ISSN: 2366-1313

# Rice Leaf Diseases Detection Using Random Forest Algorithm

[1] SHAIK JAILANI, [2] M. NARESH

[1]PG Scholar, Dept. of MCA, Newton's Institute of Engineering, Koppanur, Macherla Guntur, (A.P)

[2]Associate Professor, Dept. of CSE, Newton's Institute of Engineering, Koppanur, Macherla Guntur, (A.P)

***Abstract**: Diseases affecting rice crops in different parts of the globe cause a great deal of destruction and loss. Many different kinds of organisms, including fungus, Bacteria, and viruses, are responsible for the devastating rice diseases that plague farmers across the world. There are many different ways to categorise rice. In this study, we utilise image classification to categorise a dataset including information on three of the most prevalent rice leaf diseases in Thailand: bacterial leaf blight (BLB), brown spot rice (BSR), and bacterial leaf blight (BSR). In addition, the accuracy, precision, and recall of various image processing algorithms for classifying rice leaf diseases. These algorithms include the Random Forest classification algorithm, the decision tree classification algorithm, the gradient boosting classification algorithm, and the naive bayes classification algorithm. The random forest method had the highest performance level of 69.44 percent when used to the picture classification of rice leaf diseases.*

***Keywords**: Rice Leaf Diseases, Algorithm, Image processing, Classification, Random Forest.*

## I. INTRODUCTION

Classifying images based on statistical measures. Its goal is to classify images based on their unique characteristics. This is the resemblance of picture points to the same class [7] determined by a model based on mathematical concepts and statistical analysis.

It's possible to classify data using one of two approaches: supervised learning or unsupervised learning. The focus of this study is on discovering methods for extracting information from digital images. Supervised learning technique employs the data to determine picture classification. Data mining and information extraction in the context of image-based learning [16]. The features of picture categorization serve as a teaching tool. To make inferences about the kind of data represented by fresh images, we may use a classification model [14]. In specifically, we use the machine learning approach for classification purposes [10] by first extracting the vector of characteristics numerical picture.

The fundamental idea of classifying visual information: 1) A plethora of data-linked photos

As Nutnicha Teeyasuksaet

ISSN: 2366-1313

Chiang Mai, Thailand's Department of Livestock Development's Fifth Regional Livestock Office

The statistical and data mining model is needed at nutnichavet@gmail.com. Two) Many programmes provide features designed to make data mining easier for end users. It's compatible with high-level languages like Python and R. Parameter adjustments in data analysis are, nevertheless, crucial and essential[21].

Therefore, our researcher came up with a method to use the random forest algorithm for picture classification of rice leaf diseases[1]. We also looked into how it stacked up against other popular classification strategies including decision trees, gradient boosts, and naive bayes. The picture categorization system we developed is simpler and quicker than competing methods [17], and our experiments demonstrate that it achieves excellent accuracy scores.

## II. LITERATURE SURVEY

Numerous studies have used BPSO and SVM as a backbone for their own wrapper-based attribute selection methodologies. In order to increase the effectiveness of picking highly predictive features, authors have also suggested modifications in particle swarm.

To this end, Agrafiotis and Cedeno (2002) use PSO and a Neural Network (NN) to narrow down candidate attribute subsets. Due to the problem-specific nature of the attribute space, it might be challenging for users to determine an appropriate size for the attribute subset to be selected using this technique.

A wrapper-based attribute selection approach for locating keystroke dynamic systems was developed by Azevedo et al. (2007) using BPSO and SVM. To solve this challenge, Tu et al. (2007) developed a particle swarm optimizer–SVM–one-versus-all method–attribute selection strategy.

Wrapper-based attribute selection using BPSO and SVM was developed by Lin et al. (2008). The suggested technique iteratively optimises the SVM parameters while it explores the issue space in search of the best subset of attributes. While Lin et al. (2008) used BPSO to optimise attribute subset search, Huang & Dun (2008) used continuous valued PSO to optimise SVM parameters.

Selecting a subset of qualities based on their probability and their effect on other attributes previously contributed is the basis of an adaptive selection technique presented by Unler and Murat (2010). When compared to Tabu search and the

ISSN: 2366-1313

scatter search algorithm, the results from this technique are clearly better. The authors also say that reducing the minimum number of characteristics needed to add to the subset may increase the effectiveness of the adaptive selection technique and the quality of the solution.

In order to optimise parameters in SVM and the search for an attribute subset with high prediction accuracy at the same time, Liu et al. (2011) suggest a multi-swarm BPSO approach. This approach uses a larger number of swarm particles for the search and relies on intricate communication between the several sub-swarms, which increases the computing burden.

