And hence, India is among the leading producers and exporters of rice. However, the crop yield is significantly threatened by diseases such as bacterial Blight, brown spots, leaf blasts, etc., which can cause potential yield and quality loss if not identified properly.
Farmers majorly rely on manual field assessment, which is labour-intensive, time-consuming and often could lead to misidentification and significant economic loss. Recent advancements in fields like machine learning, deep learning, and hyperspectral imaging have automated data-driven approaches that are capable of improving detection, accuracy and efficiency. This study critically examine is more than 10 research works published between 2019 and 2025 covering traditional tools, techniques, frameworks, and algorithms for object detection and imaging techniques. And after a thorough systematic research and evaluation, a detection system was developed to identify rise, disease and help reduce yield loss.
After comparatively evaluating the research, it has been indicated that deep learning models consistently demonstrate superior performance, frequently achieving an accuracy level of more than 95%, while hyperspectral and UAV-based systems offer potential for early-stage disease identification, despite practical cost constraints.
This review identifies existing research gaps and focuses on the need for a scalable, lightweight and farmer-friendly disease detection system to support sustainable farming in India and contribute to global agricultural applications.

Keywords- Machine Learning, Deep Learning, Rice Disease Detection, YOLO, Hyperspectral Imaging, UAV, Smart Agriculture, Automation.

I. INTRODUCTION

Agricultural yield is totally dependent on effectiveness of crop health surveillance and timely disease detection. In rice cultivation, even the small infection can spread rapidly due to dense planting system and favorable climatic condition which correspond to yield loss. In India, rice serves as staple crop of economic and nutritional importance, so to prevent crop health is very important for sustaining food security and farmer’s income stability. However, rice crop are highly prone to various leaf base diseases that can potentially reduce growth and yield potential.

Rice leaf diseases are especially significant as the leaf directly operates photosynthetic efficiency and energy assimilation. Infection or any kind of disease to leave disrupt chlorophyll activity, reduces metabolic processes and directly shows effect on grain formulation and quality. Symptoms such as lesions, discoloration, or irregular spotting may appear similar to different disease categories in its early stage due to which disease identification becomes challenging for farmers. Farmers often rely on manual detection to differentiate between different disease types. Yet visual assessment can be subjective by factor such as lighting, experience, leaf maturity, and environmental stress effects, which in turn could increase the risk of misdiagnosis and inappropriate treatment.

Advancement in digital imaging, deep learning and machine learning technologies has enabled automatic detection using higher resolution image acquisition, availability of large-scale image to facilitate automated disease recognition, and analyzing different features such as texture, color, lesions morphology, and structured deformation and accordingly data driven models can detect visual patterns associated with specific rice leaf disease more accurately. Such technology support development of intelligent diagnostic system that can reduce dependency on manual expertise.

Despite significant research available in this domain, there are some research gaps to be filled with practical implementation and predictive performance. In addition to productive capability aspects, such as scalability, adaptability to field condition and economically feasible, plays vital role in successful deployment. Variation in environmental settings, software and hardware availability, proper and accurate dataset and resource constraints significantly affect system performance. Therefore, for integrating technical effectiveness and practical implementation we require a smooth transition between research setting, tools and technologies, and field level agricultural use.
So accordingly, this study presents a structured analytical investigation of rice disease detection approaches. By examining methodological, advancement, implementation, challenges, and system level consideration. The paper aims to identify existing research, limitation and outline direction for developing reliable scalable and farmer- friendly diagnostic system to diverse agriculture environment.

II. Overview of Rice Leaf Diseases

Disease is an abnormal condition that can affects structure or function of some or all parts, and usually associated with signs and symptoms. So rice leaf disease is majorly three types based on pathogens bacterial, fungal and viral diseases and few insect pests' disease.

