

Automatic Diagnosis of Rice Leaves Diseases Using Hybrid Deep Learning Model

Amjad Rehman Khan^{1,*}, Ibrahim Abunadi¹, Bayan AlGhofaily¹, Haider Ali², and Tanzila Saba¹

¹ Artificial Intelligence & Data Analytics Lab (AIDA), CCIS Prince Sultan University, Riyadh 11586, Saudi Arabia; Email: iabunadi@psu.edu.sa (I.A.); bghofaily@psu.edu.sa (B.A.); drstanzila@gmail.com (T.S.)

² Department of Statistics, University of Gujrat, Gujrat 50700, Pakistan; Email: mirzahaiderali51@gmail.com (H.A.)

*Correspondence: arkhan@psu.edu.sa (A.R.K.)

Abstract—Rice demand is increasing with the rise in population worldwide, but this crop production is negatively affected due to different fatal diseases. Reported rice disease diagnosis models are imprecise, inefficient, and Taylor made. Hence, this research presents an efficient hybrid model of different rice disease diagnoses to support the agricultural industry's economic growth. The proposed hybrid model is composed of Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and Self-Attention (SA) modules. The fitness level of the proposed model is evaluated using a test dataset, 5-fold cross-validation (CV), Hosmer-Lemeshow test, Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). Four rice leaf diseases (Bacterial blight, Blast, Tungro, Brown spot) are diagnosed from the benchmark dataset. From the 5-fold CV metric, the proposed model attained 100% with a 0.001 average accuracy and loss for training samples. Similarly, got 97.51% with a 0.110 average accuracy and loss for validation samples. The proposed model also achieved the higher Receiver Operating Curve (ROC) with Area Under Curve (AUC) rate, precision, recall, and F1-score. The model also obtained minimum RMSE, MAPE and Hosmer Lemeshow test values, revealing that the proposed model is well-fitted. The proposed model also got 100% accuracy, precision, recall, F1-score as well as for testing samples. The performance metrics exhibited that the proposed model's overall performance was perfect and could be used in agriculture for disease identification of rice leaves. The present investigation achieved a high diagnosis rate for rice leaf disease identification without over and under-fitting issues. The model has a 97.5%–100% confidence interval for detecting rice leaf disease. Finally, the proposed model will support the agriculture industry in diagnosing rice leaf diseases and monitoring their growth.

Keywords—Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), image processing, agriculture, economic growth, local investment

I. INTRODUCTION

Rice is a major crop in the agriculture industry. Though, agricultural diseases may lower yield and quality, endangering global food supplies. These diseases threaten global food production and availability [1, 2]. For instance,

millions of people worldwide lack enough food and plant pests waste nearly 12% of the global food supply. Compared to past metric tons, rice consumption has increased [3]. However, poor agricultural monitoring caused disease-related rice diseases [4]. Rice diseases are the leading cause of large economic losses. To fight plant diseases, bactericides, fungicides, and nematicides are widely used, harming the agriculture industry. Currently, rice disease diagnosis is simple but uses traditional approaches that misidentify related diseases and harm rice growth [5]. Manually detecting rice diseases is unreliable, expensive, and time-consuming. Due to environmental impact, slow detection speed, and poor accuracy, computer-based identification approaches are seldom employed. Thus, precise automated systems for rice disease diagnosis are in crucial demand [6]. To enhance rice crop production, several studies have used computer vision technologies to control crop diseases and proposed traditional/ deep learning models in the literature [7, 8].

Numerous obstacles have hindered the development of a system for automated diagnosing rice ailments. It is emphasized that both diagnosis and detection entail procedures that might make it exceedingly difficult to accurately segment the region of the rice plant where the symptoms emerge. The capture circumstances are difficult to manage, which may make it more challenging to forecast images and identify the disease [9]. Moreover, the symptoms produced by several diseases may seem similar, and methods of differentiation may be based on minute differences. Inconsistencies in the distribution of data features used to train and validate the model are common issues that pose overfitting issues [10]. This is important for automatically diagnosing plant diseases since symptoms may vary by region and systems might face overfitting issues that reduce accuracy considerably. A number of the recommended designs for diagnosing rice leaf diseases have been discovered to be offline, while few are observed in real-time [11, 12]. However, image resolution increased in real-time, which also increased computing complexity. In addition, the complexity of real-time operations grows with a wide range of diseases and their vague symptoms. The current research presents a

hybrid deep learning model with a self-attention module with the following main contributions to address these obstacles.

- a) The proposed system combines deep learning models such as Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) with the Self-Attention (SA) model based on disease identification on rice leaves.
- b) The proposed CNN+LSTM+SA model for rice leaves disease diagnosis using Mendeley data reached a 100% accuracy rate for training and 97.51% for validation samples, 100% for testing samples.
- c) The proposed model can identify rice leaves disease with a 97.51%–100% confidence interval without the problem of under and overfitting. Therefore, it will be helpful in the agriculture sector to enable early identification of rice leaf diseases.

