Data Collection and Image Classification for Rice Disease Detection using Cloud-based Application

Anmolpreet Singh Kataria
Department of Software Engineering
Delhi Technological University
New Delhi, India
axndtu@gmail.com

Sanjay Patidar
Department of Software Engineering
Delhi Technological University
New Delhi, India
sanjaypatidar@dtu.ac.in

Piyush Gupta
Department of Software Engineering
Delhi Technological University
New Delhi, India
piyush.gupta5800@gmail.com

Abstract — India is a country that has recently undergone a digital revolution with telecom tariffs going down and smartphone penetration increasing every day. It is a very well-known fact that India's progress is highly dependent on the progress of its agricultural sector and almost half of its population's livelihood comes from farming. In such a scenario, combating crop diseases becomes a very essential requirement. Rice is one of the most produced crops in India and rice crop diseases are a farmer's nightmare. With machine learning and artificial intelligence finding applications in every field, crop disease identification using image processing seems like the perfect use case of using automated systems to figure out the diseases.

This paper proposes the use of a cloud-based application that can help the farmers figure out what disease is affecting their rice crop. The application uses cloud computing that stores a Residual Neural Network bases classification model to predict the disease. At the same time, the application also allows the storage of the images on the server which can be used to create a new dataset and solve the problem of the availability of a large free dataset for rice crop diseases. The application can identify three diseases namely, Leaf Smut, Brown Spot, and Bacterial Leaf Blight. The application provides an attractive solution that can be expanded to other crops by just changing the trained model at the backend.

1. INTRODUCTION

A lot of crops are grown in India, names include Rice, Wheat, Sugarcane, Coffee, and Tea, among others. This directly stems from the fact that India has a diverse variety of climate and soil conditions. In fact, agriculture contributes 17-18% to the country's GDP and provides a means of livelihood to more than 50% of the people [1]. All of this makes any issue related to agriculture a very major obstacle to be tackled. Rice is one of the crops with the highest area under cultivation. But rice can become affected by a variety of diseases.

Diseases like Rice Blast, Sheath Blight, Brown Spot and Sheath rot are common. Diseases like these render a lot of production destroyed, leading to immense financial loss, and given India's dependence on the agriculture sector, this becomes an issue that requires immediate attention. A lot of farming land in India is owned by small farmers. Farmers are largely uneducated, and the measures taken by the government to educate them, haven't been as successful as they need to be.

The farmers currently rely on manual inspection to identify the diseases that affect their crops and must guess the disease based on their experience. Implementing remedies that can effectively tackle the diseased crop becomes an even bigger issue if the disease cannot even be identified. The digital revolution in India has had a much higher penetration than the farmer education measures.

A large proportion of people in rural India own a smartphone and use cellular data, which was made possible by the recent drop in tariffs by telecom operators like Jio and Airtel. A cloud-based application that allows the farmers to capture the photos of the crop and upload them and returns the disease that the crop has been affected by is the solution proposed by this research. The application can be modified to provide effective remedies for the disease, which would eradicate manual guessing and allow farmers to tackle the problem with ease.

The research on rice crop disease detection has been hampered by the unavailability of publicly available free datasets of labelled rice crop disease photos. The proposed application will also provide an effective medium to capture data, in the form of photos, which would help in building a more accurate and advanced model for the identification of diseases.

1.1 RICE PLANT DISEASES

Rice plants are susceptible to a variety of diseases. It becomes important for the reader to know about the major rice diseases because it provides an insight into how the disease looks in images, which allows him to better understand the task of image processing and classification. The following table summarizes the diseases in rice plants. The table tells the initial, mild, and severe symptoms of the diseases. The following part of the section shows what the plant looks like when it is affected by the disease.

