# RECOGNIZING RICE DISEASE AND FEATURE BASED CLASSIFICATION OF RICE VARIETY USING MACHINE LEARNING

**A Project Report submitted in partial fulfillment of the requirements for the award of the degree of**

**BACHELOR OF TECHNOLOGY**

**IN**

**COMPUTER SCIENCE AND ENGINEERING**

**Submitted by**


| | |
|---|---|
| **Para Meenakshi Chowdary** | **121710302062** |
| **K Sai Sujith** | **121710302061** |
| **V Nikhil Kumar** | **121710302064** |
| **Ch Sravan** | **121710302011** |

**Under the esteemed guidance of**

**SRINIVASA RAO GADU**

**Assistant professor, GIT**

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**GITAM**

**(Deemed to be University)**

**VISAKHAPATNAM**


**MAY 2021**

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**GITAM INSTITUTE OF TECHNOLOGY**

**GITAM**

**(Deemed to be University)**

# DECLARATION

We at this moment declare that the Project review entitled "**RICE DISEASE PREDICTION AND RICE GRAIN CLASSIFICATION USING MACHINE LEARNING**" is an original work done in the Department of Computer Science and Engineering, GITAM Institute of Technology, GITAM (Deemed to be University) submitted in partial fulfillment of the requirements for the award of the degree of B.Tech. In Computer Science and Engineering. The work has not been submitted to any other college or university to award any degree or diploma.

Date: May 2021.

| Registration No: | Name | Signature |
|---|---|---|
| 121710302062 | Para Meenakshi Chowdary | |
| 121710302061 | K Sai Sujith | |
| 121710302064 | V Nikhil Kumar | |
| 121710302011 | Ch Sravan | |

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**GITAM INSTITUTE OF TECHNOLOGY**

**GITAM (Deemed to be University)**

## BONAFIDE CERTIFICATE

This is to certify that the project report entitled **"RICE DISEASE PREDICTION AND RICE GRAIN CLASSIFICATION USING MACHINE LEARNING"** is a bonafide record of work carried out by Para Meenakshi Chowdary**(121710302062),** submitted in partial fulfillment of the requirement for the award of the degree of Bachelors of Technology in Computer Science and Engineering.

| PROJECT GUIDE | HEAD OF THE DEPARTMENT |
|---|---|
| Srinivas Rao Gadu | Dr.K.THAMMI REDDY |
| (ASSISTANT PROFESSOR) | (PROFESSOR) |
| CSE,GIT | CSE,GIT |
| GITAM | GITAM |

## ACKNOWLEDGEMENT

We would like to thank our project guide **SRINIVASA RAO GADU**, Assistant professor, Department of CSE, for stimulating guidance and profuse assistance. We shall always cherish our association with him for his advice, encouragement, and valuable suggestions throughout this work's progress. We consider it a great privilege to work under her guidance and constant support.

We also express our thanks to the project reviewers, **G. Srinivasa Rao**, Associate Professor, **Naveen Kumar Kuppili**, Assistant professor, Department of CSE, GITAM(Deemed to be University), for their valuable suggestions and guidance for doing our project.

We consider it a privilege to express our deepest gratitude to **Dr.K.Thammi Reddy**, Head of the Department, Computer Science Engineering, for his valuable suggestions and constant motivation, helping us complete this project.

Our sincere thanks to **Dr.C. Dharma Raj**, Principal, GITAM Institute of Technology, GITAM (Deemed to be University), for inspiring us to learn new technologies and tools.

Finally, we deem it a great pleasure to thank one and all that helped us directly and indirectly throughout this project.

Para Meenakshi Chowdary 121710302062

## TABLE OF CONTENTS

## LIST OF FIGURES



## LIST OF TABLES

# ABSTRACT

India mostly depends upon rice for its economy and for meeting its food demands. However, farmers are facing severe loss in rice production due to crop diseases. Due to a lack of knowledge about disease and pesticides available to control the disease, farmers can lose crop yield. However, finding the current disease and providing the best remedy requires expert advice or prior knowledge to prevent the disease. Nevertheless, manual detection of disease costs a large amount of time and labor, so it is inevitably prudent to have an automated system to detect disease. To solve the above problem, we are developing a Machine Learning model using a CNN algorithm to detect the rice crop disease using the image and provide a suitable remedy. After necessary pre-processing, the dataset was trained with a range of different machine learning algorithms and achieved a classification accuracy of 97.17% and 99.45% when applied to the test dataset. These remedies give information on pesticide use to control the disease. In India, most people consume rice as food, so demand is also increasing. Rice grain can be classified according to physical and chemical characteristics. Grain size and shape, chalkiness, witness, milling degree, bulk density, and moisture content are physical characteristics. While amylase content, generalization temperature, and gel consistency are some of the chemical characteristics. Here we classify rice grains based on grain shape, size, and color. Edge detection is used to find out the region of boundaries of rice grain. After extracting necessary features from the image, they are saved for further classification. The dataset was trained with a range of different machine learning algorithms and achieved an accuracy of 91.30% on the Decision Tree Classifier.