II. LITERATURE REVIEW

Currently, researchers have estimated agricultural yields using computer vision for decades. Image processing, pattern recognition, machine and deep learning models have been utilized to diagnose agricultural diseases using computer vision. SVMs identified bacterial rice leaf, sheath, and blast shape and texture. A genetic algorithm and support vector machine recognized damaged crop leaves [13]. Islam *et al.* [14] employed Naive Bayes to categorize brown rice spot, bacterial blight, and blasts using RGB values of afflicted areas. Infrared thermal imaging could detect tomato mosaic disease and wheat leaf rust. The diagnosis rate for Blast disease was above 89%, and Bacterial, Blight, Brown Spot were above 90% reported. Most crop disease diagnostics need manual feature extraction. Thus, expressiveness is confined and consequences are hard to generalize. Some methods need expensive equipment. Limitations make crop disease diagnosis difficult [15, 16].

Deep learning might enhance agricultural disease diagnosis. Deep learning is used to detect crop diseases thru image classification, object identification and spray recommendation. Lu *et al.* [17] proposed an in-field deep learning based CNN model to identify ten common rice diseases. They reported 95.48% accuracy on a dataset of 500 images. Islam *et al.* [18] proposed a modified CNN model to diagnose Paddy leaf disease from 984 images and reported 92.68% accuracy using InceptionResNet-V2. Zguven and Adem [19] identified sugar beet leaf spot disease using Faster R-CNN and reported 95.48% recognition accuracy for using 155 images test dataset. Chen *et al.* [20] enhanced VGGNet with an inception module with migration learning and diagnosed correct rice diseases 92%. Picon *et al.* [21] extended the Deep Residual Neural Network-based algorithm to diagnose multiple plant diseases and claimed 96% accuracy on a polit test conducted in Germany. Barbedo [22] investigated 14 plants and 79 conditions using a lesions-and-spots-based image classification approach that was more accurate than original images. This method claimed

accuracies 12% better than the authentic images. However, no crop had an accuracy below 75%, even with 10 plagues. These findings show that deep learning could identify and recognize plant diseases with adequate data, even when the database does not cover all realistic possibilities. Karlekar and Seal [23] recommended SoyNet for soybean leaf image disease detection. Tomato, cassava, tulip, and millet disease detection also uses deep learning. Their model has two modules. By eliminating complicated backgrounds, the first module separates leaf components from complete images. The second module presents SoyNet, a deep-learning CNN for soybean plant disease identification using segmented leaf images. They claimed 98.14% identification accuracy.

Rahman *et al.* [24] recently presented a CNN model to detect rice plant diseases and pests using camera images. Four cameras were used to gather 1426 ideas and reported 94.33% mean validation accuracy. Their model used two-stage training to reduce model size while retaining high classification precision dramatically. When stacked CNN was used instead of VGG16, test accuracy increased by 95% while model size lowered by 98%. Hence, an outstanding model free from over and under-fitting issues is needed to precisely diagnose rice leave diseases [25, 26].

Overall, deep learning can accurately diagnose diseases [27]. However, deep learning research on rice diseases focused only on a few conditions. For example, rice leaf blast, fake smut, neck blast, sheath blight, bacterial stripe disease, and brown spot are among the rice diseases seen in fields [28]. Furthermore, it is observed from the evaluation of reported techniques that one prediction model may overfit machine learning algorithms. Therefore, it is concluded that ensemble learning is the possible solution, and so it was used to predict disease diagnosis [29, 30].

Further, this paper has three main sections: Section III presents the proposed hybrid model, Section IV exhibits experimental results, detailed analysis and comparisons. Finally, Section V concludes the research.

III. MATERIALS AND METHODS

CNN has exhibited excellent performance pattern recognition applications in agriculture, such as pests' recognition, disease classification, pests spray needs and forecasting, etc. [31–33]. The LSTM model has various applications, such as image classification, speech recognition, machine translation, language modelling, stock prediction, etc. [34, 35]. The self-Attention is widely used for text classification, image data recognition, and classification [36]. Its extended architecture produces good results in NLP (Natural Language Processing) tasks than CNN and Recurrent neural network (RNN) [37, 38]. All three models have good properties of image recognition and classification.

This paper proposed hybrid model composed of CNN, LSTM, and SA to identify the rice leave diseases using their image data. The rectified linear unit (ReLU) function is employed in CNN, LSTM, and SA layers. The ReLU function is less expensive and faster than other functions. It is also considered an efficient function that is always

leveled as 0 and 1 [39]. The SoftMax activation function was used in the output layer due to multi-class classification. The SoftMax can be used for multi-class type in the dense layer / output layer [40]. The architecture of the hybrid model is exhibited in Fig. 1.

The model consists of three convolutional layers, three max-pooling layers, two dense layers, two LSTM layers, and one SA layer. The grey area in the diagram presents the convolutional layers, the brown site presents the functions such as Relu and softmax, blue colour presents the max-pooling layers, the green color with dense layers, the red colour with LSTM layer and the yellow color presents the self-attention layer. The 32 (3 × 3), 64 (3 × 3), and 64 (3 × 3) were the shapes of three convolutional layers, respectively. The (2 × 2) kernel size is used for all max pooling layers. The two layers of LSTM contain 32 filter size.

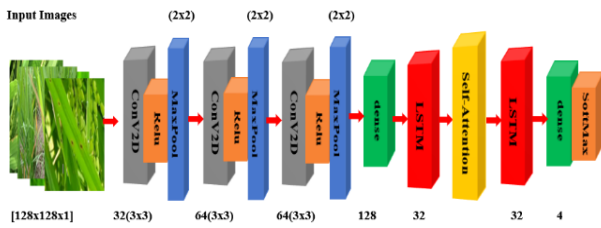


Figure 1. Architecture of hybrid model.