Xu and Rahmat-Samii (2007) explain that in the BPSO method of wrapper-based attribute selection, the location of each particle is a binary string in Hamming space with no boundary requirements. Therefore, there is no divergence issue; nonetheless, there are two circumstances when there is a premature convergence problem:

## III . Materials and Methods

Data analysis with knife is a tool for picture categorization. The most crucial part of the operation is the data node connections. The first step is to feed information into your eyes. The image's parameters may be changed, and features can be extracted [20] to form a data table with attributes and values. This is used in the context of ML algorithms [3]. Diseases of the rice leaf are the main topic of our investigation. Accessible at (https://archive.ics.uci.edu/ml/datasets/Rice+Leaf+Dise aces), this database was used in this study. It categorised three different forms of rice leaf disease:

Spotting, A. Brown Disease of rice, or BSR

One of the most common and damaging rice diseases, BSR, has been largely ignored.

Infected coleoptiles, trees, leaf sheaths, and panicle roots all suffer from brown spot, a fungal disease. Fungus (Helminthophobia Oryza) is to blame, and the fungus may be carried by the wind and dropped on immature rice seeds, where it will damage them, turn them white, and lower their quality. The most obvious harm is the many large blotches on the leaves, which may wipe out the whole harvest. Infected seeds may have abnormal shapes, such as hollowed-out grains, or develop spots or abnormal colouring.

ISSN: 2366-1313

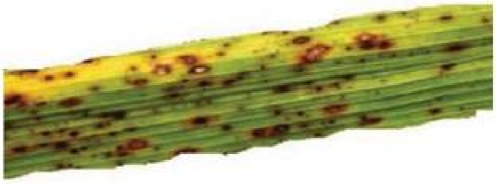

Fig. 1. Brown spot of rice disease

B. Leaf Smut Rice disease: LSR It is a common epidemic in high humidity. It is caused by fungus (Ustilaginoidea virens). This fungus destroys rice seeds then creates many fibres and seeds in that seed until it emerges out of the seed group of fibres and reproduction on this seed to be dark green colour. However, it found on some grass species and destroyed by the fungus.

Fig. 2. Leaf smut of rice disease

C. Bacterial Leaf Blight disease: BLB

Xanthomonas Oryza is responsible for causing bacterial leaf blight, often known as BLB. The result is wilting seedlings and dead leaves [15]. The disease is particularly prevalent in areas where the infected plants are surrounded by weeds and stubbles.

Especially in lowland irrigated and rainfed locations, it is possible to find it in both tropical and temperate environments.

The optimal conditions for the sickness are 25–34 degrees Celsius and relative humidities exceeding 70 percent. When there are strong gusts and continuous heavy rains, the bacteria that cause the illness may spread rapidly on lesions of affected plants through droplets of ooze. High nitrogen fertilisation of vulnerable rice types may exacerbate bacterial blight [4].

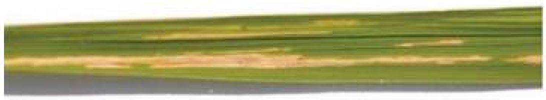

Fig. 3. Bacterial leaf blight of rice disease

There are four different types of rice leaf diseases, and each has its own unique set of symptoms that are shown in a JPEG picture series totalling 120 photographs.

The ratio of training data to test data is 70:30.

The images' filenames will be defined with the first three characters to help identify the type of rice leaf disease and create the target column in the learning process, and the training datasets will be randomly partitioned to create a model for predicting the disease [19]. Increased processing processes, such as identifying the position of the point indicating the type of disease on the rice leaf disease image, etc. [18], have resulted from the prevalence of

ISSN: 2366-1313

diseases like Brown Spot Rice (BSR), Leaf Smut Rice (LSR), and Bacterial Leaf Blight (BLB) [6, 7]. Model performance is measured using Fig. 4 as input to the prediction model.

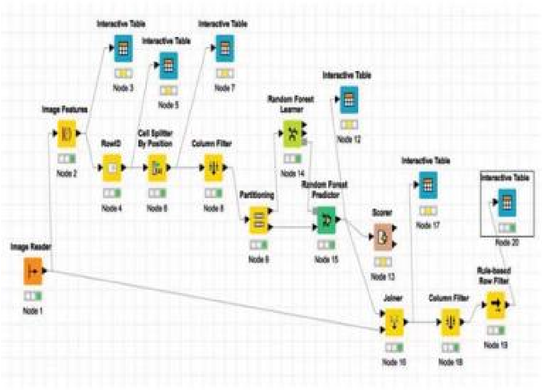


Fig. 4. Flow algorithm for image classification of rice leaf diseases

## IV. Results

Our study details the algorithms used for image classification of rice leaf diseases, including the random forest algorithm [11], the decision tree algorithm [12], the gradient boosting algorithm [13], and the naive bayes algorithm [14]. The results show that the random forest method achieves the best accuracy (69.44%), followed by the gradient boosting technique (66.67%), and finally the naive bayes algorithm (36.11%). And the best accuracy in each category is as follows: 69.23 for BLB, 80.00 for BSR, and 70.00 for LSR using Gradient Boosting, Random Forest, and Decision Tree respectively.