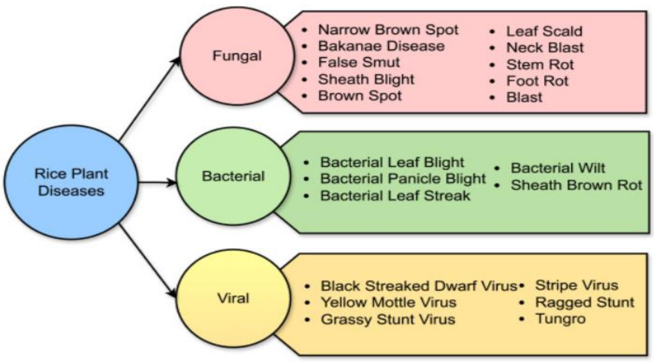

Fig 1: Rice plant diseases

So in this research we have studied 7 major diseases namely Brown Spot, Leaf scald, Sheath Blight, Bacterial Blight, Tungro, Hispa and Healthy leaf. So let's get a brief overview about these diseases:

A. Brown Spot:

Fig 2: Brown spot

- Type: Fungal.
- Caused by: Fungal agent Bipolaris oryzae also known as Helminthosporium oryzae is primary cause of brown spot disease.
- Development conditions: requires high humidity, nutrient-deficient soil mostly nitrogen and potassium level deficiency in the soil, and extended leaf wetness.
- Appearance: disease is addressed by small, circular to oval lesions on the leaf surface, lesions have a uniform brown center with a dark margin. Spots/lesions are usually scattered individually across the leaf surface.

B. Leaf scald

Fig 3: Leaf scald

- Type: Fungal.
- Caused by: Fungal agent Microdochium oryzae classified under Monographella albescens.
- Development condition: requires high humidity, excessive nitrogen fertilization, and dense crop canopy for favorable development.
- Appearance: initially appears as small, water-soaked lesions which are generally large and irregular near leaf tips/edges. Infects at tip and later spreads downward. Spot/lesion is grayish genter with sharp defined brown borders. Leaf appears partially dried/scalded.

C. Leaf Blast

Fig 4: Leaf blast

- Type: Fungal.
- Caused by: fungal agent Magnaporthe oryzae.
- Development condition: most destructive type spreads rapidly under high humidity, repeated rainfall, and moderate temperatures.
- Appearance: lesions are usually elongated or diamond shaped, these lesions have gray center with darker margin. And these lesions appear at various positions rather than at tip/margin.

D. Sheath Blight

Fig 5: Sheath blight

- Type: Fungal
- Caused by: fungal agent Rhizoctonia solani.
- Development condition: grows in typically dense crop canopies with high humidity.
- Appearance: initially lesions are oval/irregular large greenish-gray patches at lower leaf sheath. Later lesions are enlarged and extends to upper leave and no typical shapes like diamond/round.

E. Tungro

Fig 6: Tungro

- Type: Viral
- Caused by: combined effect of Rice Tungro Bacilliform Virus(RTBV) and Tungro Spherical Virus(RTSV).and transmitted by leafhoppers.
- Appearance: shows uniform yellow to orange discoloration of leaves, stunted growth and reduced tillering. Leaves appear narrow and upright. No lesions.

F. Hispa

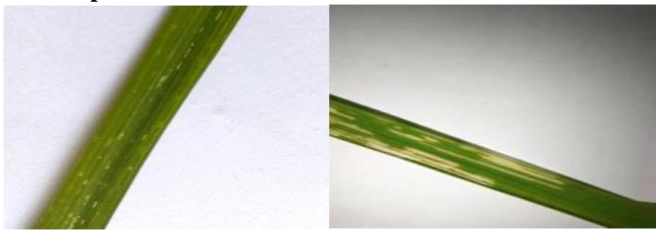

Fig 7: Hispa

- Type: infection due to insects.
- Caused by: pest Dicladispa armigera (Blue beetle).
- Development condition:
- Appearance

