Journal of Namibian Studies, 34 S2(2023): 1627-1636 ISSN: 2197-5523 (online)

Automated Detection and Classification of Rice Leaf Diseases Using Hybrid Deep Learning CNN

Josie C. Calfoforo[1*], Lenny M. Amar[2], Rayjand T. Gellamucho[3], Michelle I. Tribaco[4]

[1]Iloilo Science and Technology University – Iloilo City Campus, Iloilo City, Philippines, josie.calfoforo@isatu.edu.ph

[2]Iloilo Science and Technology University – Iloilo City Campus, Iloilo City, Philippines

[3]Iloilo Science and Technology University – Iloilo City Campus, Iloilo City, Philippines

[4]Iloilo Science and Technology University – Iloilo City Campus, Iloilo City, Philippines

Abstract

Rice (Oryza sativa) is one of the most widely cultivated crops in the Philippines. However, several diseases affect this crop, significantly reducing its quality and production. It is crucial to detect them early to stop the spread of infections. “This study proposed a deep hybrid learning CNN with transfer learning for detecting and classifying rice leaf diseases, namely bacterial leaf blight, brown spot, and leaf smut, with 40 images each”. The dataset was trained after the data pre-processing stage. A new model obtained 88.58% and 91.67% accuracy in the training and validation sets, respectively. Its performance metrics were analysed and evaluated based on the confusion matrix and classification report, such as “accuracy, recall (specificity), precision (Positive Predictive Value), and F-1 score”.

Keywords: Rice Leaf Disease Detection, Machine Learning, Hybrid Deep Learning, Convolutional Neural Network

Introduction

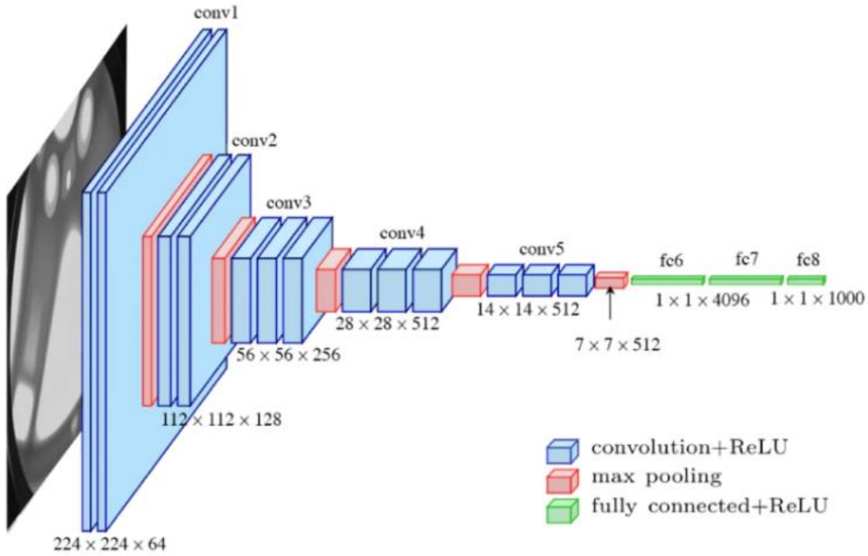
Machine learning significantly impacts everyone's life in business and academia due to modern computing technology. It enables the development of new products and gives firms knowledge of current consumer behaviour trends, practical business patterns, and agricultural economic conditions. “The agriculture industry deals with various issues, including ineffective farming methods, poor use of compost, manures, and fertilisers, inadequate water supply, various diseases that damage plants, and more. Diseases are incredibly destructive to a plant's health, affecting its growth. The attack of these many illnesses in the plants causes a significant loss in yield performance, both in quality and quantity". Since food is increasing rapidly due to worldwide population growth, automation for rice leaf disease in agriculture science is the primary concern of every country. To help this address, early detection and identification of plant diseases using leaf photos is an important and challenging research area in agriculture.

With the machine learning approach, the early detection and diagnosis for effective adapted precautionary measures of various rice leaf diseases help keep the disease from spreading (Chhillar et al., 2020). It will improve the production quantity and quality, reducing pesticide use and increasing the country's economic growth. The continuous observation of diseases that can be done manually may incur costs for farmers. Therefore, developing an automated system will support agronomists, pathologists, and even farmers to diagnose rice diseases more efficiently and take preventive measures in time.