The raw images were preprocessed using OpenCV in python [40]. First raw images are imported in Python, and finally, through OpenCV, we normalised, smoothed and reshaped the rice leave ideas for disease classification.

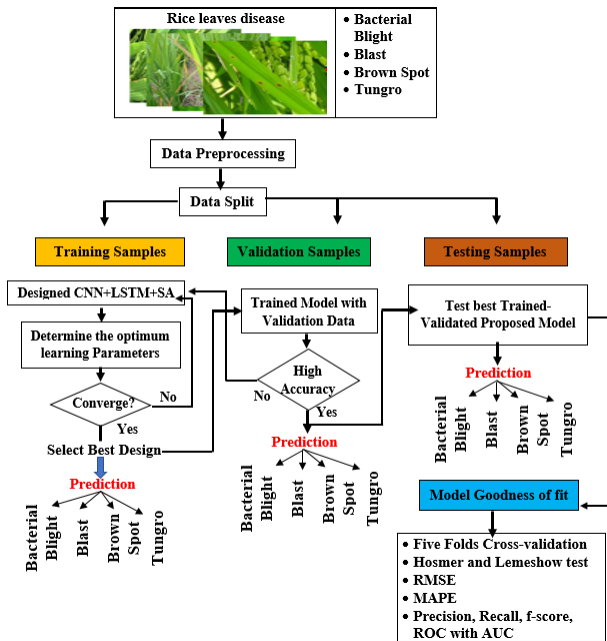


Figure 2. Proposed research framework.

The dataset comprises 5932 images divided into train, validation and test datasets. 70% (4152) images were used as training samples and 20% (1602) images were used as validation samples and 10% (178) were used as testing

samples or holdout sample groups. The 5-fold Cross-Validation (CV) approach was also applied to the proposed model to measure whether any overfitting exists. The other benefit of the CV approach is to check the predictive accuracy and to minimize the bias of train/ test datasets. The detailed framework of the proposed methodology is exhibited in Fig. 2. First, the accuracy, loss, precision, recall, F1-score, RMSE, MAPE, Hosmer and Lemeshow goodness of fit test was calculated to observe the depth results of the proposed model.

IV. EXPERIMENTAL RESULTS

A. Dataset

The images of rice leaf diseases employed in experiments were downloaded from the publicly available Mendeley dataset [41]. The dataset composed of 5932 images of rice leaf diseases (Bacterial Blight, Blast, Brown Spot, Tungro) exhibited in Fig. 3. Where 1584 (26.70%) samples were Bacterial blight, 1440 (24.25%) samples were Blast disease, 1308 (22.04%) were Tungro disease, and 1600 (26.97%) were Brown spot out of 5932 samples used in experiments.

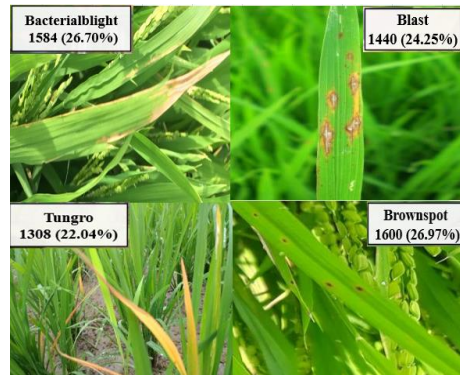


Figure 3. Four types of rice disease.

B. Analysis and Comparisons

1) Results from CNN model

The Table I present the results of CNN model of training and validation samples in the form of Precision, Recall, F1-Score, Support, and Accuracy. The rice types of training samples were 99.27% correctly classified and rice types of validation samples were 95.61% correctly classified. The difference of accuracy of training and validation samples lead the over fitting problem. The LSTM and SA layers were added to make sure the model has no problem of over fitting.

This study employs the proposed model to identify the rice leaf disease using the images dataset. 70% (4152) images were used as training samples and 20% (1602) images were used as validation samples and 10% (178) were used as testing samples or holdout samples group. The performance metrics and results are presented in Table II and Table III.

The proposed model attained 100% precision, recall, and F1-score for training samples for Bacterial blight, Blast, Brown spot and Tungro. The accuracy rate for

training samples obtained from the proposed model was 100%. The 100% accuracy was outstanding, and unexpected results were obtained from the proposed model. The proposed model was trained on a training dataset, and tested on Validation dataset and after trained-validated model tested on holdout sample or testing samples not involved in training. The 100% precision, recall, F1-score obtained for the Bacterial Blight class, Blast class, Brown Spot class, and Tungro class respectively from training samples. The 97% precision, recall, and F1-score attained for the Bacterial Blight and Blast class, 98% precision, recall, F1-score for Brown Spot class and 97% precision, 99% recall, 98% F1-score obtained for Tungro class from validation samples are presented in Table III. A 98% accuracy was obtained for the validation samples with the proposed model and 100% accuracy for training samples. The results of the proposed model regarding training and validation samples were outstanding as compare to CNN model. In Fig. 4, the accurate classified numbers of the rice leave disease have appeared in 4×4 diagonal values, and alternative incorrect classified numbers are present in the off-diagonal in the 4×4 matrix. Out of 4152 samples, 1108 samples of Bacterial blight, 1034 samples of Blast, 1111 of Brown Spot, and 899 Tungro were correctly classified from training samples. Two specimens were shown in off-diagonal, indicating an incorrect classification of the proposed model for training samples.