In addition, Naive Bayes equals 91.67 in BLB recall, whereas Random Forest equals 66.67 in BSR recall and 83.33 in LSR recall (see TABLE I for details).

TABLE I PREDICTION PERFORMANCE OF IMAGE CLASSIFICATION METHODS

| Methods | Precision | | | Recall | | | ACC |
|---|---|---|---|---|---|---|---|
| | BLB | BSR | LSR | BLB | BSR | LSR | |
| Random Forest | 63.64 | **80.00** | 66.67 | 58.33 | **66.67** | **83.33** | **69.44** |
| Decision Tree | 60.00 | 63.64 | **70.00** | 75.00 | 58.33 | 58.33 | 63.89 |
| Gradient Boosting | **69.23** | 63.64 | 66.67 | 75.00 | 58.33 | 66.67 | 66.67 |
| Naive Bayes | 34.38 | 0.00 | 66.67 | **91.67** | 0.00 | 16.67 | 36.11 |

## V. Discussion

The results of our research, it can be concluded that the random forest algorithm is the highest of accuracy for image classification by compare with the other algorithm classification. And show the effectiveness of the classification of each rice leaf disease using the random forest algorithm that can display picture of various rice leaf diseases, including with target and prediction result as shown in Fig. 5-7

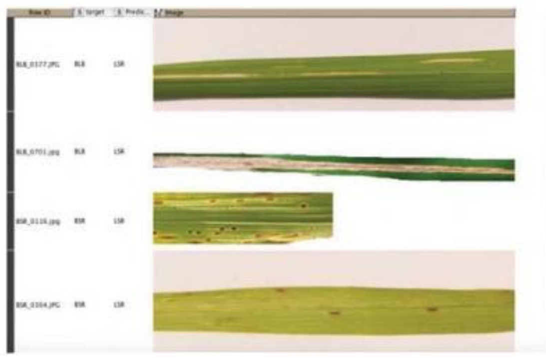

Fig. 5. Target (BLB, BSR) and Prediction (LSR) of Rice Leaf Diseases

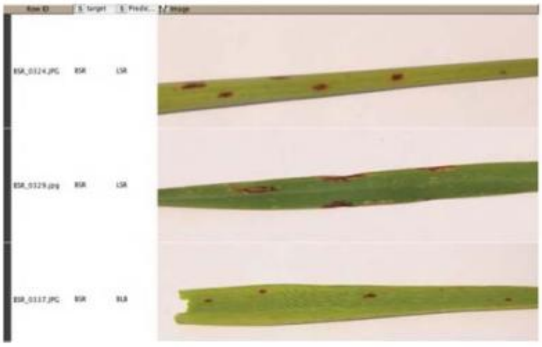

Fig. 6. Target (BSR) and Prediction (LSR, BLB) of Rice Leaf Diseases

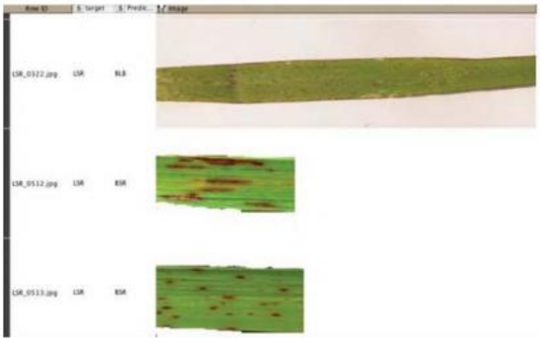

Fig. 7. Target (LSR) and Prediction (BLB, BSR) of Rice Leaf Diseases

## VI. CONCULSION

In this study, we introduce the method for image-based disease classification of rice leaf lesions. The random forest method was determined to have the greatest accuracy for picture classification at 69.44%.

In addition, this study provides a visual representation of the flow algorithm for image classification of rice leaf diseases image processing by knife method, which can be used to analyse and extract the distinctive features of each type of rice leaf disease.