# INTRODUCTION

[1]Rice Disease Prediction

Rice gives protein and energy to the greater part of the total population. Rice consumption and demand expanded with the development of the population. Despite increased food demand, rice diseases have caused a great deal of loss in yield. The use of pesticides and the deployment of blast-resistant cultivars are the main methods of combating the disease. However, excessive use of pesticides increases the cost of rice production and causes considerable environmental damage. Moreover, in practice, the diagnosis of rice blasts is often manually conducted, which is subjective and time-consuming even for well-experienced experts. In modern agricultural practices, it is crucial to managing pests and diseases using highly efficient methods with minimum damage to the environment. An automated rice disease computerized detecting system could provide information for the prevention and control of rice disease,minimize economic loss, reduce pesticide residues, and improve the quality and quantity of agriculture production. Research in practical algorithms of feature extraction and classification of rice disease is critical to achieving such a system.

India needs to work for its industrial advancements, which will involve intelligent systems that can take decisions without any human interventions. So we came up with an idea to develop an automated system using machine learning techniques. This system will contribute to agricultural development by automatically identifying and classifying diseases from the images of rice leaves.

Rice blast and brown spots were considered the most prominent diseases in 1979-81, but now brown spots and bacterial blight are considered the most prominent and dangerous rice diseases. We have identified three rice leaf disease detection (bacterial blight, brown spot, and leaf smut). The reason for choosing these three diseases is the prevalence of these diseases in India.

The features of the three diseases i.eLeaf blast, hispa and brown spot are described below.

**LEAF BLAST** :

Leaf Blast is caused by the fungus *Magnaporthe oryzae.* However, leaf impact occurrence will in general reduce as plants develop constantly grown-up plant protection from the illness.

<u>Symptoms</u>

- In the first stage of disease, symptoms appear as white to gray-green lesions or spots with dark green borders.
- Older lesions on the leaves are elliptical or spindle-shaped and whitish to gray centers with red to brownish or necrotic borders.

**HISPA** :

- Rice hispa scrapes the upper surface of leaf blades, leaving only the lower epidermis and also makes plants less vigorous by tunneling through leaf tissues.

<u>Symptoms</u>

- causes white streaks parallel to midrib leaving only the lower epidermis
- The wilting of harmed leaves
- whitish and membranous leaves

**BROWN SPOT** :

Brown spot is a contagious illness that taints the leaves, branches, and spikelets. Its most apparent harm is the various enormous spots on the leaves, which can execute the entire leaf. For contamination to happen, the leaves should be wet for 8−24 hours.

<u>Symptoms</u>

- At the initial stage lesions can be observed on the leaves. They are at first little, round, and dull earthy colored to purple-brown colored.
- At the final stage lesions become circular to oval with a light brown to gray center, surrounded by a reddish brown margin.

Leaf Blast

Brown spot

Hispa

Monitoring the occurrence of diseases and their frequencies are important to prevent the loss in crop production. Traditionally, crop disease management is carried out by manual detection of any irregularity in plants, then classification of that irregularity as disease by experts and finally recommending appropriate treatment. It causes additional time and labor as well. Unexpectedly, taking the images of the affected area of the leaf and detecting it with a pre-prepared model gives a better detection and classification of diseases output. So we worked on an methodology that makes disease prediction and classification of the three rice diseases mentioned above by using machine learning approaches with high accuracy.

[2]Rice Variety Classification

Rice is the single most important food in the Asian countries, especially in India. Averagely 60.45% of cultivation is occupied by the total cultivation in India. For the global demand, rice is exported to the other countries and global rice demand is increased by 1.95% annually. As technology is growing, industries and peoples are adapting to new technologies rather than using old techniques. People consuming rice is increasing drastically and also bagging and packaging is becoming automated. In the current grain handling system, grain type and quality are quickly evaluated by visual investigation

The conventional techniques utilized for grain size and shape estimation are dial micrometer,grain shape analyzer and graphical strategy, however these strategies are very tedious. In the dial micrometer and grain shape analyzer we can quantify the length and broadness of single grains all at once. The result of this investigation is likewise relative, tedious, having variable outcomes and exorbitant. So we proposed a visual inspection model to satisfy the customer's needs. The purpose of the image processing techniques is extracting the parameters of rice grain to classify the rice variety. The quality of the rice grain is based on several parameters. Such as grain color, shape, and size. Digital images are the key sources of the machine vision systems.

