A LEAF-BASED RICE DISEASE RECOGNITION SYSTEM USING CONVOLUTIONALNEURAL NETWORK

[1]*Dominic O. Cagadas, [2]Mongkol Ekpanyapong, [3]Christine Marie J. Madrid
University of Science and Technology of Southern Philippines Cagayan de Oro City, 9000, Philippines*
Correspondence Email: dominic.cagadas@ustp.edu.ph

ABSTRACT- *Timely detection of diseases and infection of rice plants would help farmers especially those living in remote areas in treating rice plants and hopefully can increase yield in agricultural production. Developments in a deep convolutional neural network have become the state-of-the-art solution for visual recognition. As CNN has successfully waved in image classification problems, this study detects and recognizes the diseases or infections from the leaves of rice plants. Diseases under test are Bacterial Leaf Blight caused by Bacterial Infection, Rice Blast from Fungal Infection, Rice Tungro Disease from Viral Infection, and Healthy Rice. Pretrained CNN models, Naïve Models,Resnet50, VGG16, VGG19 and Inception V3 are used as feature extractors and classifiers. Experimental results show that amongst all models used for classification, the VGG19 model has achieved an accuracy of 91.0% under fewer parameters and less time for training which makes the system novel, robust and efficient.*

Keywords: rice disease recognition, CNN models, leaf-based rice infection, rice infection

1. INTRODUCTION

Rice is a food staple for more than 3.5 billion people around the world, particularly in Asia, Latin America, and parts of Africa. Rice has been cultivated in Asia for thousands of years [1]. It feeds more people compared to any other crops over a long time. It is extravagantly diverse in the way that rice is grown and consumed by humans, where it can grow in both wet and dry environments compared to other crops that cannot survive, which also makes it unique

However, diseases in crops are inevitable. Farmers have struggled with pesticides, parasites, and plant diseases for years, their primary concerns on the farm. When a rice plant gets sick, the overall harvest for the season suffers, leaving farmers disappointed [2].

Damages in rice caused by its diseases can greatly reduce production or yield. These diseases are mainly caused by fungal, bacterial, and/or viral infections. Planting rice varieties that can resist such infections is the simplest and the most cost-effective management for diseases [3].

In [3], It is estimated that every year, farmers lose an average of 37% of their crops due to pests and diseases. To significantly reduce losses, a timely and accurate diagnosis of diseases is needed on top of good crop management. Crop problems can be caused by both living and non-living factors including but not limited to fungus and rats, soil acidity, water, wind, temperature, and even radiation.

Recently, deep learning has been thoroughly used in several visual recognition problems. Given its success, it has become the new state-of-the-art solution for distinct tasks of different domains. It has grown quickly as it has the ability to effectively encode spectral and spatial information from its own data. Its methods have achieved state-of-the-art results in different applications, such as remote sensing, image classification, urban planning, mapping, and many more [4].

Convolutional neural network (CNN), is a class of artificial neural networks that have become dominant in various computer vision tasks specifically in image classification [5]. CNN is similar to other neural networks, but they have an added complexity due to the fact that they use a series of convolutional layers such as convolutions, max-pooling, fully-connected layers, and activation functions [6]. Alex Net, VGGs, ResNet50, InceptionV3, and Google Net are the very popular CNN architectures used as subroutines in obtaining representations to solve different tasks.

In [3], timely and accurate recognition of rice plant diseases can greatly help to increase yield and reduce economic losses significantly. It is of great assistance to farmers in applying timely treatment to infections. As such, the developments of deep learning-based CNN, allowed researchers to potentially improve the accuracy of image classification problems.

In this study, we incorporate the rice disease classification into our different CNN models and observe its responses in terms of accuracy in disease detection. According to [7], different rice disease existing studies have high accuracy but have many parameters hence giving more time in training, difficulty in getting results and eating a lot of the device memory. Hence, the proposed system leads to knowing the best model fit for detecting and or classifying rice diseases for better efficiency and accuracy with fewer parameters.

The main objective of this study is to modify a CNN model system with higher accuracy that classifies and detects rice diseases based on its leaves and compare the results of recognition of the different CNN models such as VGG16, VGG19, Google net/InceptionV3, Resnet50, and some modified models. Leaf classification of plants is a useful step for plant health prediction and yield estimation [8].

Conceptual Framework

Table 1 shows the conceptual framework of the study. For the inputs of the study, leaf-based rice disease images were collected from the internet and the healthy rice images were collected from a bare rice field. For the processes of the study, a Naïve CNN model and four pre-trained CNN models were used separately for each system namely Resnet50, VGG16, VGG19, and InceptionV3. The images are fed to the CNN modelsystem for training and validation. After creating a successful model, it will be tested for correct classifications and accuracy.