TABLE I. PERFORMANCE OF CNN MODEL WITHOUT 5-FOLD CV APPROACH

Training Samples					
	Precision	Recall	F1-score	Support	Accuracy (%)
Bacterial Blight	0.97	1.00	0.99	1108	99.27%
Blast	1.00	1.00	1.00	1034	
Brown Spot	1.00	1.00	1.00	1111	
Tungro	1.00	0.97	0.98	899	
Validation Samples					
Bacterial Blight	0.90	0.99	0.94	428	95.61%
Blast	0.95	0.98	0.97	356	
Brown Spot	0.99	0.96	0.98	449	
Tungro	0.99	0.89	0.94	369	

TABLE II. PERFORMANCE METRICS

Accuracy	$\frac{T_p + T_n}{T_p + T_n + F_p + F_n}$
Precision	$\frac{T_p}{T_p + F_p}$
Recall	$\frac{T_p}{T_p + T_n}$
F1-score	$2 \times \frac{\text{Precision} \times \text{recall}}{\text{Precision} + \text{recall}}$
RMSE	$\sqrt{\frac{\sum_{i=0}^n (\text{Actual}_i - \text{Predicted}_i)^2}{n}}$
MAPE	$\frac{1}{n} \sum_{i=0}^n \left \frac{\text{Actual}_i - \text{Predicted}_i}{\text{Predicted}_i} \right \times 100$

TABLE III. PERFORMANCE OF PROPOSED MODEL WITHOUT 5-FOLD CV APPROACH

Training Samples					
	Precision	Recall	F1-score	Support	Accuracy (%)
Bacterial Blight	1.00	1.00	1.00	1108	100%
Blast	1.00	1.00	1.00	1034	
Brown Spot	1.00	1.00	1.00	1111	
Tungro	1.00	1.00	1.00	899	
Validation Samples					
Bacterial Blight	0.97	0.97	0.97	428	98%
Blast	0.97	0.97	0.97	356	
Brown Spot	0.98	0.98	0.98	449	
Tungro	0.97	0.99	0.98	369	

Similarly, out of 1602 samples, 416 samples of Bacterial blight, 345 samples of Blast, 439 of Brown Spot, and 365 samples of Tungro were correctly classified from validation samples. Only 37 examples were shown in off-diagonal, indicating the incorrect classification of the proposed model for validation samples. The confusion matrix also indicates that the proposed model has a wonderful experience with this data. Additionally, we applied a 5-folds CV approach on train and test datasets for more precise verification.

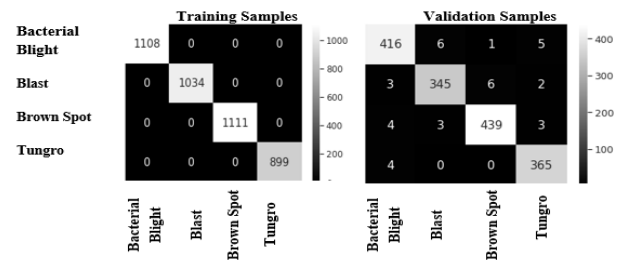


Figure 4. Confusion matrix for train and validation datasets without 5-fold CV approach.

2) Cross-validation metrics

The 5-fold CV approach has been applied to measure performance of proposed model. The graphical results of 5-fold CV approach with accuracies and losses of training and validation samples against 25 epochs are given in Fig. 5. The solid black, red, blue, green, and grey lines present the accuracy and losses of the 5 folds of training samples, respectively. Similarly, the dotted black, red, blue, green, and grey lines present the accuracy and losses of the 5 folds of validation samples, respectively. From 1 to 25 we can observe that the accuracies of each fold increase as epoch increase and loss decrease as epochs increase which is the best part of the proposed model.

By adopting the CV approach 100%, 100%, 100%, 100%, and 100% accuracies were achieved from 5-folds with training samples with the proposed model in Table IV. The 0.001, 0.003, 0.003, 0.001, and 0.002 losses were obtained from 5-folds for training samples. Similarly,

97.40%, 97.62%, 97.58%, 97.52%, and 97.47% accuracies were achieved from 5-folds with validation samples. The 0.102, 0.098, 0.121, 0.132, and 0.101 losses were obtained from 5-folds for validation samples. The average accuracy of the 5-fold CV method for training samples was 100% with a 0.001 loss rate. Similarly, the proposed model got average accuracy 97.51% by 5-fold CV method for validation samples with a 0.110 loss rate.