In this system we have predicted 6 types of rice grains. Namely Basmati, Brown Rice, BPT, Navara Rice, NLR and Sona Masuri.

# LITERATURE SURVEY

[1] Rice Disease Prediction

a). Rice Leaf Diseases Classification Using CNN

Journal: IEEE

Authors: Shreya Ghosal, Kamal Sarkar

Rice is one of the major developed yields in India which is experiencing different crop diseases at different phases of its development. It is hard to physically recognize these infections precisely with their restricted information. Late advancements in Deep Learning show that Automatic Image Recognition frameworks utilizing Convolutional Neural Network(CNN) models can be extremely advantageous in such problems.

b).Rice Disease Identification using Machine learning

Journal: Instructables

Author: Santanu Phadikar and Jaya Sil

It describes a software prototype system for rice disease detection based on the infected images of various rice plants. Images of the infected rice plants are captured by camera and processed using image growing, image segmentation techniques to detect infected parts of the plants. Then the infected part of the leaf has been used for the classification purpose using neural networks. The methods evolved in this system are both image processing and soft computing techniques are applied on a number of diseased rice plants.

**c).** Rice Disease Identification and Classification by Integrating Support Vector Machine With Deep Convolutional Neural Network

Journal: IEEE

Over the planet , rice is one among the foremost extensively consuming staple foods. But thanks to various diseases of rice, farmers can't yield rice consistent with their expectation. Therefore, the standard , quantity and production of rice are being interrupted. It is difficult for the farmers to diagnose diseases, due to the shortage of data and unavailability of professional specialists. It requires time in identifying and classifying rice diseases. Therefore, an automatic and accurate identification system has become essential to beat this problem.

d). Plant Disease Detection using Machine Learning

Journal: International Research Journal of Engineering and Technology

The identification and detection of diseases of plants is one among the most points which determine the loss of the yield of crop production and agriculture. The studies of disease are the study of any visible points in any part of the plant which helps us differentiate between two plants, technically any spots or color shades. The sustainability of the plant is one of the key points that is for agricultural development.

[2] Rice Variety Classification

A) Classification of Rice Grains Using Image Processing And Machine Learning Techniques

Journal : IEEE

In this study, the classification of rice grains was tested using image processing and machine learning techniques. There are three-grain types used for variety. But broken grains were considered as a type too. Thus a total of 4-grain types were used for variety. Rice grain images obtained by a webcam. Six attributes were extracted for each grain image.GrainAttributes are related to their shape geometry. The most successful nearest neighbor with a generalization algorithm was selected for real-time testing. Real-time test classification result's accuracy was calculated as 90.5% with the selected classifier.

B) Geometric Feature Extraction of Selected Rice Grains using Image Processing Techniques

Journal : International Journal of Computer Applications 124(8):41-46

The developed application is independent of camera position and effectively identifies seven geometrical features (Major axis length, Minor axis length, eccentricity, area, orientation, perimeter, aspect ratio) of the selected five different rice grain varieties.
Features were extracted after performing basic functions on digital image.
Based on Major axis length it further classifies the rice grains into three classes. The developed approach is useful to identify the percentage of different Major axis length rice grains in a sample. The error rate of measuring different geometric features between proposed method and experimental analysis is found between - 1.39% and 1.40%

## PROPOSED WORK

[1]Rice Disease Prediction

In this system for detecting the Rice Brownspot, Leaf Blast and Rice Hispa disease of paddy, it involves major two phases one is training the model and the other part is detecting the given image of the disease. So the first step is to train the model using the image dataset. Both healthy and disease leaf image dataset are gathered. Here we have collected 523 images of Rice Brownspot, 779 images of Leaf Blast, 565 images of Hispa and 1448 healthy paddy leaf images. These images are trained using the Convolutional Neural Network (CNN) Algorithm.

A basic CNN model contains an input layer, multiple hidden layers and output layer. The hidden layer consists of Convolution layer, Rectified Linear Unit, pooling layer and fully connected layer. The CNN architecture for the proposed model is shown in figure 4.1. The input layer takes the resized, gray scaled image and the outcome of the model is detected disease and recommendations to control the disease.

The proposed system has concentrated on detecting the paddy diseases and provides the suitable remedies, thus leading to an increase in paddy crop production. In this system it detects the most common and frequently occurring paddy diseases (leaf blast and brown spot) and provide a remedy.

