International Journal of Smart Agriculture (IJSA)

ISSN: 2995-9829 (ONLINE)

VOLUME 3 ISSUE 1 (2025)

PUBLISHED BY
E-PALLI PUBLISHERS, DELAWARE, USA

International Journal of Smart Agriculture (IJSA)

Volume 3 Issue 1, Year 2025
ISSN: 2995-9829 (Online)
DOI: https://doi.org/10.54536/ijsa.v3i1.6188
https://journals.e-palli.com/home/index.php/ijsa

A Baseline Study of End-to-End CNN Models for Paddy Leaf Disease Classification Using Extended Datasets

Lele Mohammed[1], Aminu Aliyu Abdullahi[1*]

Article Information

Received: September 20, 2025

Accepted: October 24, 2025

Published: December 06, 2025

Keywords

Accuracy, Classification, CNN, Rice Leaf Disease

ABSTRACT

Classification of plant leaf diseases is one of the interesting areas of study in machine learning. Prompt identification and treatment of paddy plant diseases can boost rice production across the world. Rice is one of the major staple food that is consumed worldwide. Its production needs to be enhanced significantly. Rice farmers suffer a great deal by means manual detection of the type of diseases that infect their farmlands. Deploying technology can greatly help farmers to tackle problem of paddy diseases. Convolutional Neural Network (CNN) is an image processing method that can be useful in ensuring higher rice production. The CNN model was used to carry out this research owing to its uniqueness in producing higher detection accuracy. This research work is crucial for its importance in making sure rice crops are cultivated in large scale throughout the world. Four classes of rice diseases are selected for this research, namely: Bacterial leaf blight, rice leaf blast, Tungro and rice brownspot. Although the rice disease dataset is not easily available, this research used datasets of about 5932 images to experiment with CNN on more datasets. The proposed model has achieved an improved accuracy of 99.12% from the benchmark paper.

INTRODUCTION

Rice is today, one of the most consumed food worldwide Shreya *et al.* (2020). Boosting rice production across the globe has therefore become imperative to cover the high demand. Farmers need to acquire proper knowledge to classify plant leaf diseases in their farmlands. The knowledge is not readily available in rural areas. Therefore, most farmers use manual method to deal with plant leaf diseases. This practice has led to the decline in rice production on account of the inaccuracies involved. Deployment of technology in agriculture has reduced a lot of errors brought about by humans. In today's world, farmers need automated systems to tackle problem of rice plant diseases. Various techniques and methods can be used to design a model that can detect and classify diseases of plant accurately. Classification of plant leaf diseases in agricultural field today, has become a notable topic of research in the aspect of disease recognition Shunmugam *et al.* (2019). Images of the affected plants can be collected from the field for use by the model in the classification process. However, obtaining large datasets is a big challenge that is why some of the previous research on image processing and classification worked with small datasets. Furthermore, CNN performs well with huge data making it more accurate and efficient. In a benchmark research conducted by Shreya *et al.* (2020), 92.46% accuracy was obtained due to small datasets. This research work has contributed 7.20% increment by experimenting with more datasets. Convolutional Neural Network (CNN) technique was used to design the proposed model as in the benchmark paper. CNN is preferred to other techniques for its high classification accuracy and also the ability to handle huge data Swathika *et al.* (2021). CNN works well with huge datasets. As a modern approach in image classification, CNN has the capacity to produce definite diagnosis Pratapagiri (2021). Image processing methods are important in recognising diseases of plant promptly before damage is done. To increase rice production greatly, robust techniques such as CNN must be deployed. Deployment of smart agriculture can boost economy and bring about sustainable development (Polwaththa *et al.,* 2025).

LITERATURE REVIEW

Swathika R. *et al.* (2021) conducted research to highlight the existing research on image recognition using CNN. The authors introduced a new model for rice disease recognition using CNN. The authors used about 3500 datasets and their model achieved 70% accuracy on classification. They have also highlighted on need for using more data on the future work. Mathew B, *et al.,* categorised CNN as the best method for image detection by virtue of its automatic feature extraction trait. Their work made comparison between different learning models and revealed that CNN has the highest classification accuracy when compared to other image processing methods.