In "Deep Learning (DL)", CNN is the most popular and commonly employed algorithm (Khan et al., 2020; Dhillon & Verma, 2020). It helps find patterns in images to recognise specific objects. Figure 1 shows that a typical CNN is similar to the "multi-layer perceptron (MLP)" and consists of numerous "convolution layers (CL)" preceding sub-sampling or "pooling layers (PL)". In contrast, the ending layers are fully connected (FC) layers architecture (Alzubaidi et al., 2021). "CNN significantly improved image identification by reducing the requirement for image pre-processing and enabling feature selection and extraction (Tunio et al., 2021). Most existing deep convolutional neural networks are trained with large rather than small datasets using deep convolutional neural networks because of overfitting when implementing the models".

In this study, Transfer Learning will be applied to generate the deep learning CNN model using the Rice Leaf Diseases Dataset downloaded from the UCI Machine Learning Repository with three (3) diseases: bacterial leaf blight, brown spot, and leaf smut containing 40 images each. This study will modify the "deep neural network (DNN)" and use this model to fit a small dataset. This learning aims to demonstrate a suitable adjusted deep learning model to be trained on ImageNet for image classification. It can be utilised to fit a minimal dataset without severe overfitting. Convolutional neural network (CNN) architectures will be the backbone of the transfer learning method. The performance parameters of a new model will be evaluated based on the performance metrics of the Confusion Matrix – "accuracy, sensitivity (True Positive Rate), specificity (recall), precision (Positive Predictive Value), and f1-score".

Figure 1. Typical CNN Architecture



Literature Review

Following the lead of earlier research, numerous scientists used hybrid models to achieve their goals. The classification of plant pictures into weeds and crops for selective herbicide spraying has been suggested as a better way to limit the use of herbicides. The development of image categorization methods made advantage of the deep learning function. For the neural network to be trained to distinguish between the classes of "Plant" and "Weed," datasets including photos of plants and weeds were prepared. With merely 250 photos of each plant in the dataset, the maximum efficiency of 96.3% was attained. The proposed model may be loaded into the Raspberry Pi rapidly, and a sprayer that is attached makes it possible to do selective spraying (Yashwanth et al., 2020).

Convolutional Neural Network (CNN) architecture was proposed based on VGG-16 with various rice leaf diseases containing 1509 images for the train set and 647 images for the test set. The model's performance, which had previously not produced adequate results on such a small dataset, has been significantly enhanced by Transfer Learning by fine-tuning the default VGGNet. Twenty-five epochs were the maximum number that was utilised in their study. The accuracy of the proposed model was obtained at 92.46% (Ghosal & Sarkar, 2020).

"A survey study was conducted on eight major rice diseases, namely bacterial leaf blight, false smut, rice hispa, blast, stemborer, sheath blight, brown spot, and brown planthopper, and explored using the CNNs technique. Their paper was divided into two major parts: survey methodology followed for conducting the work and a state-of-the-art used for rice disease detection (RDD) using the CNNs technique”. Based on their findings, the highest accuracy was obtained at 99% using ANN techniques in the case of non-CNNs, while 93.58% for CNN techniques on detecting bacterial leaf blight and brown spot disease (Sharma et al., 2021).

The research (Tunio et al., 2021) suggested a model for detecting three common rice leaf diseases: bacterial leaf blight, leaf smut, and brown spot (BLB, LS, and BS) using 119 photos. This model used a hybrid CNN-based deep learning technique. In the total experiment, deep learning tools based on Python were employed, and the accuracy was 90.8%. However, in the study by Krishnamoorthy et al. 2021, they used a pre-trained deep “convolutional neural network (CNN)” model of InceptionResNetV2 with a transfer learning approach for the detection of rice leaf diseases. The training and test sets for this study contained 4000 and 300 instances, respectively. “The suggested model's parameters were adjusted for the classification task and yielded a respectable accuracy of 95.67%. Furthermore, the author stressed that CNN is an algorithm of deep learning approaches that has been successfully used to handle computer vision challenges like image analysis, object segmentation, and picture categorization. It is a feed-forward ANN made up of three distinct layers: input, output, and hidden, and it effectively detects visual representations. Transfer learning is a method for using a previously trained CNN for a different problem. Transfer learning makes it possible to apply general-purpose learner models developed on a wide variety of images to a particular goal, such as the identification of plant diseases” (Farhood et al., 2022). Given that datasets used to identify plant diseases often have a limited number of training samples, this learning is advantageous.