III. LITERATURE SURVEY

Sr. No	Title & Authors	Year	Objectives	Work Done	Conclusions	Future Scope	Remark s
1.	Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification- Sladojevic et al.	2016	Use CNNs for plant disease classification.	Built a CNN with Caffe framework on 30,880 images.	Achieved 93% accuracy, proving CNNs outperform handcrafted features.	Apply transfer learning for broader crop coverage.	Pioneering CNN work; dataset-driven accuracy but high computational demand.
2.	Hyperspectral Imaging Combined with Deep Transfer Learning for Rice Disease Detection- Lei Feng, Baohua Wu, Yong He and Chu Zhang	2021	Use hyperspectral imaging for early disease detection.	Applied deep transfer learning to hyperspectral datasets.	Achieved >92% accuracy in pre-symptomatic detection.	Reduce cost of HSI sensors and apply domain adaptation.	Excellent for early detection but impractical for small farmers.
3.	Rice Disease Detection Using YOLOv5- MdErshadul Haque, Ashikur Rahman, Iftekhar Junaeid, Samiul Ul Hoque, Manoranjan Pau	2022	Apply YOLOv5 for real-time rice disease detection.	Trained YOLOv5 on 1,500 annotated rice disease images.	Achieved precision 90%, recall 67%, mAP 76%, enabling fast field-level diagnosis.	Improve recall using larger, more diverse datasets.	Good real-time detector but needs refinement for early-stage symptoms.
4.	Rice Leaf Disease Detection Using Machine Learning Techniques- Dr. Abdul-Wahab Sami Ibrahim, Dr. Baidaa Abdul khaliq Atya	2022	Detect bacterial blight, brown spot, and leaf smut using ML classifiers.	Compared Logistic Regression, KNN, Decision Tree (J48), and Naïve Bayes on 480 augmented images.	Decision Tree achieved the best performance (95.2% accuracy), outperforming other classifiers.	Explore ensemble learning and larger, higher-quality datasets.	Simple ML models can be effective but limited for real-field scalability.
5.	Classification of Rice Leaf Blast Severity Using Hyperspectral Imaging- Guosheng Zhang, Tongyu Xu and Youwen Tian	2022	Classify severity levels of rice blast using HSI.	Used SVM and PNN with spectral features across seasons.	Consistently high classification performance across datasets.	Integrate severity classification with UAV-based monitoring.	Good accuracy but scalability limited by sensor and processing complexity.

6.	Deep Learning for Rice Leaf Disease Detection: A Systematic Literature Review-Chinna Gopi Simhadria, Hari Kishan Kondaveetia, Valli Kumari Vatsavayib.	2024	Provide a comprehensive review of DL approaches in rice disease detection.	Analyzed CNN, YOLO, transfer learning, and UAV/HSI methods across multiple studies.	DL achieves high accuracy (up to ~93.58%) but struggles with dataset imbalance and generalization	Create standardized benchmarks and lightweight field models.	Excellent survey; highlights challenges that motivate current research.
7.	SSD-YOLO: Lightweight Network for Rice Leaf Disease Detection- Canlin Pan, Shen Wang, Yahui Wang and Chaoyang Liu	2025	Develop a lightweight YOLO model for rice disease detection on mobile/edge devices.	Designed SSD-YOLO; tested on brown spot, blast, bacterial blight.	Achieved 95.19% accuracy, mAP 95.32%, and up to 99.48% per-disease accuracy.	Extend to multi-crop, cross-regional detection systems.	Very practical for real-world deployment due to speed + accuracy.
8.	Transformer-Linked YOLO (TLI-YOLO) for Rice Disease Detection- Zhuqi Li, Wangyu Wu, Bingcai Wei, Hao Li 1, Jingbo Zhan 1, Songtao Deng and Jian Wang	2025	Improve YOLO's ability to detect small lesions.	Integrated lightweight transformer attention with YOLO.	Outperformed baseline YOLO in lesion detection and mobile deployment.	Extend to multi-disease classification across different crops.	Promising hybrid model, though metrics need standard reporting.

Across these studies, **deep learning and transformer-based models consistently outperform traditional ML methods** in both accuracy and robustness. **YOLO variants** are leading candidates for real-time detection, while **HSI and UAV approaches** enable early-stage and large-scale monitoring. However, issues such as **high computational demands**, **limited dataset diversity**, and a **lack of affordable deployment solutions** remain as major research gaps.