Table 1. Conceptual Framework of the Study

Class	Precision	Recall	F1-score	Support
Class 0 (Healthy_Rice)	1.00	0.86	0.93	81
Class 1 BacterialBlight)	0.81	0.58	0.67	66
Class 2 (Rice_Blast)	0.69	0.81	0.75	91
Class 3 (Rice_Tungro)	0.78	0.91	0.84	82
Average	**0.82**	**0.79**	**0.80**	**320**

Table 2. Classification Report of CNN Naïve Model

Class	Precision	Recall	F1-score	Support
Class 0 (Healthy_Rice)	0.78	0.89	0.83	92
Class 1 (bacterial blight)	0.72	0.78	0.75	109
Class 2 (Rice_Blast)	0.89	0.61	0.72	107
Class 3 (Rice_Tungro)	0.77	0.87	0.82	92
Average	**0.79**	**0.79**	**0.78**	**400**

Figure 1 presents the graphs of the model accuracy and model loss of the CNN Naïve Model at 50 iterations. It is evident that at 30 iterations, the lines are almost stable or linear which might indicate that the model is robust andefficient at that moment.

Table 3. Classification Report of CNN ResNet50 Model

INPUT	PROCESS	OUTPUT
Rice Disease Data Images 1. Bacterial Blight 2. Rice Blast 3. Tungro Disease 4. HealthyRice	**CNN Models** Classifiers	**Results of Classification** Test of Accuracy

2. METHODOLOGY

The overall study focuses on the training, validating, and testing of the gathered leaf-based rice disease images dataset. Most of the rice leaves diseases were collected online while the healthy rice leaves were collected in a bare field. Images were cropped and reshaped into desired image sizes. After collection, images were classified according to their disease classes. CNN modelswere then used for feature extraction and classification ofthe images.

2.2 System Design

The overall system design starts with dataset gathering of around 2000 images divided into four classes, images arethen resized as per model requirement, and trained for different procedures, validation then follows and ends with a testing procedure for each model implementation. All models are executed in the Keras library using the TensorFlow backend. Outputs are visualized using Sckit-learning library.

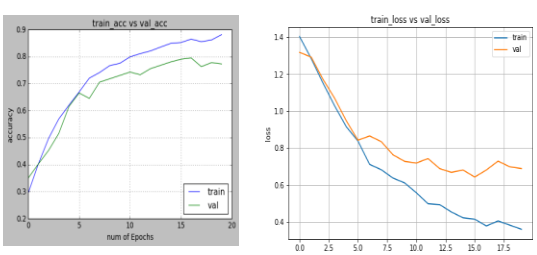

Figure 1. Accuracy and loss of Naïve Model

3. RESULTS AND DISCUSSION

3.1 CNN Naïve Model

The test images were done in 50 iterations where the overall accuracy is found to be 78%. Classification reports are also presented here. Model accuracy and model loss accuracy are presented between train and validation. Table 2 presents the classification report of the CNN Naïve model with four classes. It is evident in the resultsthat the average accuracy of the model is 78.0% across four classes. From the other reports, Class 0 Healthy Rice has the highest Recall with a percentage of 89.0% identified to be true positive healthy rice while the rest were false negative (other diseases).

CNN ResNet50 Model

The test images were done in 12 iterations where the overall accuracy is found to be 80.31%. The classification report and confusion matrix are also presented here. Model accuracy and model loss accuracy are presented between train and validation.

Table 3 presents the classification report of the CNN Resnet50 model with four classes. It is evident in the results that the average accuracy of the model is 80.31% across four classes. From the other reports, Class 3 Rice Tungro Disease has the highest recall with a percentage of 91% amongst others, meaning, among 82 samples, 75 samples were identified to be true positive (rice_tungro disease) while the rest were false negative (other diseases).

Figure 2 presents the graphs of the model accuracy and model loss of the CNN Resnet50 Model at 12 iterations. It is evident from the graphs obtained that the lines are stable or linear which might indicate that the model is robust and efficient at that moment.

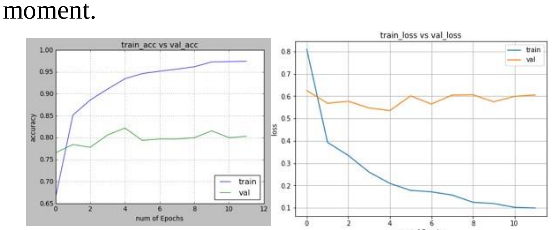

Figure 2. Accuracy and loss of ResNet50 Model

3.2 CNN VGG16 Model

The test images were done in 12 iterations where the overall accuracy is found to be 90.00%. The classification report and confusion matrix are also presented here. Model accuracy and model loss accuracy are presented between train and validation.