Disease	Symptoms	
Rice Blast	Initial	Blue-green water-soaked lesions on leaves, grey centre-brown margin boat-shaped spots
	Mild	Major area of leaves are dried up and withered, similar spots are formed on the sheath
	Severe	Burnt Appearance of the field
Brown Spot	Initial	Leaf blade and plumes get spots as small brown dots
	Mild	The expanding brown spots to merge and leaf dries up
	Severe	Brownish scorched appearance
Sheath Blight	Initial	Greenish grey spots near the water level on leaf sheath
	Mild	Expanding spots with grey, white centre and blackish-brown border
	Severe	Large spots cause the death of the whole leaf and eventually the whole plant
Bacterial Leaf Blight	Initial	Translucent spots near leaf margin
	Mild	Irregular enlarging spots expand to cover the entire leaf
	Severe	Entire leaf turns white or straw yellow
Leaf Smut	Initial	Slightly Raised angular black spots on both sides of leaves
	Mild	Heavily affected leaves turn yellow
	Severe	Leaf tips die and turn grey

Table 1: Rice Plant Diseases and their symptoms

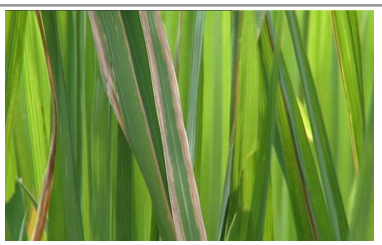
Rice Blast	
Brown Spot	
Sheath Blight	
Bacterial Leaf Blight	
Leaf Smut	

Fig 1: Different types of rice plant diseases

1.2 PROCESS OF RICE DISEASE CLASSIFICATION

Image classification for rice disease detection follows a systematic approach that is depicted by figure 2.

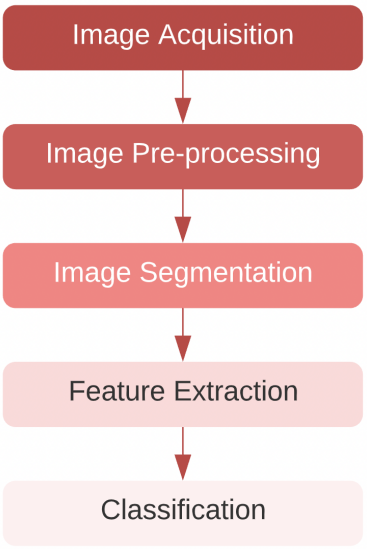


Fig 2: Process of rice disease image classification

1. Image Acquisition: The task of collecting a labelled dataset to train our classifiers is known as image acquisition. Properly labelled and varying images of all the diseases can help train a better classifier. This becomes a task due to the lack of free publicly available quality datasets. Researchers must resort to the manual collection of data and labelling to conduct their research.

2. Image Preprocessing: Images in any dataset can have a variety of features in them that are not useful for research purposes and often act as a barrier to successful classification. Features like pests on the plant, any fertilizer marks, dew, etc. act as noise for the classification models. Shadows are a major issue in image classification. Images might need to be cropped to obtain the area of interest. All these features can be removed using various filters.

3. Image Segmentation: Separating the areas of the image or dividing them into regions is the aim of image segmentation. It can help separate areas on the affected plant. For example, the brown parts of a leaf are separated from the green parts using edge detection or Otsu's Method.

4. Feature Extraction: Feature extraction is the process of identifying the distinguishing characteristics of the image. Feature extraction plays a pivotal role because the features extracted in this step are passed on to the classifier. Different colours, textures and shapes can help differentiate one disease from another.

5. Classification: Classification is the machine learning task of classifying a new image based on the images provided when training the model. The features provided by the image processing tasks along with the labels on the images help train a model that can classify features extracted from a new image based on the similarities between them.

2. RELATED WORK

This field of research, that is, the usage of machine learning in improving agricultural productivity encompasses a wide array of topics. Some of these are weed inspection, soil analysis and seed selection. Getting to know the research done already helps the reader to know what the status quo is regarding the current topic of discussion and helps him proceed further in the rigphotosrection. This section aims to provide a survey of the current progress by the survey of a few publications.