Method type	Representative models	Key improvements	Feature
Traditional ML	SVM + GLCM/ILMFD	Combines GLCM and multi-fractal dimension features	Manual feature engineering, poor generalization
	AFKMRG + Enhanced LSTM (Sahu and Minz, 2023)	Adaptive region-growing segmentation with FSJ-FOA optimization	High computational complexity
Lightweight CNN	DM-YOLO	DySample + MPDIoU	Low sensitivity to small lesions
	YOLO v5s-ours	CA attention + ASFF + ASPP	High real-time demand (30.7 fps)
	GDS-YOLO	GsConv + DySample + SCAM	23% fewer parameters
Attention Mechanisms	EMA-YOLOv8	SPD-Conv + EMA	Model size: 7.1MB
	TomatoGuard-YOLO	MPIRU + DFAF	Ultra-lightweight (2.65MB)
	RGC-YOLO	RepGhost + CBAM	Optimized for embedded devices
Multi-Scale Fusion	Bi-FAPN-YOLOv5	DenseNet-201 + Bidirectional Feature Attention Pyramid	Requires model pruning
	GhostNet_Triplet_YOLOv8	GhostNet + Triplet Attention	50.2% smaller model size

Fig 8 : Comparison of existing disease detection methods.

IV. PROPOSED SYSTEM: (YOLO V8 ARCHITECTURE)

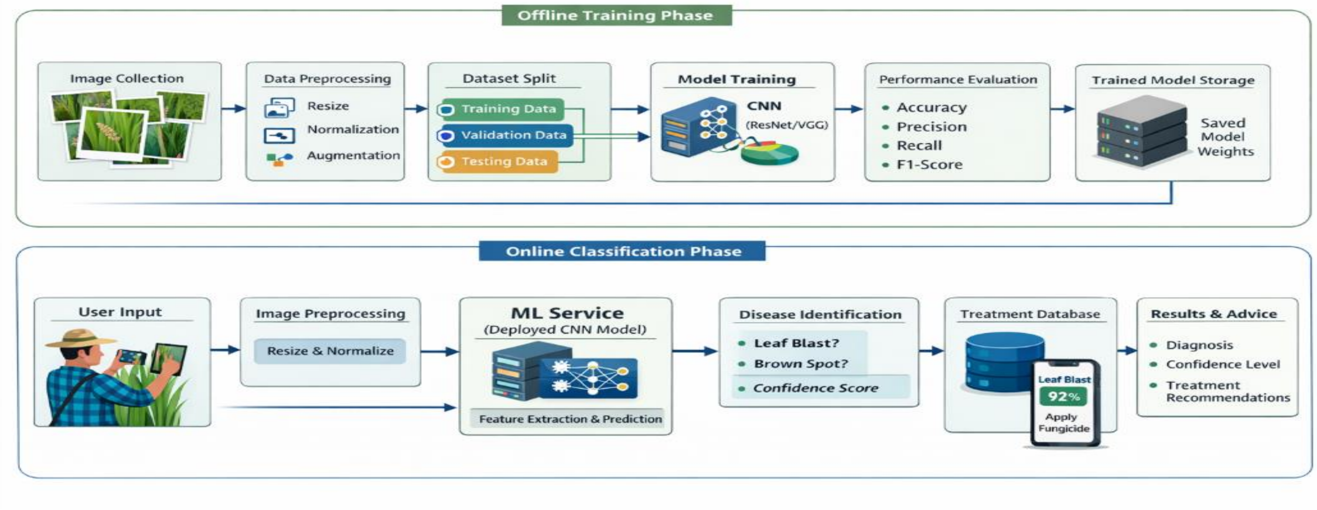

Fig 9: System Architecture

System Architecture:
The system is designed using two stage approach that applies deep learning techniques. These two stages are first the training phase and second the online or classification phase, each serving a specific purpose in building and using the model.