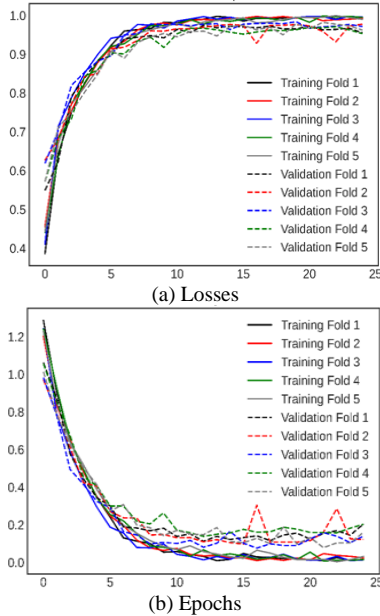


Figure 5. 5-fold CV accuracies and losses against the 25 epochs of hybrid model.

TABLE IV. PRESENTS THE PERFORMANCE OF THE PROPOSED 5-FOLD CV

Training Samples						
	Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	Mean
Accuracy	100%	100%	100%	100%	100%	100%
Loss	0.001	0.003	0.002	0.001	0.002	0.001
Validation Samples						
Accuracy	97.40%	97.62%	97.58%	97.52%	97.47%	97.51%
Loss	0.102	0.098	0.121	0.132	0.101	0.110

The CV approach also revealed that the overall performance of proposed model is good enough and has no overfitting issues. Therefore, highly fruitful for the prediction process. The CV method has also shown high predictive accuracy of the proposed models on random sampling from training and validation datasets. After the 5-fold CV, we also calculated the actual quantity of four classes from training samples, validation samples, train predicted classes, and validation predicted classes. Fig. 6 specifies a graphical depiction of actual and expected classes. It also presents the actual and prediction quantity for both train/test datasets for Bacterial Blight, Blast, Brown Spot, and Tungro for both train/ test datasets. The green, blue, sky blue, brown colors with dotted style denote the actual and predicted quantity of Bacterial Blight, Blast, Brown Spot and Tungro class respectively. The

calculated measures were used for the Hosmer and Lemeshow test of goodness, RMSE, and MAPE.

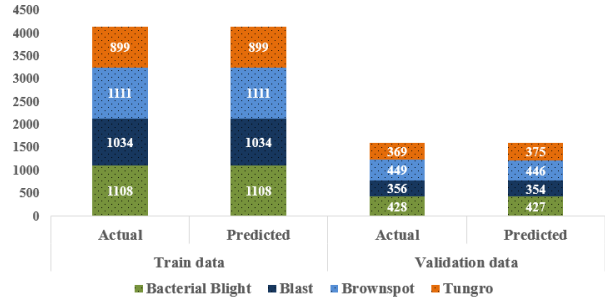


Figure 6. Actual vs prediction quantity accuracy.

C. Measurements of the Goodness Metrics

The Hosmer-Lemeshow test determines the model's goodness of fit from the observed and predicted values using the chi-square goodness of fit test [42]. The null hypothesis of the chi-square test was actual and observed or predicted classes are the same. The test statistics (p -value) measured in chi-square for the acceptance or rejection of the null hypothesis is given in Table V. The 0.0 (1) chi-square test statistics (p -value) were achieved for training samples and 0.1298 (0.9880) chi-square test statistics (p -value) were achieved for validation samples. By accepting the null hypothesis, we conclude that proposed model performance is outstanding.

TABLE V. GOODNESS OF FIT MODEL

	Chi-Square Test-Statistics	p -value	RMSE	MAPE
Training Samples	0.0	1	0.0	0.0
Validation Samples	0.1298	0.9880	3.53	0.007

The significant results of Hosmer-Lemeshow test show that model is well-fitted. The 0.0 and 3.53 values of RMSE were obtained from training and validation samples respectively. Similarly, 0.0 and 0.007 values of MAPE were obtained from training and validation samples, respectively. The lower value of RMSE and MAPE was found for the proposed model. It could be seen that (CNN+LSTM+SA) proposed model is the most successful model for rice leaves disease identification. All criteria conclude that the overall performance of the proposed model was outstanding. The model is free from the problem of overfitting and underfitting and can be utilized for prediction.

D. ROC Curve Analysis

The ROC curve analysis was also performed to evaluate the model's diagnosis accuracy of all four rice leaf diseases. It exhibited a high specificity and sensitivity rate. The ROC curve values within 0.70–0.80 are accepted, higher than 0.80 are excellent, and higher than 0.90 are rarely observed. The ROC curve with AUC based on the proposed model of the leaf's disease is presented in Fig. 7. The four blue, red, green, and black solid lines present the

ROC with AUC of rice leaf disease for training and validation samples. The Bacterial Blight, Blast, and Tungro rice leaf disease acquired 1 AUC value from ROC for training and Validation samples. Similarly, Brown Spot rice leaf disease 1 and 0.99 AUC value from ROC for training and validation samples. Overall, the AUC of the ROC curve based on the proposed model is very close and high, indicating that the proposed model's performance is wonderful.

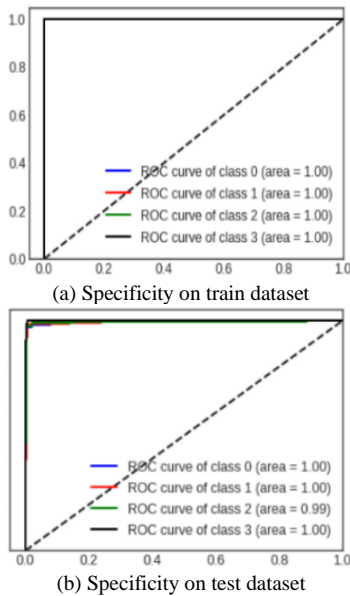


Figure 7. The ROC curves with AUC representing sensitivity analysis of the proposed model on train and test datasets.