Norhalina *et al.* (2020) performed research on paddy disease recognition using CNN. They built a five layers CNN model that produced accuracy of 93% using 3355 datasets of healthy and diseased images. They highlighted that CNN gave the best accuracy as compared to other sophisticated methods. Md. Ashiqul (n.d.) conducted a research using deep learning CNN. They studied four

[1] Department of Computer Science, Faculty of Computing, Federal University Dutse, Nigeria
[*] Corresponding author's e-mail: aminuabdullahisumaila@gmail.com

different models and achieved an accuracy of 92.68% using Inception-ResNet-V2. Shreya *et al.* (2020) built a VGG-16 architecture with transfer learning to classify four different classes of rice diseases and achieved an accuracy of 92.46%. The authors use few datasets to design their model and so they highlighted on the need to fill in the gap their research created by using more datasets to improve on the accuracy.

Plant Disease Detection Neural Network (PDDNN) architecture can be implemented using CNN model to identify diseases of plants Hari *et al.* (2019) . If the disease of plants cannot be detected on time, damage can occur beyond remedy. For prompt and accurate detection, models like CNN must be used to prevent that occurrence S. Hari *et al.,* used CNN architecture to achieve a classification accuracy of 86.00%. Rinu *et al.* (2020) used VGG16 CNN model to study thirty-eight different classes of plant diseases classification and achieved an accuracy of 94.8%. Merecelin *et al.* (2019) carried out a comprehensive research on plant leaf disease classification using CNN algorithm. Their system used 3663 datasets to achieve a classification accuracy of 87%.

Deep-CNN needs huge datasets to produce better accuracy. To obtain such data, image augmentation is used to expand the training data using different augmentation tools such as scaling, flipping, zoom and shift. Geetharamani *et al.* (2019) proposed a deep-CNN model trained on 55,636 augmented training datasets from 38 classes of plant leaves. Gajjar *et al.* (2021) built a real-time system to classify crop plant diseases using deep-CNN. Their system achieved an accuracy of 96.88%. Yang *et al.* (2017) developed a rice classification model to prove the robustness of CNN over conventional learning methods in giving out good result. The authors used 500 images of healthy and diseased paddy leaves and achieved an accuracy of 95.48%.

Sony *et al.* (2020) conducted a study on three different rice leaf diseases using datasets obtained from UCI database. Their research produced significant result. Sharma *et al.* (2021) undertook a survey on 8 crucial paddy diseases using CNN techniques. They established that CNN produces highest accuracy of 93.58% in classifying bacterial blight and brownspot. Rishabh *et al.* (n.d.) developed an RDD model for Hispa rice disease classification using deep-CNN. The authors achieved 94% accuracy using one thousand datasets. Surya *et al.* (2020) performed a research on cassava leaf plant using CNN. The authors used Tensorflow library to achieve 85.38% training accuracy and 74.96% validation accuracy. The presence of Artificial intelligence (AI) in agriculture has brought positive advancements and innovations thereby changing the ways in which farming is done (Padhiary *et al.,* 2025).

Types of Paddy Diseases

The Images used in training the proposed model consist of four different classes of rice diseases, namely: Bacterial leaf blight, Rice blast, Brown spot and Tungro. Augmentation activities like zoom and rotate were performed on the datasets to create additional data needed for training of the model. Figure 1 (a)- (d) depict four classes of rice diseases that were worked upon using Tensorflow library of Python.

Figure 1: (a) Bacterial blight (b) Rice blast (c) brownspot (d) Tungro

Bacterial Leaf Blight

Bacterial leaf blight (Rice leaf blight) is a severe rice disease that is caused by a bacterium called Xoo (X. oryzae pathovar oryzae). It causes yellowish ooze or wilting of leaves. Crop destruction due to this disease could be up to 75% (Yen *et al.,* 2020).

Rice Blast

Rice blast also called neck blast is a deadly paddy fungal illness that is caused by a fungi called Magnaporthe oryzae. It causes high economic loss of about 30% of the world annual yield Upadhyay (2020). Blast of rice can be recognised by the appearance of lesions on leaves, pedicles and seeds.

Brown Spot

Rice brown spot disease is one of the vital illnesses of rice plant that is caused by a fungi called D. oryzae. It can attack seedlings and grown plants. It is one of the destructive rice disease that could lead to total loss of yield in the farmlands. It mostly appears in regions where water is scarce Mau Yosep *et al.* (2020). Symptoms of brown spot includes oval shaped brown spots and discolouration of stems.