1. **Offline(Training) Phase**
   This phase operates on developer's side and focuses on developing reliable and accurate prediction model
   - **Raw image or dataset collection:**
     Large number of diseased rice leaf images are collected under different conditions with varying lights, angles and backgrounds. More the number of images, higher the chance for model to learn diverse pattern
   - **Data pre-processing and augmentation:**
     The collected images are prepared for training by resigning them into a uniform size and applying data augmentation techniques like rotation, flipping, zooming etc.
   - **Dataset splitting:**
     Here the normalized dataset is divided into three parts, training, validation and testing in ratio of 70:15:15 each split place crucial role in model development
   - **Model training:**
     70% division of dataset is used to train the model. This dataset is used to teach deep learning model, typically a CNN to identify disease pattern in the leaf images.
   - **Validation process:**
     During training validation, dataset is used to check models performance and hyper tone the parameters. This helps avoid overfitting and ensure better accuracy.
   - **Performance evaluation:**
     after successfully completing the training, the model is tested on new/unseen data to evaluate its accuracy using metrics like accuracy, precision, recall, and F1 score

   Once the model performs or meets the required metrics, its trained weights are saved so that it can be used later in real world application

2. **Online real time classification phase**
   - **User input:**
     Here the end-user captures or uploads image of rice leave to the deep learning model
   - **Image pre-processing:**
     The uploaded image is normalizing and resized to required size and noise is reduced.
   - **Disease identification:**
     Here the input image uploaded by user is sent to deployed machine learning server where it analyses the image and predict the type of disease
   - **Research generation:**
     Once the system successfully identifies the disease, the API display the name of disease and also provides a confidence score and treatments for the disease

Object Detection Architecture of Yolo V8

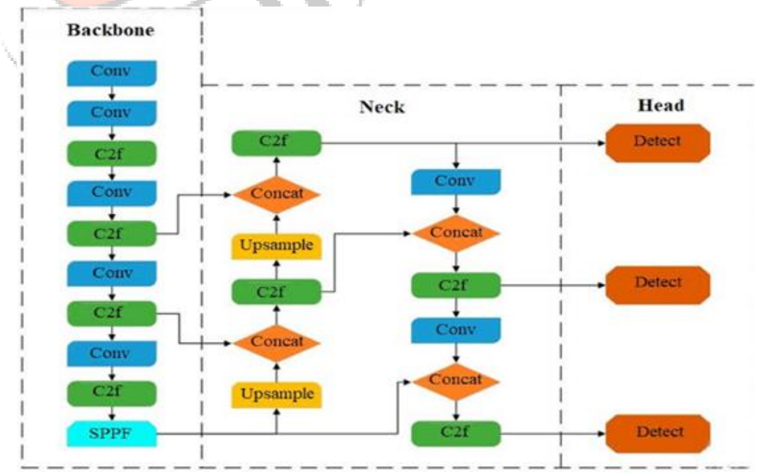

Fig 10: YOLO v8 Architecture

Object detection architecture of Yolo V8 is a cutting-edge object detection system designed for high speed and accuracy. It follows modular structure divided into three main components: **Backbone, Neck, And Head** each place different role in object detection.

1. **Backbone (Feature Extraction)**
The left section of given diagram is the backbone part. It is foundation of model responsible for extracting feature maps from raw input. It takes the input image from user and extract important features such as edges, shapes, patterns, textures etc.

Main components:

- Conv (convolution layer): detect basic patterns (edges, color).
- C2F Block: enhances feature and improves detection accuracy by integrating high-level features with contextual information.
- SPPF (spatial pyramid, pooling fast): provides multi-scale representation of captures feature map.

Output: different levels of feature map low level to high-level information

2. **Neck (feature fusion)**
This acts as bridge; it fixes features from different layer to detect object of different sizes. Fusing features from different scales to ensure the model detects object of varying sizes.

Main operations:

- Up sample: increases resolution helps detect small object with help of nearest neighbor.
- Concat(concatenation): merges feature from different layer.
- Conv+ C2F: refine combined features.

This layer combines zoomed in and out view to understand the full picture

3. **Head (detection layer)**
This layer carryout's final production. It is towards the right side of diagram, it predicts boundary boxes, class labels, confidence score.