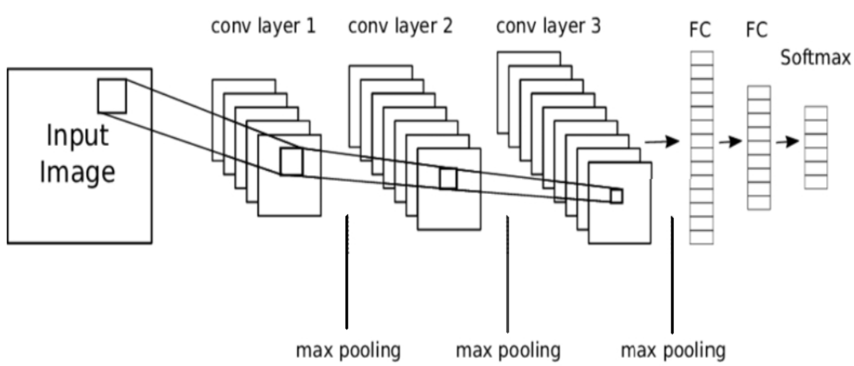


Fig 4.1.Convolutional Neural Network Architecture

Proposed Architecture

Firstly the data set is loaded and it is preprocessed by lessening the size and changing it over to the grayscale, at that point the highlights are separated utilizing the convolutional layer of CNN these highlights will help the model in training. Now the model is used to detect the disease.

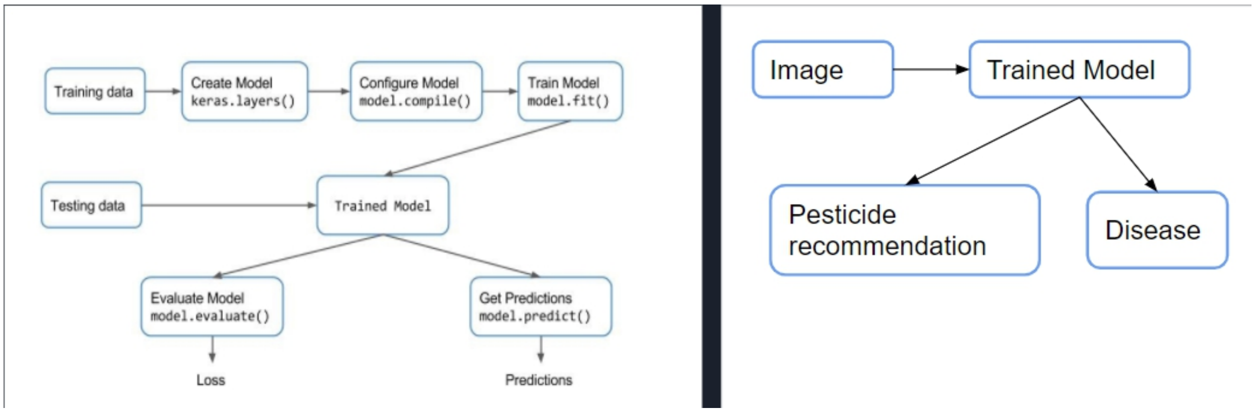


Fig 4.2. Proposed Architecture

In the testing phase, the same data preprocessing is done and the image data is uploaded to CNN algorithm then the given image parameters are compared with the training model parameters and then the model is able to detect the disease and provide the remedies. Some of the results obtained from our model are shown in figure 4.3.