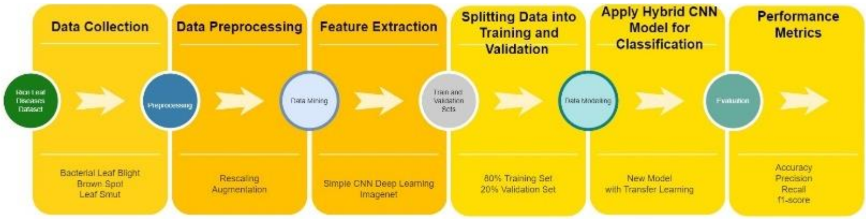
The previous review studies show that different types of technologies have been used to enhance the Deep Learning Hybrid Model of CNN, varying from most superficial to complex. Some researchers have utilised existing machine learning algorithms, whereas others have used their methods and algorithms to recognise images. This study will use a Deep Neural Network (DNN) to modify this model and be trained on ImageNet for image classification, a hybrid deep learning CNN-based approach. As a result, this will effectively be utilised to fit a tiny dataset without severe overfitting.

Methodology

This section presents a block diagram for the proposed work of this study. Figure 2 shows the data collection, pre-processing, feature extraction, splitting of data into training and validation sets, applying the Hybrid CNN model for classification, and evaluation of the performance metrics based on the confusion matrix and classification report, such as accuracy, recall, precision, and F-1 score.

Diagram

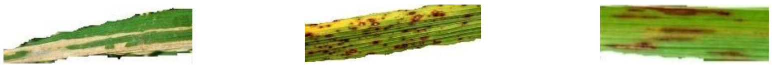
Figure 2. Block Diagram for Proposed Work



Data Collection

The dataset on rice leaf diseases was downloaded from UCI Machine Learning Repository with three (3) classes labelled bacterial leaf blight, brown spot, and leaf smut containing 120 images, as shown in Figure 3.

Figure 3. Classes of Rice Leaf Diseases



(a) Bacterial Leaf Blight (b) Brown Spot (c) Leaf Smut

Data Pre-processing

The data pre-processing of rice leaf diseases with 120 images was applied with a geometric transformation-based technique. This method includes altering, translating, rotating, and scaling an image, which changes the quality and size of an image dataset.

Feature Extraction

The feature extraction approach plays a vital role in agriculture for image classification and extracting relevant information from input images, such as shape, colour, and texture.

Splitting Data into Training and Validation Sets

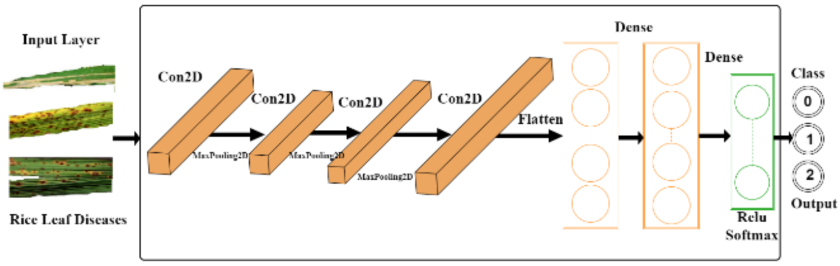
This section's dataset comprises 120 images of rice leaf diseases labelled as bacterial leaf blight, brown spot, and leaf smut, divided into training and validation sets.

Hybrid Deep CNN Architecture

Figure 4 shows a Hybrid Deep "Convolutional Neural Network (CNN)" model built from a typical CNN architecture concept containing 120 rice leaf images with three classes of diseases. A dataset on rice leaf diseases was trained using four convolutional layers, three max-pooling layers, and two dense layers using activation functions such as *read* and *softmax*.

Thus, a new model obtained an accuracy of correctly classified after all data training and validation based on the number of epochs executed.

Figure 4. Hybrid Deep CNN with Transfer Learning Architecture



Data Validation (Performance Metrics)

To validate the result of the dataset, a new model was evaluated based on the following performance metrics of the Confusion Matrix:

- *Accuracy.* "Using the evaluation metrics assessed the classification models to measure how many items are correctly classified instances, incorrectly classified instances, Kappa statistic, mean absolute error, root mean squared error, relative absolute error, and root relative squared error". The formula for accuracy is:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (1)$$

- *TP-Rate. "*Rate of true positives (instances correctly classified as a given class)".

$$\text{TPR (sensitivity)} = \frac{\text{TP}}{\text{TP} + \text{FN}} \qquad (2)$$

- *FP-Rate.* "Rate of false positives (instances falsely classified as a given class). It is known as the false positive rate (FPR)" and is calculated as follows:

$$\text{FPR (1-specificity)} = \frac{\text{FP}}{\text{TN} + \text{FP}} \qquad (3)$$

- *Precision.* "The proportion of actual positive instances of a class is divided by the total instances classified as that class of rice leaf disease".

$$Precision = \frac{TP}{TP + FP} \qquad (4)$$

- *Recall.* "This rate measures the prediction accuracy of TP over the actual positive instances in the dataset".

$$Recall = \frac{TP}{TP + FN} \qquad (5)$$

- *F1-Score.* "This rate measures the accuracy of the prediction or its harmonic average combined with precision" and recall calculated as:

$$F1_{score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

Experimental Results and Analysis

"The results prove the validity of the proposed model for detecting and classifying rice leaves with bacterial leaf blight, brown spot, and leaf smut diseases using a Hybrid Deep Learning CNN-based technique. The experiment was done in Google Colab using Python-based deep learning packages to build a model".