Tungro

Rice tungro disease is one of the disastrous paddy viral disease in the world, most especially in South and Southeast Asia. It is caused by amalgamation of two viruses, namely: Rice tungro baciliform and rice tungro Spherical viruses (Bunawan *et al.,* 2014). These viruses can cause serious economic damage. Symptoms of rice tungro includes; Stunting of leaves, dark brown specks, yellow or orange leaves.

Dataset Collection

The dataset used in this research work was collected from Mendeley dataset repository. The link is https://data.mendeley.com/datasets/fwcj 7stb8r/ 1. Total images collected were 5932, comprising 4 classes of rice leaf diseases, namely: Bacterial leaf blight, rice blast, brownspot and rice tungro. The dataset was split into 80:20 ratios for training and validation respectively. 4746 files were allocated for training of the model and 1186 files for validating it. The CNN technique was used to design the model. The size of the image used was 64X64 pixels, RGB colour image.

Research Framework

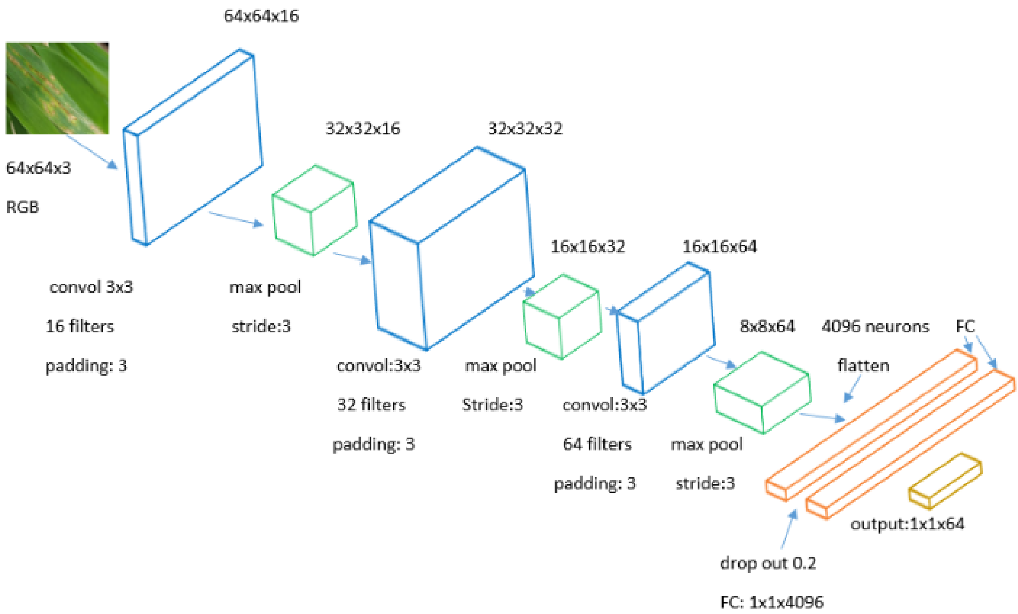

Figure 2: Research Framework

MATERIALS AND METHODS

There are three main parts that form the CNN architecture, namely: Convolutional layers, pooling layers, and fully connected layers.The convol layer and the corresponding pooling layers make up the feature extraction process while the FC layer forms the classification process. The convol layer extracts features from the images by performing mathematical operations. The pooling layer reduces the size of the output of the convol layer called feature map. The reduction is to costs of computation. Although, the dropout layer and the activation function are also regarded as layers in the CNN framework, they are generally considered as significant parameters. The former aid the model in dealing with overfitting while the latter helps in adding non-linearity to the neural network. Five layers have been used in this project work. Three convolutional layers with their corresponding pooling layers and two fully connected layers. The fc layer was followed by ReLu activation function. The first convolutional layer uses 16 filters, second uses 32 filters and third layer uses 64 filters respectively. Maxpooling layer and dropout layers were also used. Total trainable parameters used were 286,052. The developed model was built using python and executed on the google colab platform using keras and tensorflow version 2.6.0 respectively. The keras library reads data from the directory to process the input image (64X64), and the fit () function trains the model.

Implementation

The proposed model was implanted as follows:

Setup

The experiments were carried out on google colab with windows 10 with 8GB RAM, 64-bit OS. Keras and Tensorflow versions 2.6.0 were used to run the codes on the Google colab.