2020 research by Sanjay et al. [2] compares the classification of rice plant diseases by an SVM classifier, a convolutional neural network, and a residual neural network. Removing image preprocessing as the focus and zooming in on improving the accuracy of classification, they use a ResNet to classify the images. The dataset used in the research is Rice Leaf Disease Dataset from the UCI Machine Learning Repository [3]. They succeed in achieving an accuracy of 95.83% using the ResNet.

Research conducted by Jayanti et al. [5] uses Fuzzy C-mean clustering to perform edge-detection on diseased leaves, extracts the features using SURF techniques and feeds them into an artificial neural network. Another research conducted by Prajapati et al. [4] performed on a dataset of diseased rice leaves succeeded in achieving an accuracy of 93.33% on the training dataset and an accuracy of 88.57% on the test dataset using 10-fold cross-validation.

A study by Jayme Garcia et al. [6] describes the techniques for image processing. They also provide a survey of different machine learning models applied to the classification of plant diseases. Another comparative study by Shah et al. [7] compares research done in the detection and classification of rice plant diseases. The authors provide a study on various image processing techniques used in this field and the consequent machine learning techniques for the same.

3. PROPOSED WORK

This section provides details on our proposed methodology for implementing a cloud-based application that can classify rice plant diseases and serve as a data collection model too. We won't be diving too deep into the image preprocessing and classification aspect of the process, and rather focus on the delivery of the results to the farmers using a mobile application and online website.

3.1 DATASET

A dataset publicly available in the UCI machine learning "Rice leaf disease dataset" library was used to test the implementation of the model in the application. The dataset consists of 120 images from three diseases namely, leaf smut, brown spot, and bacterial leaf blight. Figure 3 shows the distribution of the images in the dataset.

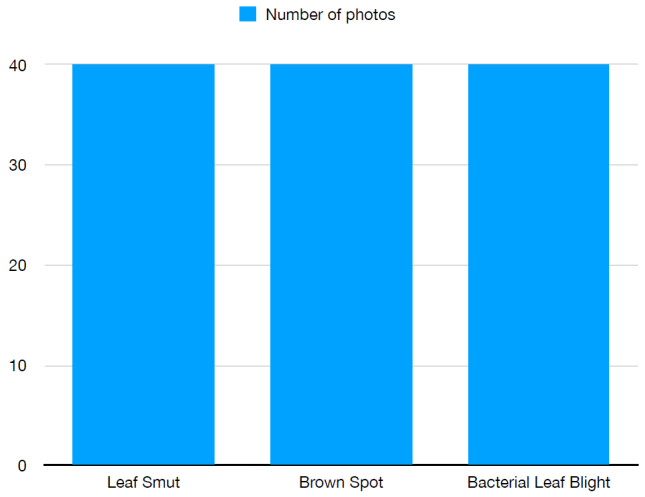


Fig 3: Distribution of the dataset

3.2 IMAGE PREPROCESSING

The image preprocessing methodology is shown briefly as a part of figure 3. The images have undergone treatment with a transformation function. The transformation function has performed basic preprocessing steps like increasing contrast and flipping the images.

To normalize the intensity of all pixels and help the classifier learn and perform faster, we have used a normalization function [2]. Using this function the images are rescaled and enhanced. This provides a standard learning rate for our training model making it more accurate and efficient.

3.3 CLASSIFICATION

The classification uses a Residual neural network for the multi-class classification of the photos into three classes, namely, leaf smut, brown spot, and bacterial leaf blight. The table below depicts the architecture of the ResNet, as used by the research conducted in 2020 [2].