E. Model Testing

The 10% (178) images were used as testing samples or holdout samples group to measure the trained-validated model performance. The holdout samples were separate from the part of trained samples or not include in training process. The proposed model was tested on testing samples and obtained a 100% accuracy rate with 100% precision, recall, F1-score given in Fig. 8. Out of 178 images, all were correctly classified by the proposed model. Overall, the performance of the proposed model is outstanding.

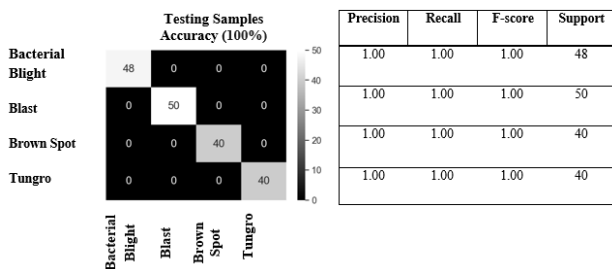


Figure 8. The Confusion matrix with accuracy and classification report of testing samples with proposed model.

F. Discussion

The researchers need an easy and reliable method to identify the rice's leaf disease confidently. To obtain high confidence, this study was designed and achieved the

desired results. However, it is observed that from the previously proposed model's high accuracy was not obtained. This study used four types of rice leaf diseases for the discrimination process. The python language tool was used for all analysis processes. First raw images are imported in python, then through OpenCV, we normalized the resize the images, smooth the images, and reshape the images of rice leaf disease for classification purposes. The rice types from individual CNN model of training samples were 99.27% correctly classified and rice types of validation samples were 95.61% correctly classified. The difference of the accuracy of training and validation samples lead the over fitting problem. The LSTM and SA layers were added to make sure the model has no problem of over fitting. The CNN+LSTM+SA proposed model was designed to identify the four rice leaf diseases. 70% (4152) images were used as training samples and 20% (1602) images were used as validation samples and 10% (178) were used as testing samples or holdout samples group to test model's prediction error. The proposed model achieved 100% precision, recall and F1-score for Bacterial blight, Blast, Brown spot, and Tungro respectively for training samples. The accuracy rate for training samples obtained from the proposed model was 100%. The 97% precision, recall, and F1-score attained for the Bacterial Blight and Blast class, 98% precision, recall, F1-score for Brown Spot class and 97% precision, 99% recall, 98% F1-score obtained for Tungro class from validation samples are presented in Table III. A 98% accuracy was obtained for the validation samples with the proposed model and 100% accuracy for training samples. The 5-fold CV approach has been applied to measure the performance of the proposed model. By adopting the CV approach 100%, 100%, 100%, 100%, and 100% accuracies were achieved from 5-folds with training samples with the proposed model in Table III. The 0.001, 0.003, 0.003, 0.001, and 0.002 losses were obtained from 5-folds for training samples. Correspondingly, 97.40%, 97.62%, 97.58%, 97.52%, and 97.47% accuracies were achieved from 5-folds with validation samples. The 0.102, 0.098, 0.121, 0.132, and 0.101 losses were obtained from 5-folds for validation samples. The average accuracy of the 5-fold CV method for training samples was 100% with a 0.001 loss rate. Comparably, the proposed model got average accuracy 97.51% by 5-fold CV method for validation samples with a 0.110 loss rate. The insignificant results of the Hosmer-Lemeshow test indicate that the model is well-fitted. The 0.0 and 3.53 values of RMSE were obtained from training and validation samples. Likewise, 0.0 and 0.007 values of MAPE were obtained from training and validation samples. The lower value of RMSE and MAPE was found for the proposed model. The proposed model was tested on testing samples and obtained a 100% accuracy rate with 100% precision, recall, F1-score. Out of 178 images were correctly classified by the proposed model. Overall, the performance of the proposed model is outstanding. Therefore, the proposed model is the most successful for rice leaves disease identification. It is also free from overfitting/ underfitting issues and can be used for disease diagnosis.

V. CONCLUSION

This paper has presented an automated diagnosis of rice leaves diseases using a hybrid deep learning model. This study applied the training-validation approach to diagnose four rice leaf disease types and measure the proposed model performance. After getting the best results of the proposed model, we applied different criteria to verify the model's goodness of fit. The 5-fold cross-validation method, Hosmer Lemeshow test, RMSE, MAPE, ROC with AUC. From the 5-fold CV method proposed model attained 100% with a 0.001 average accuracy and loss for training samples.

Similarly, it got 97.51% with a 0.110 average accuracy and loss for validation samples. The proposed model also exhibited higher ROC with AUC rate, precision, recall, and F1-score. While the model obtained minimum values of RMSE and MAPE. The proposed model obtained a 100% accuracy rate with 100% precision, recall, F1-score for testing samples. The proposed model shows better rice leaves disease diagnosis after evaluation thru standard criteria. The proposed model could identify rice leaves disease with a 97.5%-100% confidence interval. Finally, our findings will be helpful for rice plant disease diagnosis and plant recognition in agriculture.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Conceptualization, A.R. K., and I. A.; methodology, T. S., H. A.; software, H. A and B. A.; validation, I. A, A. R. K., B.A. and T.S.; writing—original draft preparation, A.R.K., I.A and H.A; writing—review and editing, T.S, B.A.; visualization, H.A.; supervision, A.R.K.; project administration, T.S., B.A.; funding, I.A., A.R. K., H.A. All authors have read and agreed to the published version of the manuscript.