Image Collection

The datasets comprising four classes of rice diseases, namely; Rice blast, Bacterial blight, tungro and brownspot were used in this project. Total of 5392 images were wholly acquired from Mendeley datasets repository.

Pre-Processing And Augmentation

The images collected from Mendeley were sized 64X64 pixels and augmentation approaches like rotation, zoom and rescaling were used to expand the data for training purposes.

Proposed Model Training

The files were split into 80:20 ratios and then mounted for training and validation of the model. The model was stopped at exactly 20th epoch. This is because the error rate of the testing data was negligible at that point. To reduce losses and speed up result, Adam optimiser was used. The developed model has followed the classification process from the image collection stage up to the classification as shown in Figure 1 below.

Justification

CNN model was chosen for its ability to produce higher accuracy and stability.

RESULTS AND DISCUSSIONS

Values Obtained

The proposed model has used 5932 files, out of which 4746 were used training and 1186 for validation. The model has produced an improved classification accuracy of 99.12% which is higher than that of the previous work conducted in the base paper. The batch size, image size and the kernel used were 16, 64x64, and 3x3 pixel respectively.

The model was run for 25 epochs using 4746 and 1186 training and testing data respectively. After running the datasets successfully, the model produced an improved accuracy of 99.12%. The accuracy was obtained after taking average of 20 runs as shown in Table 1.

Table above, shows how the average value of the classification accuracy was calculated using 20 runs. The batch size used was 16, image size was 64x64 pixels, and the number of the epochs used were 25. Adam optimizer was used to get faster result and also minimize the losses. The figure 3 below, depicts the training and testing accuracy of the proposed model.

Table 1: The average value of the classification accuracy

Run no	Result
1	99.49%
2	99.24%
3	99.49%
4	98.65%
5	99.07%
6	99.41%
7	98.57%
8	98.74%
9	99.58%
10	99.24%
11	99.16%
12	98.82%
13	99.58%
14	98.99%
15	98.90%
16	99.16%
17	99.24%
18	99.16%
19	99.24%
20	98.74%
Total	= 1982.47
% Average	=1,982.47/20 = 99.12%

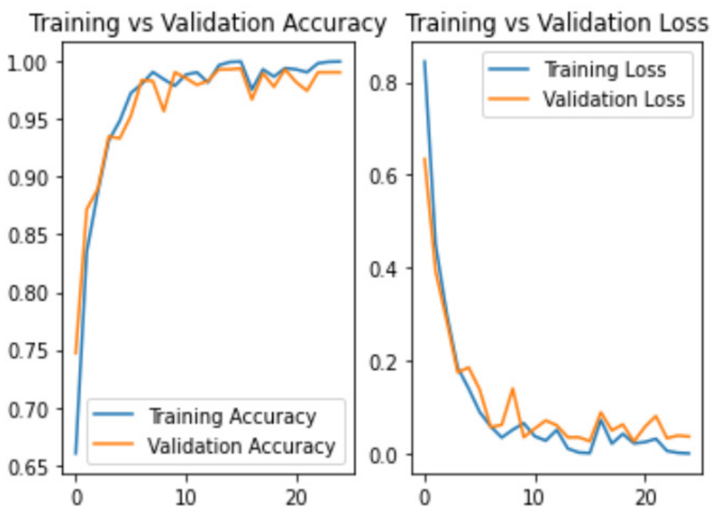

Figure 3: Training and validation accuracy plot

Error Analysis

The figure 4 below, shows the confusion matrix depicting classified and misclassified images.

The confusion matrix depicts the distribution of model classification on each type of disease.

- Bacterial blight: Leaf blight used 294 images for testing. 4 were misclassified. 2 were misclassified as tungro and the other 2 were misclassified as rice blast.
- Rice blast: Total of 292 images of rice blast were classified. Only 2 were misclassified. 1 was misclassified as tungro and the other as brownspot.
- Brownspot: 330 images of brownspot were correctly classified as brownspot, while 3 were misclassified as rice blast.