Layer name	Output size	34-layer Resnet
conv1	112 × 112	7 × 7, 64, stride 2
conv2_x	56 × 56	3 × 3 max pool, stride 2 $\begin{bmatrix} 3\times3, & 64 \\ 3\times3, & 64 \end{bmatrix} \times 3$
conv3_x	28 × 28	$\begin{bmatrix} 3\times3, & 128 \\ 3\times3, & 128 \end{bmatrix} \times 4$
conv4_x	14 × 14	$\begin{bmatrix} 3\times3, & 256 \\ 3\times3, & 256 \end{bmatrix} \times 6$
conv5_x	7 × 7	$\begin{bmatrix} 3\times3, & 512 \\ 3\times3, & 512 \end{bmatrix} \times 3$
	1 × 1	Average pool, 1000-d fc, softmax
FLOPS		3.6×10^9

Table 2: ResNet Architecture

The model is fine-tuned using various learning rates, and the learning rate with minimum loss is kept, the rest are discarded. In the end, the SoftMax layer classifies the leaves according to the highest probability of occurrence.

3.4 CLOUD-BASED APPLICATION

This section gives an insight into the tools and technologies used in developing the cross-platform application. Our aim is to keep our application lightweight so that it can be installed on any device and used no matter the specifications of the device. For this, we shift all the processing and classification on the server-side and on the clientside only the photos are uploaded and the results are shown after processing.

The proposed methodology allows the farmers to click a photo of the leaves and upload them to the server whether through the mobile application or the website. The server processes the images using the machine learning model saved and returns the results to the farmer. The application also displays remedies for fighting the disease that the rice plant has been affected by. The application support three diseases namely, Leaf Smut, Brown Spot and Bacterial Leaf Blight.

The application also functions as a data collection source for further improvement of the trained model and data analysis. The data collection aspect of this application may also be useful for various research purposes such are identification of trends of diseases, the discovery of new diseases etc.

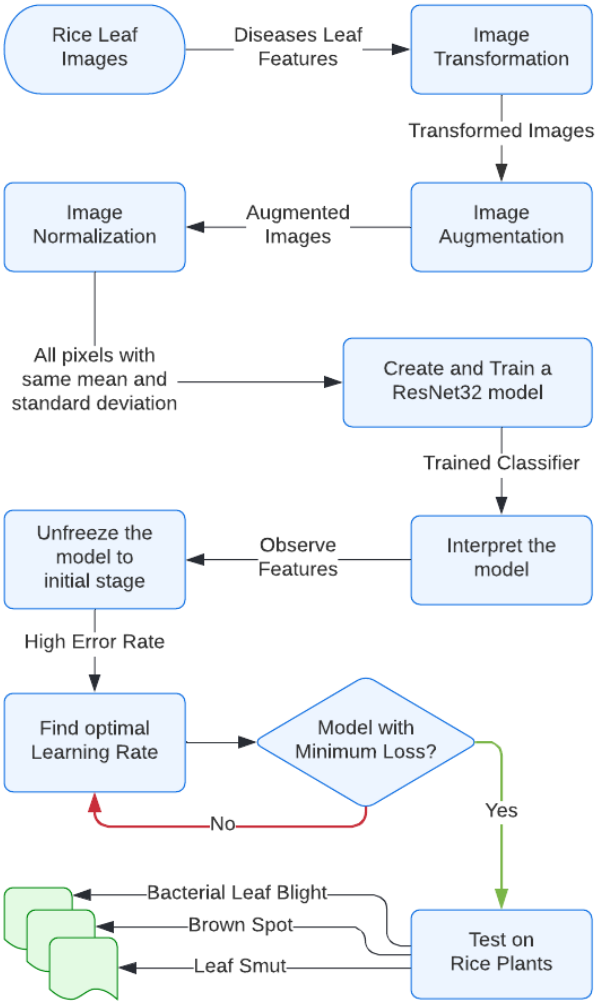


Figure 4: Image Classification Flowchart

3.5 TOOLS AND TECHNOLOGIES

We have chosen Flutter for developing the front end of the application because, out of the two top contenders for cross-platform app development, flutter provides better performance and documentation. Render provides a very lucrative free tier that allows us to test out the app at very low costs and the availability of render across the globe allows us to scale the solution to other countries. Thus we hosted our classification model on render.

