Paddy Leaf Disease Prediction Using Machine Learning

Rajeswari Nakka[1*], Vinay Kagitha[2], Teja SreeKinthala[3], Eliyam Kodali[4], Vinod Jampani[5]

Abstract

The economic stability of a nation rests largely on its agricultural sector. The health of the agricultural industry hinges heavily on the ability to produce disease-free rice. Procedures and techniques that are productive and efficient in increasing harvest yield are contingent upon a number of preconditions. Agriculture, including paddy farming, is essential to human societies. Our species has an ethical obligation to safeguard agriculture's vitality and growth potential. Plant diseases have been on the rise recently, in part because of the increased use of man-made chemicals and pesticides. You can't just brush off these illnesses in crops, because they could pose a threat down the line. Sometimes it's hard to recognize certain disorders because of a lack of technical understanding. In this research, we introduce machine learning-based technique for identifying rice leaf diseases. Rice Blast, bacterial blight, and leaf smut are the three most common leaf diseases that reduce rice yields. As input, we provided visuals of the sickened leaves. Leaves will display symptoms unique to the disease the plant is suffering from if that disease is present. The data set was then trained using the machine learning algorithm SVM (Support Vector Machine) after the pre-processing was complete

KeyWords:support vector machine, Agriculture, Rice leaf diseases, machine learning

DOI Number: 10.48047/NQ.2022.20.16.NQ880303 **NeuroQuantology2022; 20(16):2932-2937**

Introduction

Estimates for total rice production in Indiain the 2021-22 crop year (July-June) show a data high of 130.29 million tonnes, up from the previous record of 124.37 million tonnes. Over 2.7 billion people rely on paddy as their main source of nutrition; it is included in the world's three most dominant food crops. It is estimated that rice accounts for 80% of all irrigated crops(s BRRI 2019). It's distributed over the majority of India's states. The states that rely most heavily on rice production are West Bengal, Madhya Pradesh,UttarPradesh, Bihar, Orissa, Andhra Pradesh, Assam, Tamil Nadu, Kerala, Maharashtra, Punjab, and Karnataka, which together account for 92% of both area and production. Natural disasters and illnesses that attack paddy leaves have far-reaching effects on paddy farming and production, even more so than on the economy as a whole. Researchers atthe Tamil Nadu Agricultural University found that the paddy leaf disease known as Leaf Blast would cause an estimate70–80% of grain to be lost. The disease, which has already spread to 80 nations where rice is grown after being first documented inIndia in 1918, goes by several names. Brown Spot, also known as Helminthosporiumoryzae, is another important disease that has a negative impact on productivity in the Indian states like West Bengal, Orissa, Andhra Pradesh, and Tamil Nadu. In severe circumstances, it can impair crop yield by as much as 50 percent and can be found both in the nursery and the main field, causing blight of seedlings.The rice disease known as "leaf smut" occurred by a fungal bacteria named Entylomaoryzae. On both sides of infected leaves, this fungus leaves behind black, angular patches (sori) that are slightly elevated. Leaf sheath spots are an unusual adverse effect. The length of the black speck's ranges from 0.5 milli meters to 5 milli meters, while their width varies from 0.5 milli meters to 1.5 milli meters. A single leaf may sport a number of distinct spots, which do not merge into one another. It is common for leaves affected with leaf smut to crack open when wet, releasing the black spores inside.

Corresponding author: Rajeswari Nakka

Leaves that have been seriously diseased turn a sickly yellow and the edges of the leaves dwindle and become a greyish colour. When it comes to actually feeding humans, rice is far and by the most dominant crop in the world. This is especially true in Asia, home to the largest concentration of the world's poor. Increasing agricultural productivity is essential for the beneficial development of a country like India. Timeline-wise, farming exists primarily to produce food for the populace. Leaf diseases, in whatever form they manifest, are a common reason for lowered rice quality, yield, and economic growth. Thus, considering the current farming area, where it does require a great deal of effort to implement, we recommend using a more efficient method, such as remote sensing, to monitor crops.