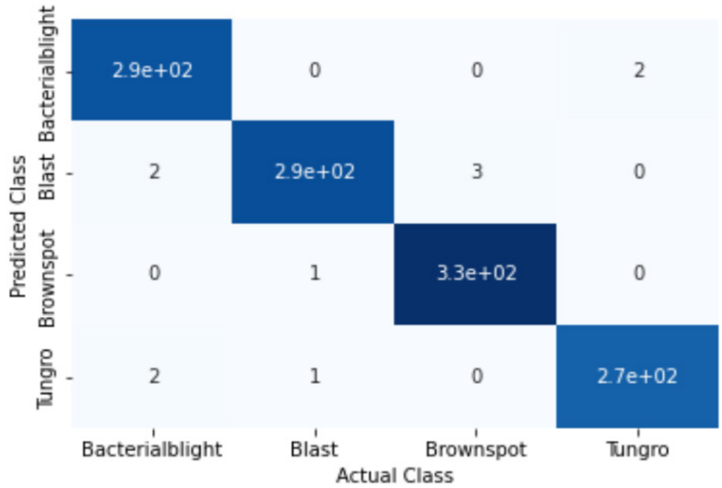

Figure 4: Confusion matrix

• Tungro: 270 images of tungro disease were correctly classified. 2 images were misclassified as leaf blight.

Classification Chart

The proposed model has classified the testing data as presented in fig 5 below.

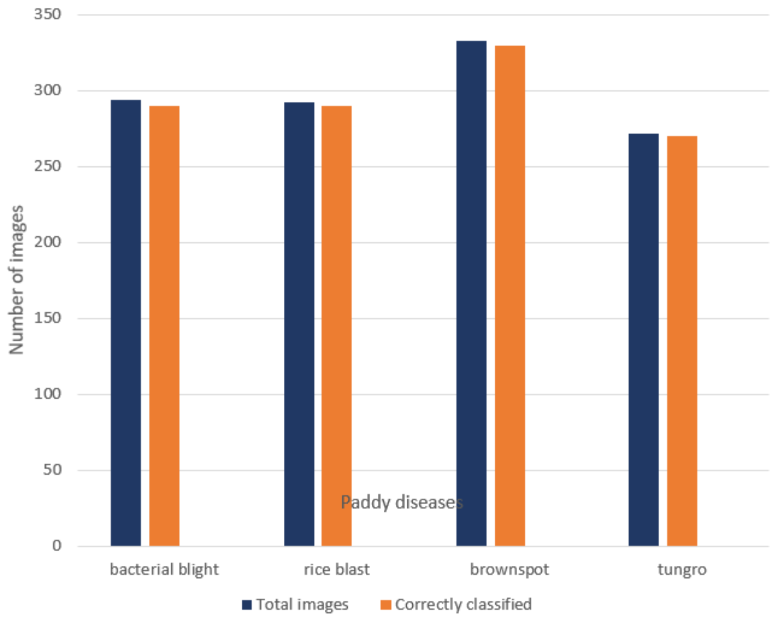

Figure 5: Classified data chart

Classification Metrics

Three classification metrics, namely: Recall, F1-score and precision value were used to validate the proposed model for accuracy. The results are shown in table 2 below.

Table 2: Classification Matrix table

	Precision-Value	Recall	F1-score
Leaf Blight	0.9862	0.9931	0.9896
Rice Blast	0.9932	0.9832	0.9882
Brownspot	0.9901	0.9970	0.9940
Tungro	0.9926	0.9889	0.9907

Comparison of Validation Metrics

The figure 6 below, shows the comparison chart for the validation metrics used.

Table 1 and figure 6 show validation matrix value for the four classes of rice disease. The model depicts superb performance across the four rice diseases, with exceptional values. Brown Spot performs the highest F1-score of 0.9940 and recall of 0.9970, showcasing superior result. Leaf Blight follows greatly with an F1-score of 0.9896, performing high in both precision (0.9862) and recall (0.9931). Rice tungro did well, with optimised precision of 0.9926 and recall of 0.9889, giving an F1-score of 0.9907. Finally rice Blast has the highest precision of