Figure 5 and Figure 6 illustrate the accuracy of the proposed Hybrid CNN in the training and validation sets obtained at 88.58% and 91.67%, respectively, which occurs on the 50^{th} epoch. Also, the training and validation loss was recorded on each epoch of model execution. The performance metrics of the hybrid model were created based on output accuracy and the number of epochs executed, such as 30, 50, and 100. Nevertheless, the results show the efficiency of a new model applied.

Due to the small dataset used in this study, the researcher encountered overfitting during the experiment. The deep-learning CNN model was modified and trained on ImageNet for image classification without severe overfitting for detecting rice bacterial leaf blight, brown spots, and leaf smut diseases.

Fig. 5. Training Loss vs Validation Loss

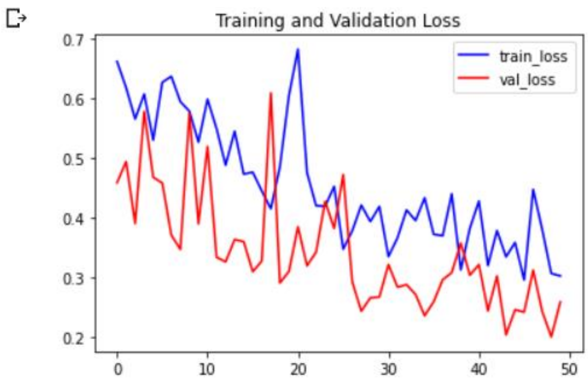
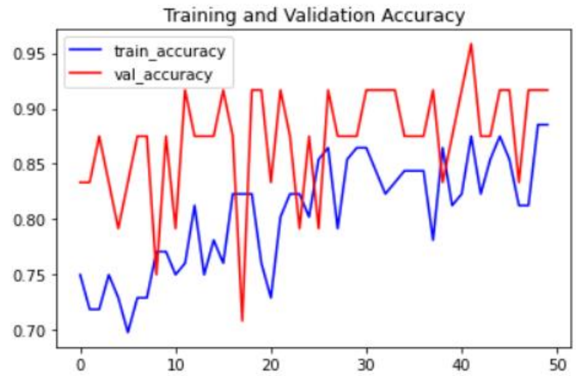


Figure 6. Training Accuracy vs Accuracy



"The accuracy, recall, precision, and F-1 score were used in this study as evaluation measures to analyse the results of a new model. With each class's negative and positive classifications, a confusion matrix is generated to determine the accurate predictions made by the model for identifying rice leaf diseases. The models' performance was evaluated based on the confusion matrix and the classification report, as shown in Figure 7 and Figure 8".

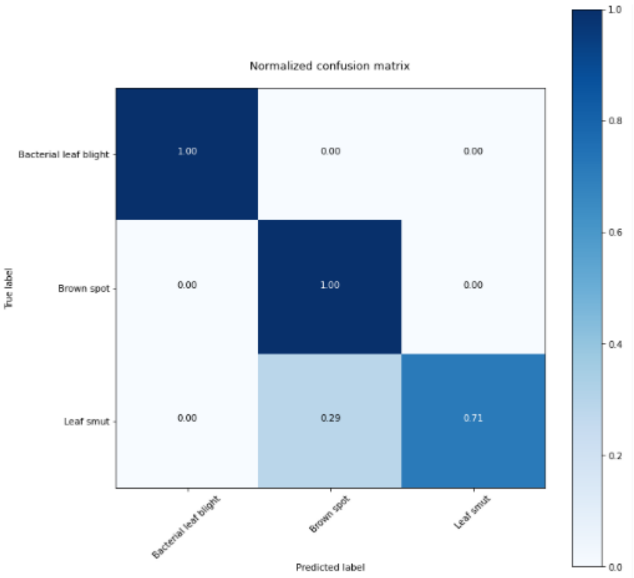
Figure 7. Classification Report Metrics for Rice Leaf Diseases

```
1/1 [==============================] - 0s 362ms/step
              precision    recall  f1-score   support

           0       1.00      1.00      1.00         8
           1       0.82      1.00      0.90         9
           2       1.00      0.71      0.83         7

    accuracy                           0.92        24
   macro avg       0.94      0.90      0.91        24
weighted avg       0.93      0.92      0.91        24
```