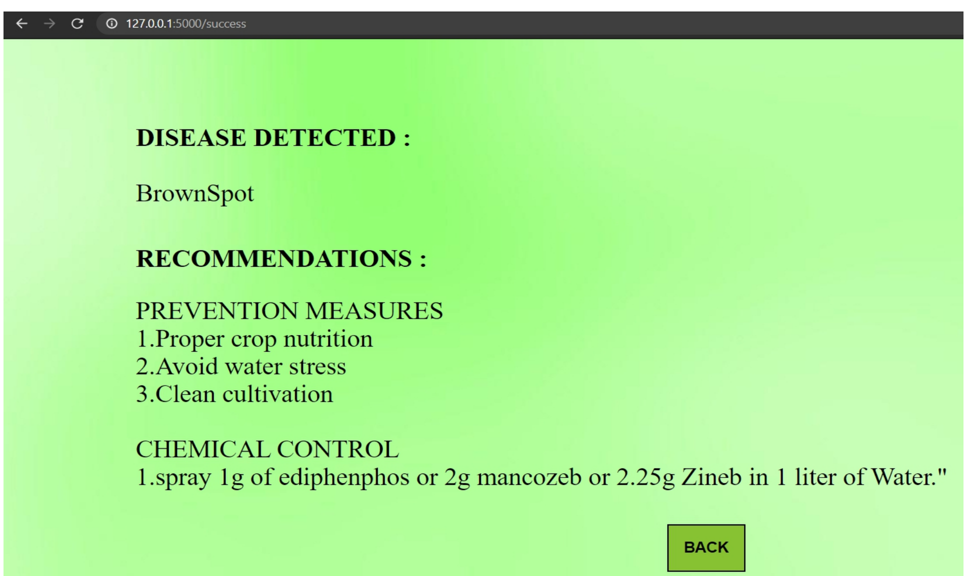


Fig 4.3. Results obtained from Rice Disease Prediction

[2]Rice Variety Classification

The collection of data sets is the primary process. This includes extraction of various parameters of individual rice grains like length,width,l/w ratio,shape,color. Rice varieties used in this project are BPT, Basmathi, Sonnamasuri, NLR,Nawara Rice,Brown Rice. Any rice sample that has been encountered in the system will be first classified and then will be segregated into its respective variety. This would be implemented by making use of different machine learning algorithms like Logistic Regression, Decision Trees, etc. These algorithms would give different recall values to the data set used. However, the algorithm with the maximum recall value would be chosen for further evaluation of the rice types.

The system is divided into smaller steps. The modules are as follows:
1) Capturing rice sample images

The first module involves capturing rice sample images with the help of a phone or tablet. The rice image is taken by keeping a one rupee coin as a reference for 2*2 dimensions.

2)Extracting relevant features

This module implements various image processing algorithms on the captured images using OpenCV in Python language. These features were then stored for further classification

3)Applying Machine Learning Algorithms

Machine Learning algorithms like Logistic Regression, Decision Tree Classifier, Naive Bayes and KNeighbours Algorithms were applied to the data set to get better classification results

4) Classifying rice grains based on their types

The results of the machine learning algorithms were then used to classify rice grains based on their types.

Rice grains are mainly classified concerning their texture, color, grain shape, etc.. The trained model identifies the different rice varieties and gives information about the nutrients available and their advantages.

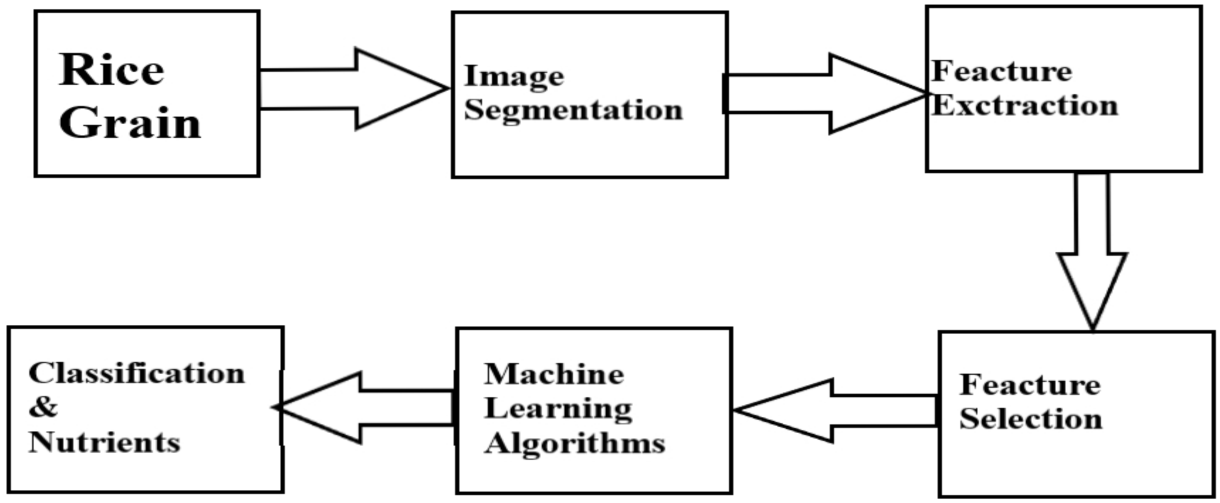


Fig 4.4 : Proposed Architecture of rice variety classification

A. Image Pre-processing

This filter is applied to remove noise which occurs during the acquisition of image and also sharpens the image. Threshold algorithm is utilized to segment the rice grains from the dark background.