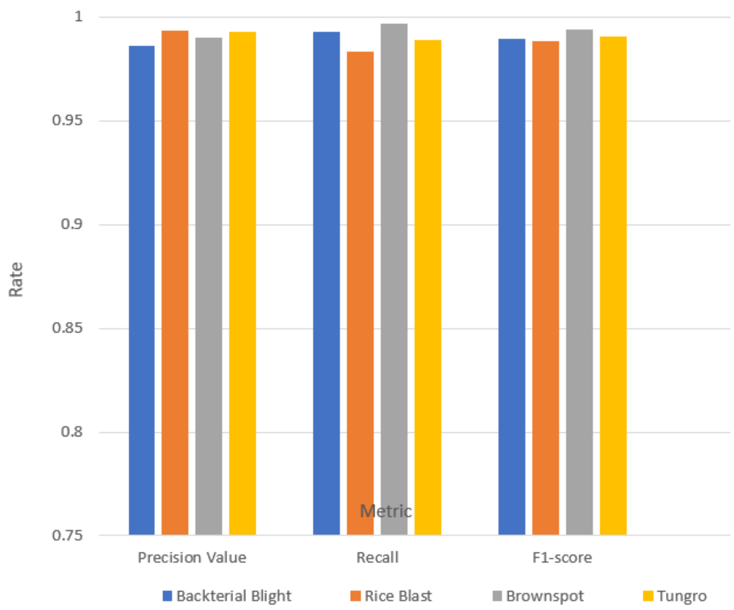

Figure 6: Validation metrics comparison chart

0.9932 but a negligibly lower recall of 0.9832, culminating in a decreased F1-score of 0.9882.

Generally, the model is greatly reliable, with negligible false positives or negatives.

Percentage Increase

After obtaining the final test accuracy the proposed method was compared in the model in [1] to determined the percentage increase.

Percentage Increase = (Difference/Original Value) * 100 % ={(99.12%)- (92.46%)/(99.12%) x 100
= 7.20% Increase.

CONCLUSION

This project work has used convolutional neural network (CNN) algorithm to classify four classes of different rice diseases of leaf blight, rice blast, brownspot and tungro. The model used 5932 images in total, with 4746 images as training data and 1186 as testing data. The model was trained with 80:20 ratios. 80% for the training and 20% for validation respectively. The model has stopped at epoch 25th after detecting negligible error rate in the testing data and a stable accuracy that was not improving. The proposed model has finally achieved an accuracy of 99.12%. Achieving additional 6.66% accuracy (7.2% percentage increase) over the model used in the bechmark paper (Shreya & Kamal, 2020). This has further cemented the fact that CNN offers higher accuracy when used with more datasets. Therefore CNN is data-hungry.

In future research, hybrid model could be developed to overcome the limitations of CNN. This will safely allow for deployment of lightweight model with high accuracy, reduced processing power and low-latency. The model can be deployed on resource constraint environment for use on edge devices. It can also be used to boost agricultural production in Nigeria through deployment of AI in the field of rice farming.

REFERENCES

Annem, S., & Chevula, J. (2020). Paddy diseases recognition using convolutional neural network. *International Journal of Engineering Applied Sciences and Technology, 5*, 358-362. https://doi.org/10.33564/IJEAST.2020.v05i03.057.

Bunawan, H., Dusik, L., Bunawan, S., & Mat Amin, N. (2014). Rice tungro disease: From identification to disease control. *World Applied Sciences Journal, 31*(6), 1221-1226. https://doi.org/10.5829/idosi.wasj.2014.31.06.610.

Deisy, M. F. C. (2019). Disease detection and classification in agricultural plants using convolutional neural networks — A visual understanding. *Proceedings of the 6th International Conference on Signal Processing and Integrated Networks* (pp. 1063-1068).

Gajjar, R., Gajjar, N., & Thakor, V. J. (2021). Real-time detection and identification of plant leaf diseases using convolutional neural networks on an embedded platform. *Visual Computer, 37*, 2099-2111. https://doi.org/10.1007/s00371-021-02164-9.

Geetharamani, G., & Pandian, A. J. (2019). Identification of plant leaf diseases using a nine-layer deep convolutional neural network. *Computers & Electrical Engineering, 76*, 323-338. https://doi.org/10.1016/j.compeleceng.2019.04.011.

Hari, S. S., Sivakumar, M., Renuga, P., Karthikeyan, S., & Suriya, S. (2019). Detection of plant disease by leaf image using convolutional neural network. In *International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN)* (pp. 1-5). https://doi.org/10.1109/ViTECoN.2019.8899748.

Islam, M., Shuvo, M., Shamsojjaman, M., Shazid, H., Hossain, M., & Tania, K. (2021). An automated convolutional neural network-based approach for paddy leaf disease detection. *International Journal of Advanced Computer Science and Applications, 12*(3), 100-105. https://doi.org/10.14569/IJACSA.2021.0120134.