**Acknowledgement**

This our research was supported by Department of Computer Science, Faculty of Science, Major University and The Fifth Regional Livestock Office, Department of Livestock Development.

## REFERENCES

[1] Archana, K. S., & Sahayadhas, A. (2018). Automatic rice leaf disease segmentation using image processing techniques. Int. J. Eng. Technol, 7(3.27), 182-185.

[2] Bosch, A., Zisserman, A., & Munoz, X. (2007, October). Image classification using random forests and ferns. In 2007 IEEE 11th international conference on computer vision (pp. 1-8). Ieee.

[3] Chaudhary, A., Kolhe, S., & Kamal, R. (2016). An improved random forest classifier for multi-class classification. Information Processing in Agriculture, 7(4), 215-222.

[4] Gnana Manickam, S. S., Priyadarshini, V. B., Narayanan, N. N., Vasudevan, P., & Kavitha, S. (1999). An overview of bacterial blight disease of rice and strategies for its management. Current Science, 1435-1444.

[5] Li, P., Wu, Q., & Burges, C. (2007). Mcrank: Learning to rank using multiple classification and gradient boosting.

ISSN: 2366-1313

Advances in neural information processing systems, 20, 897-904.

[6] Lu, D., & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. International journal of Remote sensing, 28(5), 823-870.

[7] Mekha, P., Osathanunkul, K., & Teeyasuksaet, N. (2016, December). Gene classification of dengue virus type based on codon usage. In 2016 International Computer Science and Engineering Conference (ICSEC) (pp. 1-6). IEEE.

[8] Mekha, P., Osathanunkul, K., & Teeyasuksaet, N. (2017). Comparison of Gene Classification Methods for Dengue Virus Type Based on Codon Usage. FEU Academic Review, 11, 7-7.

[9] Mekha, P., & Teeyasuksaet, N. (2015, November). Using random forest based on codon usage for predicting Human Leukocyte Antigen gene. In 2015 International Computer Science and Engineering Conference (ICSEC) (pp. 1-6). IEEE.

[10] Mekha, P., & Teeyasuksaet, N. (2019, January). Deep learning algorithms for predicting breast cancer based on tumour cells. In 2019 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTIDAMT-NCON) (pp. 343-346). IEEE.

[11] Murphy, K. P. (2006). Naive bayes classifiers. University of British Columbia, 18(60).

[12] Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. Frontiers in neurorobotics, 7, 21.

[13] Ok, A. O., Akar, O., & Gungor, O. (2012). Evaluation of random forest method for agricultural crop classification. European Journal of Remote Sensing, 45(1), 421-432.

[14] Patil, J. K., & Kumar, R. (2011). Advances in image processing for detection of plant diseases. Journal of Advanced Bioinformatics Applications and Research, 2(2), 135-141.

[15] Perumal Samy, S., Bharani, M., Sudha, M., Nagarajan, P., Arul, L., Saraswathi, R., ... & Ramalingam, J. (2010). Functional marker assisted selection for bacterial leaf blight resistance genes in rice (Oryza sativa L.). Plant breeding, 129(4), 400-406.

[16] Phadikar, S., & Sil, J. (2008, December). Rice disease identification using pattern recognition techniques. In 2008 11th International Conference on

ISSN: 2366-1313

Computer and Information Technology (pp. 420-423). IEEE.

[17] Prajapati, H. B., Shah, J. P., & Dabhi, V. K. (2017). Detection and classification of rice plant diseases. Intelligent Decision Technologies, 11(3), 357-373.

[18] Rastogi, A., Arora, R., & Sharma, S. (2015, February). Leaf disease detection and grading using computer vision technology & fuzzy logic. In 2015 2nd international conference on signal processing and integrated networks (SPIN) (pp. 500-505). IEEE.

[19] Sanyal, P., & Patel, S. C. (2008). Pattern recognition method to detect two diseases in rice plants. The Imaging Science Journal, 56(6), 319-325.

[20] Shoombuatong, W., Mekha, P., Waiyamai, K., Cheevadhanarak, S., & Chaijaruwanicha, J. (2013). Prediction of human leukocyte antigen gene using k-nearest neighbour classifier based on spectrum kernel. Science Asia, 39(1), 42-49.

[21] Shoombuatong, W., Mekha, P., & Chaijaruwanich, J. (2015). Sequence based human leukocyte antigen gene prediction using informative physicochemical properties. International journal of data mining and bioinformatics, 13(3), 211-224.

[22] Prasadu Peddi (2021), "Deeper Image Segmentation using Lloyd's Algorithm", ZKGINTERNATIONAL, vol 5, issue 2, pp: 1-7.