Figure 8. Confusion Matrix for Rice Leaf Diseases



CONCLUSIONS

With the rapid development of image processing and deep learning techniques, many researchers studied the image identification and classification of certain plant diseases. “This paper proposed a deep hybrid learning CNN with transfer learning for detecting and classifying rice leaf diseases, namely bacterial leaf blight, brown spot, and leaf smut, with 120 images. A new model achieved an accuracy of 88.58% in the training set and 91.67% in the validation set”. Its performance evaluation metrics were analysed and evaluated based on the confusion matrix and classification report, such as recall, precision, and F-1 score.

For future work, other pre-trained models of CNN, such as VGG19, VGG16, AlexNet, MobileNet, Inception, EfficientNet, and ResNet, will be used for this dataset or the own created dataset. This paper will be the basis for the researchers to conduct the second phase of this research study.

Bibliography

1. Alzubaidi, L., Jinglan Zhang, A. J. Humaidi, Ayad Al-Dujaili, Ye Duan, Omran Al-Shamma, J. Santamaría, Mohammed A. Fadhel, Muthana Al-Amidie, & Laith Farhan (2021). "Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions." *Journal of Big Data 2021 8:1* 8 (1): 1–74. https://doi.org/10.1186/S40537-021-00444-8.
2. Chhillar, A., Sanjeev Thakur, & Ajay Rana. (2020). "Survey of Plant Disease Detection Using Image Classification Techniques." *ICRITO 2020 - IEEE 8th*

International Conference on Reliability, Infocom Technologies and Optimisation (Trends and Future Directions), June 1339–44. https://doi.org/10.1109/ICRITO48877.2020.9197933.

3. Dhillon, Anamika, and Gyanendra K. Verma. 2020. "Convolutional Neural Network: A Review of Models, Methodologies and Applications to Object Detection." *Progress in Artificial Intelligence* 9 (2): 85–112. https://doi.org/10.1007/S13748-019-00203-0/METRICS.
4. Farhood, H., Ivan Bakhshayeshi, M. P., Nabi Rezvani, & Amin Beheshti. (2022). "Recent Advances of Image Processing Techniques in Agriculture." *Artificial Intelligence and Data Science in Environmental Sensing*, January, pp. 129–53. https://doi.org/10.1016/B978-0-323-90508-4.00007-1.
5. Ghosal, S., & Kamal Sarkar (2020). "Rice Leaf Diseases Classification Using CNN with Transfer Learning." *2020 IEEE Calcutta Conference, CALCON 2020 - Proceedings*, February, pp. 230–36. https://doi.org/10.1109/CALCON49167.2020.9106423.
6. Khan, A., Anabia Sohail, U. Z., & Aqsa Saeed Qureshi. (2020). "A Survey of the Recent Architectures of Deep Convolutional Neural Networks." *Artificial Intelligence Review 2020 53:8* 53 (8): 5455–5516. https://doi.org/10.1007/S10462-020-09825-6.
7. N, Krishnamoorthy, L. V. Narasimha Prasad, C. S. Pavan Kumar, Bharat Subedi, Haftom Baraki Abraha, & V. E. Sathishkumar. (2021). "Rice Leaf Diseases Prediction Using Deep Neural Networks with Transfer Learning." *Environmental Research* 198 (July): 111275. https://doi.org/10.1016/J.ENVRES.2021.111275.
8. Sharma, Rishabh, Vinay Kukreja, and Virender Kadyan. 2021. "Rice Diseases Detection Using Convolutional Neural Networks: A Survey." *2021 International Conference on Advance Computing and Innovative Technologies in Engineering, ICACITE 2021*, March, 995–1001. https://doi.org/10.1109/ICACITE51222.2021.9404620.
9. Tunio, M. H., Li Jianping, M. H. F. B., & Imran Memon (2021). "Identification and Classification of Rice Plant Disease Using Hybrid Transfer Learning." *2021 18th International Computer Conference on Wavelet Active Media Technology and Information Processing, ICCWAMTIP 2021*, pp. 525–29. https://doi.org/10.1109/ICCWAMTIP53232.2021.9674124.
10. Yashwanth, M., Muramalla Lahari Chandra, K. P., Danish Showkat, & P. Satish Kumar (2020). "Agriculture Automation Using Deep Learning Methods Implemented Using Keras." *2020 IEEE International Conference for Innovation in Technology, INOCON 2020*, November. https://doi.org/10.1109/INOCON50539.2020.9298415.