3.5.1 FLUTTER

Flutter is a cross-platform app development SDK that allows us to develop applications with beautiful interfaces. Flutter is a kit backed by Google that allows developers to develop front ends for applications that work great on both android and iOS. Flutter uses dart as the language and complies it with native code for both platforms. Flutter also supports Web and Desktop using the same codebase. This allows the users to reduce the amount of code written and maintain the apps using the same codebase. Thus we used flutter as our frontend.

3.5.2 DOCKER

It is a cloud service that provides the end-user with Paas (Platform as a service) applications. provides the user with OS-level virtual environments also known as containers which are made to provide compute, storage and network services needed by an application. We used it for storing our classification model so that the classification can be performed on the server-side instead of the client-side making the whole process of classification reliable and efficient.

3.5.3 RENDER

It is the cloud service that allows us to host Saas (Software as a service) applications. It provides a host of solutions that allows the users to eradicate the need for an actual physical server and provides the server as a service. Thus we deployed our docker image that provides serverless computing for our classification model on render. It provides all the capabilities required by a web application or a mobile application. It can be accessed through https://rpdlc.onrender.com/ [7].

3.6 IMAGE CLASSIFICATION ON APPLICATION

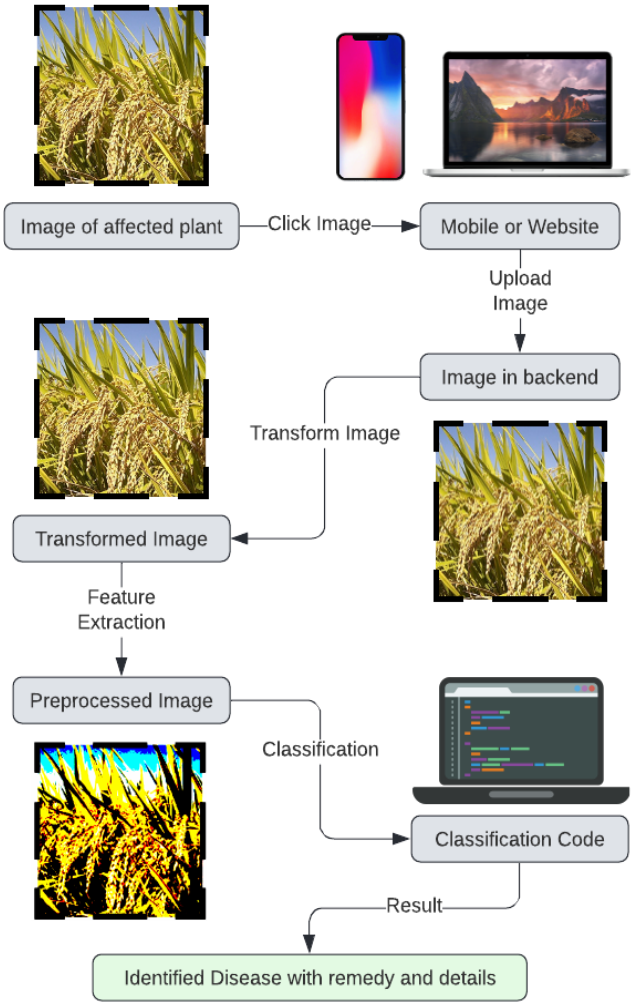


Fig 5: The cloud-based application classification process

The application follows a series of steps to classify the image and return the results to the farmer. This section describes the process. The procedure is also described by the flow chart in figure 5.

1. The farmer downloads the application from Google Play Store or Apple App Store or can simply visit our website. The application allows the farmer to click photos of the affected leaf or upload them. Note that data connection is a requirement for the application to work because classification needs to be done on the server-side.

2. The photos are sent to the docker container hosted at render. A python script processes the image as per the procedure defined above and extracts the features such as the colour, texture and shape of the spots.