FUNDING

This work was supported by the research SEED project “Intelligent and trusted metaheuristic optimization model for agriculture using ubiquitous sensors network”, (SEED-CCIS-2022{109}) under Artificial Intelligence & Data Analytics Research Lab. CCIS. Prince Sultan University, Riyadh Saudi Arabia.

REFERENCES

- [1] F. Saeed, M. A. Khan, M. Sharif, *et al.*, “The deep neural network features fusion and selection based on PLS regression with an application for crops diseases classification,” *Applied Soft Computing*, vol. 103, 107164, 2021.
- [2] N. Cherukuri, G. R. Kumar, O. Gandhi, *et al.*, “Automated classification of rice leaf disease using deep learning approach,” in *Proc. the 5th IEEE International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 2021, pp. 1206–1210.
- [3] M. H. Tunio, L. Jianping, M. H. F. Butt, *et al.*, “Identification and classification of rice plant disease using hybrid transfer learning,” in *Proc. the 2021 18th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, 2021, pp. 525–529.
- [4] M. J. Hasan, S. Mahbub, M. S. Alom, *et al.*, “Rice disease identification and classification by integrating support vector machine with deep convolutional neural network,” in *Proc. the 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*, 2019, pp. 1–6.
- [5] A. Rehman, T. Saba, M. Kashif, *et al.*, “A revisit of internet of things technologies for monitoring and control strategies in smart agriculture,” *Agronomy*, vol. 12, no. 1, pp. 127–142, 2022.
- [6] A. H. Al-Badri, N. A. Ismail, K. Al-Dulaimi, *et al.*, “Hybrid CNN model for classification of Rumex obtusifolius in grassland,” *IEEE Access*, vol. 10, pp. 90940–90957, 2022.
- [7] T. Saba, S. U. Khan, N. Islam, *et al.*, “Cloud-based decision support system for the detection and classification of malignant cells in breast cancer using breast cytology images,” *Microscopy Research and Technique*, vol. 82, no. 6, pp. 775–785, 2019.
- [8] S. Ghosal and L. Sarkar, “Rice leaf diseases classification using CNN with transfer learning,” in *Proc. IEEE Calcutta Conference (CALCON)*, 2020, pp. 230–236.
- [9] S. Ramesh and D. Vydeki, “Rice blast disease detection and classification using machine learning algorithm,” in *Proc. the Second IEEE International Conference on Micro-Electronics and Telecommunication Engineering (ICMETE)*, 2018, pp. 255–259.
- [10] R. Sharma, V. Kukreja, and V. Kadyan, “Hispa rice disease classification using convolutional neural network,” in *Proc. the Third IEEE International Conference on Signal Processing and Communication (ICPSC)*, 2021, pp. 377–381.
- [11] S. K. Upadhyay, “Rice leaves disease detection and classification using transfer learning technique,” in *Proc. the Second International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, 2022, pp. 2151–2156.
- [12] K. Gupta, A. Garg, V. Kukreja, *et al.*, “Rice diseases multi-classification: An image resizing deep learning approach,” in *Proc. the IEEE International Conference on Decision Aid Sciences and Application (DASA)*, 2021, pp. 170–175.
- [13] V. Singh and A. K. Misra, “Detection of plant leaf diseases using image segmentation and soft computing techniques,” *Information Processing in Agriculture*, vol. 4, no. 1, pp. 41–49, 2017.
- [14] T. Islam, M. Sah, S. Baral, *et al.*, “A faster technique on rice disease detection using image processing of affected area in agro-field,” in *Proc. the 2018 2nd International Conference on Inventive Communication and Computational Technologies (ICICCT)*, 2018, pp. 62–66.
- [15] T. Akram, M. Sharif, and T. Saba, “Fruits diseases classification: Exploiting a hierarchical framework for deep features fusion and selection,” *Multimedia Tools and Applications*, vol. 79, no. 35, pp. 25763–25783, 2020.
- [16] H. Yar, T. Hussain, M. Agarwal, *et al.*, “Optimized dual fire attention network and medium-scale fire classification benchmark,” *IEEE Transactions on Image Processing*, vol. 31, pp. 6331–6343, 2022.
- [17] Y. Lu, S. Yi, N. Zeng, *et al.*, “Identification of rice diseases using deep convolutional neural networks,” *Neurocomputing*, vol. 267, no. 1, pp. 378–384, 2017.
- [18] A. Islam, N. R. Shuvo, M. Shamsojjaman, *et al.*, “An automated convolutional neural network based approach for paddy leaf disease detection,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 1, pp. 280–288, 2021.
- [19] M. M. Ozguven, and K. Adem, “Automatic detection and classification of leaf spot disease in sugar beet using deep learning algorithms,” *Physica A: Statistical Mechanics and Its Applications*, vol. 535, 2019.
- [20] W. L. Chen, B. L. Yi, L. N. Fung, *et al.*, “RiceTalk: Rice blast detection using internet of things and artificial intelligence technologies,” *IEEE Internet of Things Journal*, vol. 7, no. 2, pp. 1001–1010, 2020.
- [21] A. Picon, A. Alvarez-Gila, M. Seitz, *et al.*, “Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild,” *Computers and Electronics in Agriculture*, vol. 161, pp. 280–290, 2019.
- [22] J. G. A. Barbedo, “Plant disease identification from individual lesions and spots using deep learning,” *Biosystems Engineering*, vol. 180, pp. 96–107, 2019.