Since leaf disease identification and control are so crucially automated, pen-and-paper calculations have become a reality of great magnitude.Therefore, this work can describe the answer towards cost reduction by eliminating the need for human monitoring and expert requirement for automatic identification of leaf diseases across a wide area.Using leaf symptoms, this research suggests a strategy for forecasting and categorizing those aforementioned three rice illnesses. The novelty of this research is in its application of machine learning methods to the problem of reliably diagnosing illnesses in rice leaf samples..(Nakka Rajeswari et al, 2021) Accurate disease detection and timely control activities required to ensure good quality and high output.

Literature Survey

To prevent rice plant diseases, (Manoj Mukherjee et al,2012) provided a substructure for processing photos of paddy leaves using the circle graph. This structure allows for early detection of infections, allowing for remedial action to be taken before significant output losses occur. Initially, a picture of the leaf was taken and then edited. Subsequently, MATLAB operations were used to convert the image from its original RGB colour space to grayscale, and then to extract its histogram. The obtained images were proposed as data for disease classification and severity ranking. Disorder and its phase were identified with the help of agrarian experts, who then organised a consultation treatment unit.

To help those who aren't familiar with digital technology, such as farmers, find the information they need on the Internet, (Namita Mittal et al,2014) proposed an Icon-centric Information Retrieval framework. Farmers were also able to quickly and accurately diagnose the nature of crop disease, its origins, and its manifestations through the use of pattern identification and digital visualization, without having to wait for advisors to go to their fields. The results of these experiments validated the validity of this approach. The results based on training with 25 photos from each illness category. Increasing the total number of training photographs has the potential to further enhance the results.(D.N.V.S.L.S. Indira et al, 2022)The product of images is used for extracting features where the Convolutional neural network (CNN) features are computed, and then it is given as input to the deep residual network aimed at product recommendation.

In order to automate the process of identifying leaf diseases, (Gayathri and Neelamegam 2018) proposed a framework based on image processing techniques. Processes such as pre-processing, picture collection, analysis, and paddy leaf disease arrangement were imposed to aid in the identification of these afflictions. Features are extracted using an approach that combines the Discrete Wavelet Transform (DWT), the Gray Level Co-occurrence Matrix (GLCM), and the scale-invariant feature transform. Finally, to classify the healthy and diseased plants, the extracted characteristics were fed into various classifiers such as K-Nearest Neighbours (KNN), neural network (NN)(BaburaoMarkapudi et al,2021) backpropagation, multiclass SVM, and Naive Bayesian. To better understand how to classify leaf diseases, researchers looked at a number of different approaches.

(Harshadkumar et al,2017) analysed a model for spotting and categorising paddy diseases based on images of sick rice plants. BLB, Leaf smut, and Brown spot are three diseases that have been linked to rice plants. These photos were taken using a digital camera on rice farms. There was a total of four approaches to background elimination, and three segmentation strategies, that were evaluated. It was suggested that centroid feeding centric KMC be used to accurately extract features when segmenting the sick area from a leaf picture. KMC's output was improved by eliminating green pixels in the infected area. Colour, texture, and form were used to demand radically different physical characteristics. Multi-class classification was accomplished using SVM.

- Leaf smut, characterised by small black linear lesions on leaf blades and possibly grey and dry leaf tips.

• Bacterial blight, characterised by the development of elongated lesions along the leaf margins and tips, and subsequent discoloration from white to yellow to grey as a result of fungal infection.

Brown spot, also called rice leaf blight, is characterised by lesions that range in size from tiny dots to larger brown spots.

Fig 1: Rice Blast

Fig 2:Bacterial Blight

Fig 3:Brown Spot

A.K. Singh, director of the Indian Council of Agricultural Research, claims that the disease has destroyed anywhere from 6% to 20% of the cultivated land in some areas. Damage to rice crops from diseases can significantly lower harvests. The majority of these conditions are brought on by microorganisms like bacteria, viruses, or fungi. Healthy rice plants can produce more grain-bearing bunches, a process known as tillering, so yield losses are "likely to be limited to 1-2 per cent," according to Singh. Additional space for the rice plant to expand into means more grain-bearing branches, and thus a higher yield, which could mitigate the negative effects of the loss.