3. Another python script uses the trained model to predict the highest probability disease and returns it to the frontend whether it is the mobile device or our website.

4. The application then displays the results received from the backend and prints the remedies for the disease.

3.7 DATA COLLECTION

The application can serve to collect new images. This will solve the problem of the availability of datasets for this purpose. Additionally, the images collected will be of higher quality and presents the opportunity to collect images of new diseases for agricultural research purposes. Any images uploaded by the farmers are stored on the servers. The images can be accessed and annotated by manual annotation, which can serve as a new dataset for training.

4. RESULTS AND ANALYSIS

This section gives the details of the experiment performed and the results obtained. We used a publicly available dataset “Rice Leaf Disease Dataset” from the UCI Machine Learning Repository [3]. The dataset had 40 photos for each of the 3 types of diseases. The learning rate for the ResNet was set to 1e-03, which gave the lowest loss [2].

The highest accuracy achieved was 95.83% using the ResNet. The other performance measures are summarized in the table given below.

Performance Measure	Value
Accuracy	95.83%
Recall	0.94
Precision	0.94
F-measure	0.94

Table 3: Results of the classification model

5. CONCLUSION

Rice crop diseases cause an immense loss to the agricultural dependent communities in India and worldwide, and in turn cause the economies huge losses. Thus, combating these diseases becomes an essential step in progress towards agricultural advancement. A cloud-based application that allows the farmers to figure out what disease is affecting their crop is an efficient solution to this problem. The paper proposes the architecture for the mobile application and uses an approach that allows the farmers to get remedies for the disease right on their mobile phones.

Such an application, amid the current national progress towards creating a digital India, is a perfect recipe to solve a problem that plagues the farmer's community. We have used a residual neural network that is trained on a dataset of 120 images and predicts the disease by one photo with an accuracy of 95.83%. The application uses cloud computing to build a solution that can be scaled on demand.

6. FUTURE WORK

The application we have proposed in this research has immense potential to be applied to an ever-expanding field of plant protection. The machine learning model can be trained to identify other diseases, which may or may not belong to the rice plant. The application can be scaled to provide a solution for generalised plant disease prevention. The application stores the model in the cloud, so the model can be updated if newer advancements with better accuracy are made in the rice crop disease field.

7. REFERENCES

[1] "India economic survey 2018: Farmers gain as agriculture mechanisation speeds up, but more R&D needed". The Financial Express. 29 January 2018. Retrieved 8 January 2019.

[2] Patidar, Sanjay & Pandey, Aditya & Shirish, Bub & Sriram, A.. (2020). Rice Plant Disease Detection and Classification Using Deep Residual Learning. 10.1007/978-981-15-6315-7_23.

[3] Shah, J.P., Prajapati, H.B., Dabhi, V.K.: Rice leaf diseases data set. Accessed 14 Mar 2019.

https://archive.ics.uci.edu/ml/datasets/Rice+Leaf+Diseases.

[4] Prajapati, H.B., Shah, J.P., Dabhi, V.K.: Detection and classification of rice plant diseases. J. Intell. Decis. Technol. 11(3), 357–373 (2017). IOS Press 2017.

[5] Jayanthi, G., Archana, K.S., Saritha, A.: Analysis of automatic rice disease classification using image processing techniques. Int. J. Eng. Adv. Technol. (IJEAT) 8(3S), 2249–8958 (2019)

[6] J. Garcia, J A. Barbedo, "Digital image processing techniques for detecting, quantifying and classifying plant diseases," Springer Plus, vol. 2, pp. 1-12, 2013.

[7] Shah, Jitesh & Prajapati, Harshadkumar & Dabhi, Vipul. (2016). A survey on detection and classification of rice plant diseases. 1-8. 10.1109/ICCTAC.2016.7567333.

[8] Classify Rice Plant Disease into Leaf Smut, Brown Spot and Bacterial Leaf Blight. https://rpdlc.onrender.com/