Table 4 presents the classification report of the CNN VGG16 model with four classes. It is evident in the results that the average accuracy of the model is 90.00% across four classes. From the other reports, Class 0 healthy rice has the highest Recall with a percentage of 96% amongst others, meaning, among 81 samples, 78 samples were identified to be true positive (healthy rice) while the rest were false negative (other diseases).

Table 4. Classification Report of CNN VGG16 Model

Class	Precision	Recall	F1-score	Support
Class 0 (Healthy_Rice)	1.00	0.96	0.98	81
Class 1 (bacterial blight)	0.79	0.86	0.83	66
Class 2 (Rice_Blast)	0.89	0.85	0.87	91
Class 3 (Rice_Tungro)	0.93	0.94	0.93	82
Average	**0.90**	**0.90**	**0.90**	**320**

Figure 3 presents the graphs of the model accuracy and model loss of the CNN VGG16 Model at 12 iterations. It is evident from the graph that at n iterations, the lines are almost stable or linear which indicates that the model is robust and efficient at that moment.

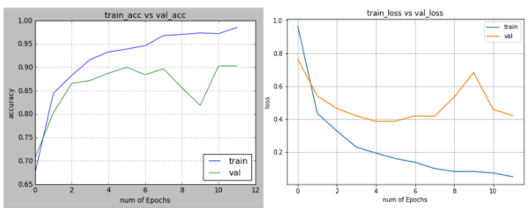

Figure 3. Accuracy and loss of VGG16 Model.

3.3 CNN VGG19 Model

The test images were done in 12 iterations where the overall accuracy is found to be 91.0 %. The classification report and confusion matrix are also presented here. Model accuracy and model loss accuracy are presented between train and validation.

Table 5 presents the classification report of the CNN VGG19 model with four classes. It is evident in the results that the average accuracy of the model is 91.00% across four classes. From the other reports, Class 0 Healthy Rice has the highest Recall with a percentage of 96% amongst others, meaning, among 81 samples, 78 samples were identified to be true positive (healthy rice) while the rest were false negative (other diseases).

Table 5. Classification Report of CNN VGG19 Model

Class	Precision	Recall	F1-score	Support
Class 0 (Healthy_Rice)	0.97	0.96	0.97	81
Class 1 (bacterial blight)	0.84	0.86	0.85	66
Class 2 (Rice_Blast)	0.90	0.88	0.89	91
Class 3 (Rice_Tungro)	0.92	0.93	0.94	82
Average	**0.91**	**0.91**	**0.91**	**320**

Figure 4 presents the graphs of the model accuracy and model loss of the CNN VGG19 Model at 12 iterations. It is evident from the graph that at n iterations, the lines are almost stable or linear which indicates that the model is robust and efficient at that moment.

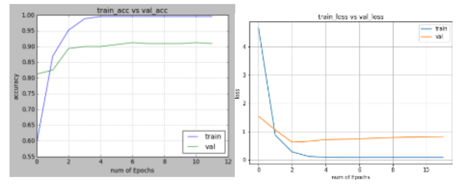

Figure 4. Accuracy and loss of VGG19 Model

3.4 CNN Inception V3 Model

The test images were done in 20 iterations where the overall accuracy is found to be 85.00%. The classification report and confusion matrix are also presented here. Model accuracy and model loss accuracy are presented between train and validation.

Table 6 presents the classification report of the CNN VGG19 model with four classes. It is evident in the results that the average accuracy of the model is 85.00% across four classes. From the other reports, Class 1 Rice Blast disease has the highest Recall with a percentage of 98% amongst others, meaning, among 66 samples, 61 samples were identified to be true positive (healthy rice) while the rest were false negative (other diseases).

Table 6. Classification Report of CNN VGG19 Model

Class	Precision	Recall	F1-score	Support
Class 0 (Healthy_Rice)	1.00	0.94	0.97	81
Class 1 (BacterialBlight)	0.63	0.98	0.77	66
Class 2 (Rice_Blast)	0.93	0.68	0.78	91
Class 3 (Rice_Tungro)	0.95	0.85	0.90	82
Average	**0.88**	**0.86**	**0.85**	**320**

Figure 5 presents the graphs of the model accuracy and model loss of the CNN InceptionV3 Model at 20 iterations. It is evident from the graphs obtained that the lines are not stable or not linear which might indicate that the model is partially robust and efficient at that moment.