Pyriculariaoryzae is a fungus that causes rice blast. This ailment is not selective and can infect paddy at anytime during its development (Leaf, neck and node). Leaves, nodes, and panicles are especially vulnerable, but the disease affects the entire plant except the roots (neck). Infected nodes are characterised by the presence of irregular black areas all around the nodes, which cause the parts of the plant above the nodes to become disjointed. Poor or low grain filling is the result of neck blast, which is caused by a fungus attacking the peduncle during the flower emergence stage. The prevalence of paddy leaf diseases presents significant challenges to farmers. Because of this, rice production has been drastically cut. Illnesses affecting the paddy plant's leaves can be detected with this method.

There is currently no reliable disease detection model for paddy leaf issues. It's possible to simulate certain rice leaf diseases with existing models. There are a large number of models in place that cover the entire range of leaf diseases. Our proposed system is unique, however, in its ability to predict diseases in rice leaves.

Constraints of previous research

There needs to be a stepping up of effort and a sharpening of accuracy in the current execution in order to locate the desired outcomes. For an image analysis, having precise time evidence is crucial. Increasing the length of records is a desirable goal in pursuit of greater precision. fewer items new diseases have been found. That's why it's crucial to broaden the scope of efforts in order to find new diseases.

Classifiers used

Diseases in paddy crops were forecast using the support vector machine (Support vector machine). Outlier detection, classification, and regression are just some of the many applications of the supervised learning models (SVM) family of algorithms.(BalaBrahmeswaraKadaru et al,2017) SVM employs an iterative process to generate an optimal hyperplane in an effort to minimise an error. Finding the hyperplane with the highest marginal gain is at the heart of SVM. To put the data set into meaningful categories. SVM is an outstanding algorithm for classification tasks. SVM performed flawlessly when extracting the one-dimensional feature vector.

It is determined that a hybrid network is the best method for accurately classifying images of rice leaf diseases for the purpose of prediction. This procedure entails two stages. Example processes include classification and feature extraction. For the purposes of feature extraction, images have their own quirks. SVM performs and generalises well on data that is excluded in the training set.Properly, it's effective even on data that lies outside the bounds of typical generalisation. (BalaBrahmeswarakadaru et al, 2018)In SVM, the kernel function is evaluated and performed for each support vector, so it classifies a single sample in a fraction of a second. Moreover, SVM's in-built functionality can be used; the algorithm is implemented in popular programming environments like Python and Matlab; and SVM can be applied to both non-linearly separable data with a soft margin and linearly separable data with a hard margin. and outlined, making their integration into SVMs a breeze.

The N-dimensional space used in SVM corresponds to the total characteristic or attributes in the dataset. Locate the best hyperplane for data partitioning next. The SVM can only perform binary classification. This is why we employ SVM for data classification.

When compared to other classifiers, the SVM Classifier has a lower computational complexity, and it can be used when there is an imbalance between the number of positive and negative examples because it can normalise the data or project into the space of the decision boundary.

Using a feature selection strategy improves classification throughput and precision. The process of feature extraction involves selecting and representing only the most important aspects of the original data in a lower dimensional space. Since a classifier cannot recognise an image based on poorly selected features, feature selection is a crucial post-processing step.

Software and hardware requirements

Paddy crop disease prediction requires Windows 7 or later. We advise a minimum of 8GB of RAM, a 500GB hard drive or SSD, and a third-generation high-end or Risen processor from Intel. To use Anaconda, you'll need Python 3.6 or later. In this case, we use the Jupyter notebook editor. This time around, we use the Flask framework.

Dataset preparation

One of the many things done with this dataset is learning how to extract different features from images inorder to use them for disease detection and classification in rice plants. The second concerns what kinds of image processing operations can yield the required data, and the third focuses on the image features that can serve as useful input for classification.

Rice Plant Disease Diagnosis and Classification Through the use of SVM, a model for classifying data was created. In order to divide the data into distinct categories, SVM locates a hyper-plane. This hyper-plane is just a line in standard 2-dimensional geometry. The N-dimensional space used in SVM corresponds to the total characteristics or attributes in the dataset. Locate the best hyperplane for data partitioning next.

The first step in preparing the dataset is to collect images of the disease in question, followed by pre-processing.

Diseases depicted as examples

To begin this collection of disease images, it is necessary to first acquire photographs of rice plants. This article focuses on three paddy leaf diseases: rice blast, bacterial blight, and leaf smut.