B. Image Segmentation

Segmentation is applied to separate the contacting features of rice grains without losing the integrity of a single component. Dilation process follows the erosion process. The goal of dilation is to grow the eroded features to their original shape without rejoining the separated features. Edge detection is used to find out the boundaries of rice grains. We use a canny algorithm to detect the edges.

C. Feature Extraction

Features indicate length, width and color of rice grains. We implement various image processing algorithms on the captured images using OpenCV in Python language. These features were then stored for further classification.

D. Feature Selection

Classification requires all standard, measured and calculated results. The classification of rice grains as per the database is shown in the following tables.

| length | width | l/wratio | shape | color | rice |
|---|---|---|---|---|---|
| 8 | 2 | 4.76 | Slender | Shadow | Basmathi |
| 8 | 2 | 4.76 | Slender | Grullo | Basmathi |
| 8 | 2 | 4.16 | Slender | Shadow | Basmathi |
| 8 | 2 | 4.16 | Slender | Grullo | Basmathi |
| 8 | 2 | 4.16 | Slender | Camouflage Green | Basmathi |
| 8 | 2 | 4.16 | Slender | Beaver | Basmathi |
| 8 | 2 | 4.16 | Slender | Shadow | Basmathi |
| 8 | 2 | 4.16 | Slender | Grullo | Basmathi |
| 8 | 2 | 4.16 | Slender | Camouflage Green | Basmathi |
| 8 | 2 | 4.16 | Slender | Beaver | Basmathi |
| 8 | 2 | 4.76 | Slender | Camouflage Green | Basmathi |
| 8 | 2 | 4.93 | Slender | Cinereous | Basmathi |
| 8 | 2 | 4.93 | Slender | Laurel Green | Basmathi |
| 5 | 2 | 2.611 | Medium | Timberwolf | BPT |
| 4 | 2 | 2.44 | Medium | Silver | BPT |

Fig 4.5 : Dataset of rice varieties.