V. Experimental Analysis

To evaluate the effectiveness of the proposed rice leaf disease detection model, a series of experiments were conducted using a multi-class dataset consisting of seven categories: Brown Spot, Healthy Rice Leaf, Hispa, Leaf Blast, Leaf Scald, Sheath Blight, and Tungro. The dataset was divided into training and validation subsets to ensure unbiased performance evaluation. The model was trained for 50 epochs using standard optimization techniques. Performance was assessed using training loss, validation loss, Top-1 accuracy, Top-5 accuracy, and confusion matrix analysis. Result comparison with different machine, learning algorithms is shown in the table below.

Table 1 Result Comparison with different machine learning algorithm

Machine Learning Algorithm	Correctly classified Instance (%)	Percentage
Naïve bayes	93.5%	73.2
Support Vector Machine	94%	84.9
J80 Decision tree	80%	76%
J48 With bagging	95.34%	92.3%
Random Forest	97.3%	95.5%

Fig 11: Result Comparison with different machine learning algorithm

A. Training and Validation Performance:

This figure contains 4 plots:

- Train Loss
- Validation Loss
- Top-1 Accuracy
- Top-5 Accuracy

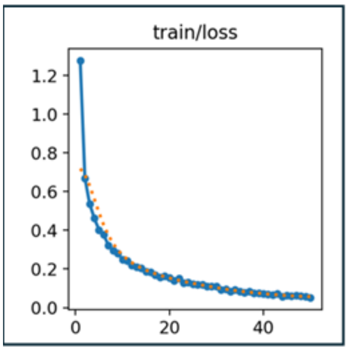

Fig 12: trail/loss

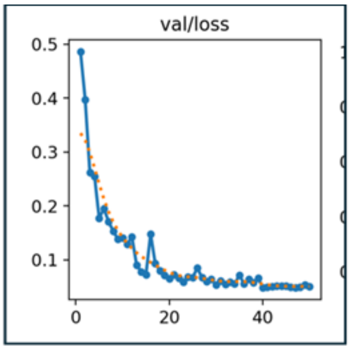

Fig 13: Val/loss

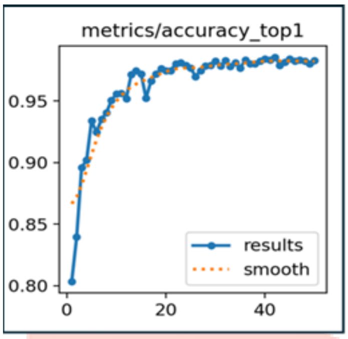

Fig 14: top-1 accuracy

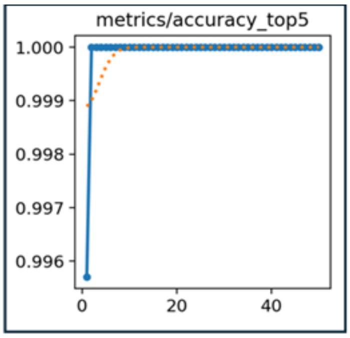

Fig 15: top-5 accuracy

a. **Train Loss (train/loss)**

- Loss starts high (~1.25)
- Rapidly decreases in first 10 epochs
- Gradually stabilizes around ~0.05
i.e. The model is learning properly, No instability or exploding gradients and Smooth convergence. After ~25 epochs, improvement becomes minimal → model has nearly converged

b. **Validation Loss (val/loss)**

- Starts around ~0.48
- Drops sharply
- Stabilizes around ~0.05
i.e. Validation loss follows the same pattern as training loss, No large gap between train and validation loss and No overfitting

c. **Top-1 Accuracy (metrics/accuracy_top1)**

- Starts around ~80%
- Quickly rises above 90%
- Final accuracy ≈ **98%**

i.e. Top-1 accuracy means: Model's first prediction is correct. Model correctly predicts the disease in first attempt 98% of the time.

d. **Top-5 Accuracy (metrics/accuracy_top5)**

- Starts ~94%
- Reaches almost **100%**

Top-5 accuracy means: Correct class is among top 5 predicted classes. Since Model has only 7 classes, this is expected to be near 100%.