Pre-processing:

This preliminary processing will augment the figure and quality of portrayals in the dataset. A lot of how well your feature extraction works depends on the number, variety, and quality of the image samples you use. Performing pre-processing allows for a large increase in the number of samples without a corresponding decrease in image quality. Cropping these massive images with so many samples by hand is being done so that we can have more images to work with overall.

Proposed work

The primary focus here is using the support vector machine which is one of the machine learning algorithm to identify diseased rice leaf tissue. This data is mined from an existing collection. The following are some advantages of a well-organized

procedure

At the time of image analysis, the user plays no role whatsoever. Improved capacity for accurate diagnosis. As opposed to other methods, this one is entirely hands-free.

It provides natural recycling processes for the known diseases. It is crucial to keep tabs on the incidence and prevalence of diseases in order to quickly identify infected plants, administer treatment, and, pre-eminently,devise upcoming planning for preventing the illness in order to limit losses. Conventional crop disease management entails experts manually spotting abnormalities in plants, labelling them as diseases, and then recommending treatments. When dealing with large farms, this chain of tasks becomes extremely difficult. It's not just a waste of time, but also of effort. Diseases can be detected and classified much more accurately by comparing portraits of the infected area of leaves with a pre-trained model. This paper proposes a method that can be used to predict and classify the three rice diseases mentioned above. Training set includes 90% of the data, whereas the test set comprises the remaining 10%. The proposed method exemplified a scheme for categorising rice plant diseases, which helps farmers make more precise calculations and boosts output.

Block diagram

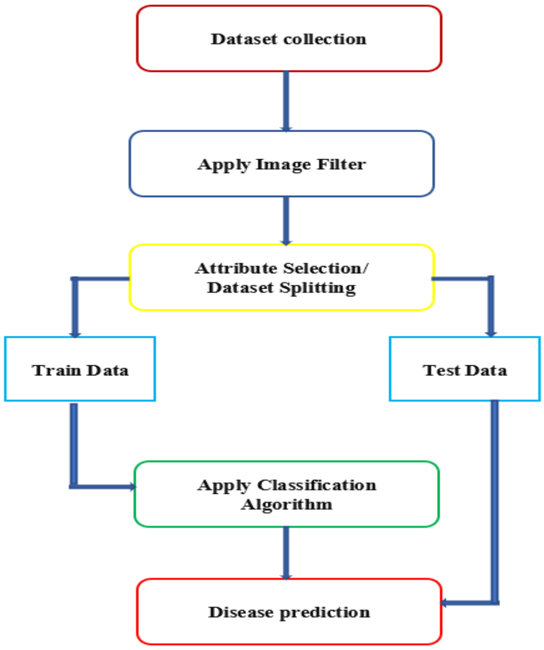

Fig 4: Flow Chart

The preceding block diagram depicts a typical work process. The dataset is then expanded in size, an image filter is applied, attributes are chosen, and the collected information is divided into training and testing sets. Next, a categorization model is applied to the classified training data. Three diseases—leaf smuth, rice blast, and bacterial blight—are being predicted using the classification model after initial data testing. Disease forecasting for paddy crops follows this procedure.

Conclusion

In this paper, we exhibit a machine learning technique for identifying leaf smut, bacterial blight, and Rice blast disease, three common diseases that can affect rice plants. The frameworks designed by (Shruthi Aggarwal et al, 2022) utilizes around thousands test spot pictures of infected leaves of the rice crop gathered from the field, giving accuracy of 79% for Bayes and 68% for SVM classifier oriented framework. As we observe the results, we achieved an overall accuracy of 83%.we achieved a precision value of 83% in identifying the bacterial leaf blight and a recall value of 83% and f1-score too. The weighted average of the precision,recall, f1-scores is 0.81, 0.81 and 0.81 respectively. The macro average of the f1-score, recall, precision is 0.80, 0.82, 0.80 respectively. Similarly, the precision,recall, f1-score values of the brown spot disease is 0.70, 0.88, 0.78 and for the disease named leaf smut is 0.86, 0.75, 0.80 are the precision, recall, f1-score values respectively. When compared to While predicting the leaf without any defects we achieved a 93% accuracy. The Mean accuracy of the data set considered and tested is 80.5%.