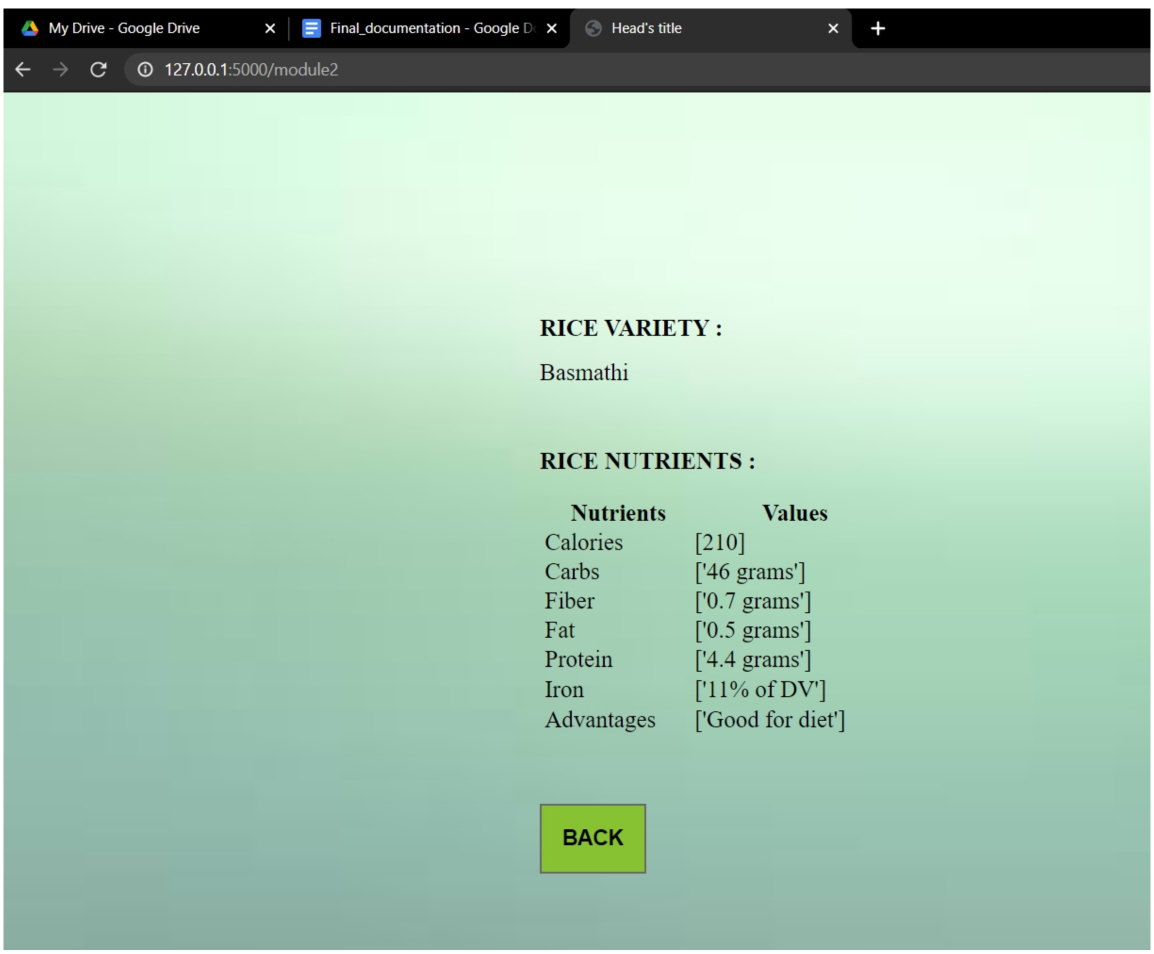


Fig 4.6 : Results of Rice Variety Prediction

# IMPLEMENTATION

[1] Rice Disease Prediction

Because of its capacity to remove great highlights, CNN has been utilized widely in AI and example acknowledgment research expressing that a multi-layer neural organization has incredible learning capacity, and that the learned highlights can extract and communicate crude information helpfully for order. CNN provides an end-to-end learning solution that avoids image pre-processing, and extracts relevant high-level features directly from raw images. The CNN architecture was inspired by the visual cortex of cats in Hubel and Wiesel's early work. In particular, Krizhevsky performed object classification and won first place in the ImageNet Large Scale Visual Recognition Challenge 2012 using a deep CNN. Since , similar CNN architectures have been successfully developed to solve a variety of image classification tasks. With full consideration of CNN's excellent performance, we propose a method that uses CNN for rice disease image feature extraction and disease classification.

STEP 1 : Import Libraries

```
# Import Libraries
import os
import glob
import matplotlib.pyplot as plt
import numpy as np
import tensorflow as tf
# Keras API
import keras
from keras.models import Sequential
from keras.layers import Dense,Dropout,Flatten
from keras.layers import Conv2D,MaxPooling2D,Activation,AveragePooling2D,BatchNormalization
from keras.preprocessing.image import ImageDataGenerator

tf.compat.v1.logging.set_verbosity(tf.compat.v1.logging.ERROR)
```

STEP 2 : Load train and test data into separate variables

# Loading train and test data into separate variables

train_dir ="C:\\Users\\balas\\Desktop\\rice_disease\\LabelledRice\\Labelled"

test_dir="C:\\Users\\balas\\Desktop\\rice_disease\\RiceDiseaseDataset\\validation"

STEP 3 : Pre-processing our raw data into usable format.

```python
# Pre-processing images with ImageDataGenerator function parameters
train_datagen=ImageDataGenerator(rescale=1./255,        #transform every pixel value from range [0,255] -> [0,1]
                                 shear_range=0.2,       #randomly applying shearing transformations
                                 zoom_range=0.2,        #randomly zooming inside pictures
                                 horizontal_flip=True)#randomly flipping half of the images horizontally
test_datagen=ImageDataGenerator(rescale=1./255)         #whatever preprocessing done to test images same is done to test images
```

```python
#Generating augmented data from train and test directories
img_width,img_height =256,256        #set height and width and color of input image
input_shape=(img_width,img_height,3)
batch_size =32                       #refers to the number of training examples utilized in one iteration
train_generator =train_datagen.flow_from_directory(train_dir, #Takes the path to directory,and generates batches of augmented data
                                                   target_size=(img_width,img_height),
                                                   batch_size=batch_size)
test_generator=test_datagen.flow_from_directory(test_dir,
                                                shuffle=True,
                                                target_size=(img_width,img_height),
                                                batch_size=batch_size)
```

```
Found 492 images belonging to 4 classes.
```

Fig 5.1 : Pre-processing image code

- Resizing image values between (0 –1) called normalization.
- Whatever pre-handling you do with the train it ought to be done to test parallelly.
- Every one of these boundaries are put away in the variable "train_datagen and test_datagen".

And then generate augmented data from train and test directories.