B. Confusion Matrix Analysis:

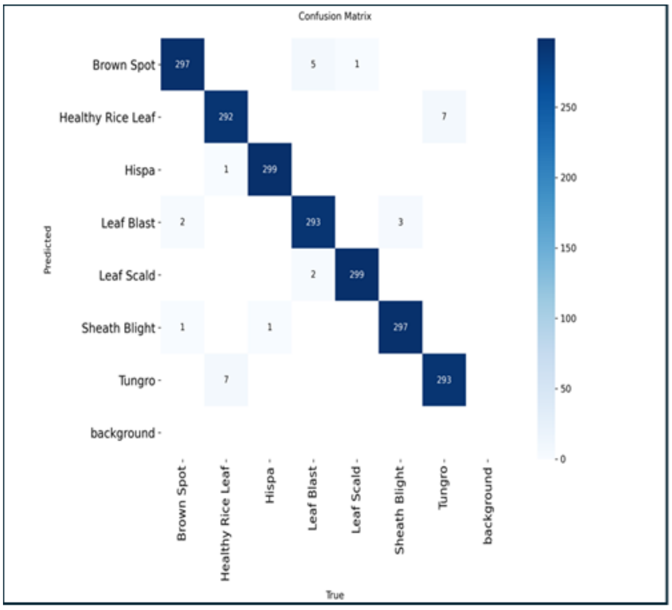

Fig 16: Normalized Confusion Matrix

Normalized Confusion Matrix:
This shows percentage-based performance per class.

Disease	Accuracy
Brown Spot	0.99(99%)
Healthy	0.97 (97%)
Hispa	1.00 (100%)
Leaf Blast	0.98 (98%)
Leaf Scald	1.00 (100%)
Sheath Blight	0.99 (99%)
Tungro	0.98 (98%)

Most classes above 97% accuracy. Leaf Scald and Hispa are perfectly classified.
Very small confusion between:

- Brown Spot & Leaf Blast
- Hispa & Brown Spot

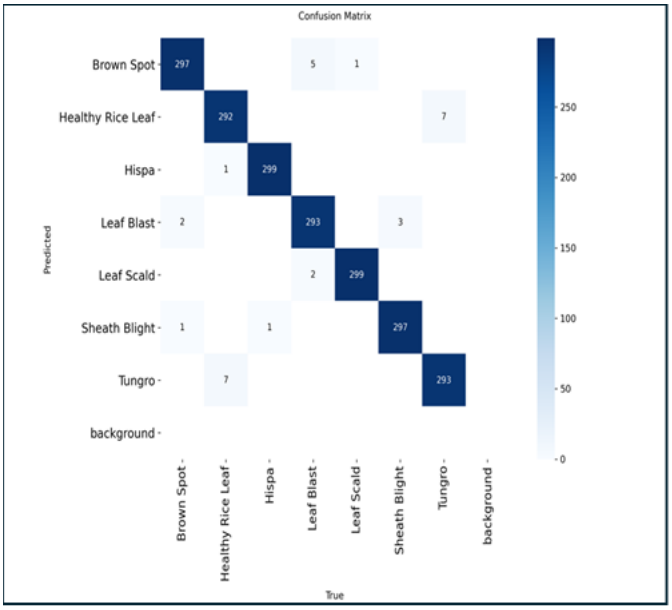

Fig 17: Confusion Matrix

Raw Confusion Matrix (Actual Counts):
This shows exact number of correctly and incorrectly classified images.
Examples:

- Brown Spot → 297 correctly classified
- Leaf Blast → 293 correct classified
- Leaf Scald → 299 correct classified
- Tungro → 293 correct classified
- Healthy Rice Leaf → 292 correctly classified (few misclassified)
- Sheath Blight → 297 correctly classified
- Hispa → 299 correctly classified

Very few misclassifications (1–6 images only).