The results are as follows

```
Classification report for -
GridSearchCV(estimator=SVC(),
             param_grid=[{'C': [1, 10, 100, 1000], 'kernel': ['linear']},
                         {'C': [1, 10, 100, 1000], 'gamma': [0.001, 0.0001],
                          'kernel': ['rbf']}]):
              precision    recall  f1-score   support

           0       0.83      0.83      0.83        12
           1       0.70      0.88      0.78         8
           2       0.86      0.75      0.80        16

    accuracy                           0.81        36
   macro avg       0.80      0.82      0.80        36
weighted avg       0.81      0.81      0.81        36
```


eISSN1303-5150

References

Mukherjee Manoj, Titan Pal and DebabrataSamanta, "Damaged V. N. paddy leaf detection using image processing", Journal of Global Research in Computer Science, vol. 3, no. 10, pp. 07-10, 2012.

Bangladesh Rice Research Institute(BRRI), Rice in Bangladesh. [Online].Availablehttp://www.knowledgebank-brri.org/riceinban.php.Accessed: 16-Feb-2019].

Harshadkumar Prajapati B, Jitesh P Shah and Vipul K Dabhi,"Detection and classification of rice plant diseases", Intelligent Decision Technologies, vol. 11, no. 3, pp. 357-373, 2017.

Ghosal S, Sarkar K. Rice leaf diseases classification using cnn. with transfer learning. 2020 IEEE Calcutta Conference (CALCON) 2020;230–236.

Gayathri Devi T and Neelamegam P, "Image processingbased. rice plant leaves diseases in Thanjavur, Tamilnadu", Cluster Computing, pp. 1-14, 2018.

C.-L. Chung, K.-J. Huang, S.-Y. Chen, M.-H. Lai, Y.-C. Chen, and Y.-Fu Kuo, "Detecting Bakanae disease in rice seedlings by machine vision," Computers and Electronics in Agriculture, vol. 121, pp. 404–411, 2016.

R. R. Atole and D. Park, "A multiclass deep convolutional neuralnetwork classifier for detection of common rice plantanomalies," International Journal of Advanced Computer Science and Applications, vol. 9, pp. 67–70, 2018.

K. Ahmed, T. R. Shahidi, S. M. IrfanulAlam, and S. Momen, "Riceleaf disease detection using machine learning techniques," in Proceedings of the 2019 International Conference on. Sustainable Technologies for Industry 4.0, Dhaka, Bangladesh, December 2019.

Shruti Aggarwal,1M. Suchithra,2N. Chandramouli,3Macha Sarada,4Amit Verma,5D. Vetrithangam,6Bhaskar Pant,7and BirukAmbachew Adugna8 "Rice Disease Detection Using Artificial Intelligence and Machine Learning Techniques to Improvise Agro-Business", Volume 2022 | Article ID 1757888 |https://doi.org/10.1155/2022/1757888

N Rajeswari, E. Keerthi, K.Mahesh, K.Pavan Sai "An Automatic GUI Based Grape Leaf Disease Identification Using CNN" UGC Journal Vol. 26 No.3,December 2021.

BalaBrahmeswara.Kadaru et al. "An Overview of General DataMining Tools" ", International Research Journal of Engineering and Technology (IRJET) Volume 4, Issue 9 , September 2017,930-936.

BalaBrahmeswara.Kadaru et al. "A Novel Ensemble decision tree classifier using hybrid feature selection measures for Parkinson's disease prediction" ", Int.J.Data Science (IJDS),ISSN:2053-082X,Vol.3,No.4,2018.

BaburaoMarkapudi, Kavitha Chaduvula, D.N.V.S.L.S. Indira &Meduri V. N. S. S. R. K. Sai Somayajulu, "Content-based video recommendation system (CBVRS): a novel approach to predict videos using multilayer feed forward neural network and Monte Carlo sampling method", published in the journal of Multimedia Tools and Applications (Springer Nature), published on 11th August 2022. PP:1-27, (2021).

D.N.V.S.L.S. Indira, Babu Rao Markapudi, Kavitha Chaduvula, Rathna Jyothi Chaduvula, "Visual and buying sequence features-based product image recommendation using optimization based deep residual network", published in the journal of Gene Expression Patterns, Volume 45, 2022, September 2022, ISSN 1567-133X