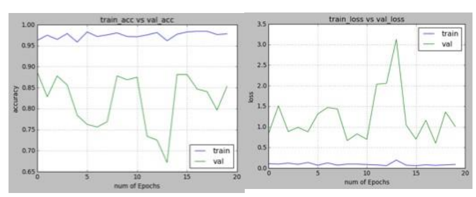

Figure 5. Accuracy and loss of Inception V3 Model

3.5 Summary of CNN Models

Table 7 presents the summary of all the CNN Models used in this study at different parameters. From the results presented, it is clearly evident that the best model fit for this research study in classifying the rice diseases based on leaves, after fine-tuning all models is CNN VGG19 Model at 91.00% accuracy.

Table 7. Summary of Results of CNN Models

Model	No. of Parameters	Testing Time (s)	Accuraby
Naïve CNNModel	831,2014	------	79.00 %
ResNet 50	8,196	2,946	80.31 %
VGG-16	16,388	7,474	90.00 %
VGG-19	3,228,420	9,665	91.00 %
InceptionV3	23,082,148	30,092	85.00 %

4. CONCLUSION AND RECOMMENDATIONS

The study has successfully implemented the leaf-based rice diseases recognition system using a convolutional neural network in four classes namely: Rice Blast Disease for Fungal Infection, Bacterial Leaf Blight for Bacterial Infection, Rice Tungro Diseases for Viral Infection and Healthy Rice. Different CNN Models were used to compare efficient and accurate results. As per the accuracyresults of the different models, it is concluded that the optimal model for classifying rice diseases under thisstudy is VGG19 with an accuracy of 91.0%. Also, it isevident from the results that the VGG19 model gave fewer parameters compared to other models as well as itstraining time.

It is recommended in this study to collect more data images of leaves contaminated with the different rice diseases to achieve higher accuracy. Moreover, fine-tuning of other models is also encouraged to lessen the parameters for training.

For commercial purposes, it is recommended to develop the system for mobile application as it will be the most accessible way of technology for the farmers especially those in remote places. Such technology will help them in the timely detection of rice diseases contaminated in their farms which would allow them to increase yield in farms by applying the right cure for the rice infection.

5. REFERENCES

[1] National Geographic, "Food Staple". [Online]. Available: https://education.nationalgeographic.org/resource/food-staple/

[2] G. Gamboa, et al. (2023) " Ricefield Health Monitoring System Using a Drone with AI Interface", Science International (Lahore) [Online]. 35(3), 181-184. Available: https://www.researchgate.net/profile/Analyn-Morite/publication/370979511_RICEFIELD_HEALTH_MONITORING_SYSTEM_USING_A_DRONE_WITH_AI_INTERFACE/links/646d627dcde97a39292b696c/RICEFIELD-HEALTH-MONITORING-SYSTEM-USING-A-DRONE-WITH-AI-INTERFACE.pdf

[3] International Rice Research Institute, Rice Knowledge Bank, " How To Manage Pest and Disease", [Online]. Available: http://www.knowledgebank.irri.org/step-by-step-production/growth/pests-and-diseases

[4] R. F. Nogueira, R. de Alencar Lotufo and R. Campos Machado. (2016, June) "Fingerprint Liveness Detection Using Convolutional Neural Networks", IEEE Transactions on Information Forensics and Security [Online]. 11(6), 1206-1213. Available: https://ieeexplore.ieee.org/document/7390065.

[5] Yamashita, R., Nishio, M., Do, R. K., & Togashi, K. (2018). "Convolutional neural networks: An overview and application in radiology", Insights into Imaging [Online]. *9*(4), 611-629. Available: https://doi.org/10.1007/s13244-018-0639-9

[6] Data Analytics, "Different Types of CNN Architectures Explained: Example" [Online]. Available: https://vitalflux.com/different-types-of-cnn-architectures-explained-examples/

[7] Chen, Y., Emer, J. & Sze, V. (2016, June). "Eyeriss: A Spatial Architecture for Energy- Efficient Dataflow for Convolutional Neural Networks", Association for Computing Machinery [Abstract]. 44(3),367-379. Available: https://dl.acm.org/doi/abs/10.1145/3007787.3001177

[8] N. Ahmed, U. Khan & S. Asif (2016). "An Automatic Leaf Based Plant Identification System", Science InternationalLahore [Online]. 28(1),427-430. Available:https://www.researchgate.net/profile/Nisar-Ahmed-31/publication/299183724_An_Automatic_Leaf_Based_Plant_Identification_System/links/56ef96b608aed17d0 9f87b5b/An-Automatic-Leaf-Based-Plant-Identification-System.pdf