C. Overall Model Evaluation:

- **98%** Top-1 Accuracy
- Nearly **100%** Top-5 Accuracy
- Very low validation loss(~ 0.05)
- No overfitting
- High per-class accuracy **(>95%)**
- Strong generalization

The experimental results indicate that the proposed deep learning model successfully learns discriminative features for multi-class rice disease classification. The smooth convergence of loss curves, high classification accuracy, and strong per-class performance demonstrate the robustness and reliability of the system. The small number of misclassifications can be attributed to similarity in disease patterns at early stages. However, the overall performance confirms the suitability of the proposed approach for practical agricultural applications.

VI. CONCLUSION

In this research and image processing technique and machine learning algorithm is used to classify and detect rice disease into seven categories, namely **Brown Spots, Leave Scald, Leaf Blast, Sheath Blight, Tungro, Hispa and Healthy Leave**. The aim of this research is to detect the disease in the plant and categorize them using **YOLOv8** as it is a lightweight and easy to use algorithm with high rate of accuracy. For this purpose, image processing, feature extraction, and machine learning algorithm is used in the proposed system.

For the performance analysis, propose algorithm was compared with different machine learning algorithms like classification, regression, Yolo V5, SVM, Naïve Bayes, etc. With **more than 7000 image dataset** that is more than thousand images per disease, 90% to 10% ratio is used to split the data into training and testing data. It was observed that accuracy of proposed system is slightly higher than other algorithm. However, for other algorithm it drops rapidly when tested using real time captured images which may or may not be present in training model, this shows that the accuracy of machine learning model with real Time comparison is different from splitting technique. As since Yolo V8 uses partial CNN technique in its algorithm, the accuracy remains same for various datasets as it minimizes the data overfitting issue in the collected dataset for the performance standpoint. The accuracy is tested using n-fold cross validation method, it was observed that accuracy is higher and proposed algorithm is correctly classifying the disease with more than 98% of accuracy rate.

VII. REFERENCES

[1] D. B. A. k. A. Dr. Abdul-Wahab Sami Ibrahi, "Detection of Diseases in Rice Leaf Using Deep Learning and," Webology, vol. 19, no. 1, p. 11, january 2022.

[2] M. E. Haque, A. Rahman, I. Junaeid, S. U. Hoque and M. paul, "RICE LEAF DISEASE CLASSIFICATION AND DETECTION USING YOLOv5," arXiv.org, Bathurst, Australia; Feni, Bangladesh, 2022.

[3] Z. Li, W. Wu, B. Wei, B. Wei, H. Li, J. Zhan, s. Deng and J. Wang, "Rice Disease Detection: TLI-YOLO Innovative Approach for enhanced detection and mobile compatibility," Sensors, vol. 25, no. 2494, p. 26, 25 April 2025.

[4] L. Feng, B. Wu, Y. He and C. Zhang, "Hyperspectral Imaging Combined With Deep Transfer Learning for Rice Disease Detection," Frontiers in Plant Science, vol. 12, no. Article 693521, p. 14, 29 September 2021.

[5] G. Zhang, T. Xu and Y. Tian, "Hyperspectral Imaging-Based Classification of Rice Leaf Blast Severity Over Multiple Growth Stages," Plant Methods (Springer Nature), vol. 18, no. Article 123, p. 10, 2022.

[6] H. Zhu, C. Lin, G. Liu, D. Wang, S. Qin, A. Li, J.-L. Xu and Y. He, "Intelligent Agriculture: Deep Learning in UAV-Based Remote Sensing Imagery for Crop Diseases and Pests Detection," Frontier, vol. 15, no. Article 1435016, p. 15, 24 October 2024.

[7] C. Pan, . S. Wang, Y. Wang and C. Liu, "SSD-YOLO: A Lightweight Network for Rice Leaf Disease Detection," Frontier, vol. vol.16, no. Article 1643096, p. 12, 18 August 2025.

[8] C. G. Simhadri, H. K. Kondaveeti, V. K. Vatsavayi, A. mitra and P. Ananthachari, "Deep Learning for Rice Leaf Disease Detection: A Systematic Literature Review on Emerging Trends, Methodologies and Techniques," Information Processing in Agriculture, vol. Vol. 12, no. Pages 151–168, 2025.