- [23] A. Karlekar and A. Seal, "SoyNet: soybean leaf diseases classificatio," *Computers and Electronics in Agriculture*, vol. 172, 2020.
- [24] M. S. Rahman, M. K. H. Sohag, L. Rahman, *et al.*, "Genetic fingerprinting for the protection of local rice (*Oryza sativa* L.)," *Journal of Advanced Biotechnology and Experimental Therapeutics*, vol. 3, no. 3, pp. 182–193, 2020.
- [25] T. Saba, A. Rehman and J. S. AlGhamdi, "Weather forecasting based on hybrid neural model," *Applied Water Science*, vol. 7, no. 7, pp. 3869–3874, 2017.
- [26] R. Deng, Y. Jiang, M. Tao, *et al.*, "Deep learning-based automatic detection of productive tillers in rice," *Computers and Electronics in Agriculture*, vol. 177, 2020.
- [27] I. Abunadi, "Deep and hybrid learning of MRI diagnosis for early detection of the progression stages in Alzheimer's disease," *Connection Science*, vol. 34, no. 1, pp. 2395–2430, 2022.
- [28] H. Pallathadka, P. Ravipati, G. S. Sajja, *et al.*, "Application of machine learning techniques in rice leaf disease detection," *Materials Today*, vol. 51, pp. 2277–2280, 2022.
- [29] A. H. Al-Badri, N. A. Ismail, K. Al-Dulaimi, *et al.*, "Classification of weed using machine learning techniques: A review — Challenges, current and future potential techniques," *Journal of Plant Diseases and Protection*, vol. 129, pp.745–768, 2022.
- [30] R. Deng, M. Tao, H. Xing, *et al.*, "Automatic diagnosis of rice diseases using deep learning," *Frontiers in Plant Science*, vol. 12, 2021.
- [31] V. P. Bhartiya, R. R. Janghel, and Y. K. Rathore, "Rice leaf disease prediction using machine learning," in *Proc. the 2022 2nd International Conference on Power, Control and Computing Technologies (ICPC2T)*, 2022, pp. 1–5.
- [32] S. A. Naz, M. U. Javed, N. Javaid, *et al.*, "Short-term electric load and price forecasting using enhanced extreme learning machine optimization in smart grids," *Energies*, vol. 12, no.5, pp. 866–878, 2019.
- [33] J. Chen, D. Zhang, *et al.*, "Using deep transfer learning for image-based plant disease identification," *Computers and Electronics in Agriculture*, vol. 173, 2020.
- [34] X. H. Le, H. V. Ho, G. Lee, *et al.*, "Application of long short-term memory (LSTM) neural network for flood forecasting," *Water*, vol. 11, no. 7, pp. 1387–1399, 2019.
- [35] M. A. Khan, T. Akram, M. Sharif, *et al.*, "CCDF: Automatic system for segmentation and recognition of fruit crops diseases based on correlation coefficient and deep CNN features," *Computers and Electronics in Agriculture*, vol. 155, pp. 220–236, 2018.
- [36] K. Aurangzeb, F. Akmal, M. A. Khan, *et al.*, "Advanced machine learning algorithm based system for crops leaf diseases recognition," in *Proc. the 2020 6th Conference on Data Science and Machine Learning Applications (CDMA)*, 2020, pp. 146–151.
- [37] T. Saba, I. Abunadi, M. N. Shahzad, *et al.*, "Machine learning techniques to detect and forecast the daily total COVID-19 infected and deaths cases under different lockdown types," *Microscopy Research and Technique*, vol. 84, no. 7, pp. 1462–1474, 2021.
- [38] S. Abbas, M. A. Khan, M. Alhaisoni, *et al.*, "Crops leaf diseases recognition: A framework of optimum deep learning features," *Comput. Mater. Contin.*, 2023.
- [39] M. S. Anari, "A hybrid model for leaf diseases classification based on the modified deep transfer learning and ensemble approach for agricultural AIoT-based monitoring," *Computational Intelligence and Neuroscience*, 2022.
- [40] Opencv-python 4.7.0.72. Wrapper package for OpenCV python bindings. [Online]. Available: <https://pypi.org/project/opencv-python/>
- [41] P. K. Sethy. Rice leaf disease image samples. [Online]. Available: <https://data.mendeley.com/datasets/fwcj7stb8r/1>
- [42] D. W. Hosmer, T. Hosmer, and S. A. Lemeshow, "Goodness-of-fit tests for the multiple logistic regression model," *Communications in Statistics*, vol. 10, pp. 1043–1069, 1980.

Copyright © 2023 by the authors. This is an open access article distributed under the Creative Commons Attribution License ([CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/)), which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.