Lu, Y., Yi, S., Zeng, N., Liu, Y., & Zhang, Y. (2017). Identification of rice diseases using deep convolutional neural networks. *Neurocomputing, 247*, 50-58. https://doi.org/10.1016/j.neucom.2017.06.023.

Matthew, B., & Ghidary, S. S. (2003). Convolutional neural networks for image processing: An application in robot vision. *Australian Joint Conference on AI*, 641-652. Springer. https://doi.org/10.1007/978-3-540-39468-8_74.

Mau, Y., Ndiwa, A., & Oematan, S. (2020). Brown spot disease severity, yield and yield loss relationships in pigmented upland rice cultivars from East Nusa Tenggara, Indonesia. *Biodiversitas Journal of Biological Diversity, 21,* 2073-2081. https://doi.org/10.13057/biodiv/d210443.

Norhalina, S., Muhammad, A., Rosziati, I., Wan, N. S., & Wan, H. N. (2020). An efficient convolutional neural network for paddy leaf disease and pest classification. *International Journal of Advanced Computer Science and Applications, 11*(7), 421-429. https://doi.org/10.14569/IJACSA.2020.0110716.

Padhiary, M., Kumar, K., Hussain, N., Roy, D., Barbhuiya, J. A., & Roy, P. (2025). Artificial intelligence in farm management: Integrating smart systems for optimal agricultural practices. *International Journal of Smart Agriculture (IJSA), 3*(1). https://doi.org/10.54536/ijsa.v3i1.3674.

Polwaththa, K. P. G. D. M. (2025). Enhancing protected agriculture with smart technologies: A review of innovations for climate-resilient farming. *International Journal of Smart Agriculture (IJSA), 3*(1), 1–24. https://doi.org/10.54536/ijsa.v3i1.5313.

Pratapagiri, S., Gangula, R., Srinivasulu, B., Sowjanya, B., & Thirupathi, L. (2021). Early detection of plant leaf disease using convolutional neural networks. In *2021 3rd International Conference on Electronics Representation and Algorithm (ICERA)* (pp. 77-82). https://doi.org/10.1109/ICERA53111.2021.9538659.

Rinu, R., & Manjula, S. H. (2021). Plant disease detection and classification using CNN. *International Journal of Research in Technology and Engineering (IJRTE), 10*(3), 152-156. https://doi.org/10.35940/ijrte.c6458.0910321.

Sharma, R., Das, S., Gourisaria, M., Rautaray, S., & Pandey, M. (2020). A model for prediction of paddy crop disease using CNN. In *Proceedings of the 2020 5th International Conference on Artificial Intelligence and Data Processing* (pp. 190-195). https://doi.org/10.1007/978-981-15-2414-1_54.

Sharma, R., Kukreja, V., & Kadyan, V. (2021). Rice diseases detection using convolutional neural networks: A survey. In *2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)* (pp. 995-1001). https://doi.org/10.1109/ICACITE51222.2021.9404620.

Shreya, G., & Kamal, S. (2020). Rice leaf diseases classification using CNN with transfer learning. *IEEE CALCON Conference,* 230-236. https://doi.org/10.1109/CALCON49167.2020.9106423.

Shunmugam, R., Ramesh, & Dharmar, V. (2019). Recognition and classification of paddy leaf diseases using optimized deep neural network with Jaya algorithm. *Information Processing in Agriculture, 7*(2), 129-138. https://doi.org/10.1016/j.inpa.2019.09.002.

Surya, R., & Gautama, E. (2020). Cassava leaf disease detection using convolutional neural networks. In *2020 6th International Conference on Science in Information Technology (ICSITech)* (pp. 97-102). https://doi.org/10.1109/ICSITech49800.2020.9392051.

Swathika, R., Srinidhi, S., Radha, N., & Sowmya, K. (2021). Disease identification in paddy leaves using CNN-based deep learning. In *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)* (pp. 1004-1008). https://doi.org/10.1109/ICICV50876.2021.9388557.

Upadhyay, K., & Bhatta, B. (2020). Rice blast (Magnaporthe oryzae) management: A review. *Agricultural Journal, 15,* 42-48. https://doi.org/10.36478/aj.2020.42.48.

Yen, H. (2020, April 1). *Rice bacterial blight. Encyclopedia Britannica.* https://www.britannica.com/science/rice-bacterial-blight.