STEP 4 : Building CNN Model

```python
# CNN Model building
model = Sequential()
model.add(Conv2D(32, (5, 5),input_shape=input_shape,activation='relu'))
model.add(MaxPooling2D(pool_size=(3, 3)))
model.add(Conv2D(32, (3, 3),activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Conv2D(64, (3, 3),activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Flatten())
model.add(Dense(512,activation='relu'))
model.add(Dropout(0.25))
model.add(Dense(128,activation='relu'))
model.add(Dense(num_classes,activation='softmax'))
model.summary()
```

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 252, 252, 32)      2432
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 84, 84, 32)        0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 82, 82, 32)        9248
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 41, 41, 32)        0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 39, 39, 64)        18496
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 19, 19, 64)        0
_________________________________________________________________
flatten (Flatten)            (None, 23104)             0
_________________________________________________________________
dense (Dense)                (None, 512)               11829760
_________________________________________________________________
dropout (Dropout)            (None, 512)               0
_________________________________________________________________
dense_1 (Dense)              (None, 128)               65664
_________________________________________________________________
dense_2 (Dense)              (None, 4)                 516
```

Fig 5.2 : CNN Model code and summary

CNN shrinks the parameters and learns highlights and stores significant data yield shape, which is diminishing after each layer.

STEP 5 : Start Training CNN with Parameters and fit the model.

- Adam optimizer learning rate is 0.001
- Loss function categorical_crossentropy is used for our Multi-class classification problem and its metrics is “accuracy”
- Fit_generator is used to train the CNN model by using the validation data parameter for fine-tuning the model.

```python
#Start training CNN with parameters
#Generating validation augmented data from train directories
validation_generator = test_datagen.flow_from_directory(
                        test_dir, # same directory as training data
                        target_size=(img_height, img_width),
                        batch_size=batch_size)
opt=keras.optimizers.Adam(lr=0.001)
model.compile(optimizer=opt,loss='categorical_crossentropy',metrics=['accuracy'])
```

```python
train=model.fit(train_generator,steps_per_epoch=15, epochs=50, verbose=1, callbacks=None,
    validation_data=validation_generator, validation_steps=None, validation_freq=1,
    class_weight=None, max_queue_size=10, workers=1, use_multiprocessing=False,
    shuffle=True, initial_epoch=0
)
```

```
Epoch 1/50
15/15 [==============================] - 81s 5s/step - loss: 0.3557 - accuracy: 0.8739 - val_loss: 0.3315 - val_accuracy: 0.8720
Epoch 2/50
15/15 [==============================] - 79s 5s/step - loss: 0.3767 - accuracy: 0.8761 - val_loss: 0.3249 - val_accuracy: 0.8699
Epoch 3/50
15/15 [==============================] - 79s 5s/step - loss: 0.3497 - accuracy: 0.8630 - val_loss: 0.2694 - val_accuracy: 0.8963
Epoch 4/50
15/15 [==============================] - 79s 5s/step - loss: 0.3845 - accuracy: 0.8435 - val_loss: 0.7160 - val_accuracy: 0.7337
Epoch 5/50
15/15 [==============================] - 78s 5s/step - loss: 0.4111 - accuracy: 0.8478 - val_loss: 0.3379 - val_accuracy: 0.8801
Epoch 6/50
15/15 [==============================] - 79s 5s/step - loss: 0.3763 - accuracy: 0.8543 - val_loss: 0.4005 - val_accuracy: 0.8679
Epoch 7/50
15/15 [==============================] - 78s 5s/step - loss: 0.3694 - accuracy: 0.8783 - val_loss: 0.3090 - val_accuracy: 0.9004
Epoch 8/50
15/15 [==============================] - 77s 5s/step - loss: 0.2732 - accuracy: 0.9043 - val_loss: 0.2647 - val_accuracy: 0.9126
```

Fig 5.3 : CNN Model training

After training the model we got classification accuracy of 97.17 %

```python
print(max(train.history['accuracy']))
```

```
0.9717391133308411
```

Fig 5.4 : Training accuracy

STEP 6 : Saving Model weights

```
# Save model
from keras.models import load_model
model.save('crop.h5')
```

STEP 7 : Predictions

- we need to preprocess our image to predict
- First, we resize the image ===> image(256,256)
- convert image to array by this it will add channels===>image(256,256,3)
- Tensorflow works with batches of images.
- Predict_classes help to predict the image belonging to which class.

Confusion Matrix

```
[67]
from sklearn.metrics import confusion_matrix
import numpy as np
from sklearn.metrics import f1_score
from sklearn.metrics import recall_score
from sklearn.metrics import precision_score
cm = confusion_matrix(label1, predection1)
recall = recall_score(label1, predection1, average=None)
precision_score(label1, predection1, average=None)
f1 = f1_score(label1, predection1, average=None)
print(cm)
print("RECALL : ",np.mean(recall))
print("PRECISION : ",np.mean(precision))
print("F1 SCORE : ",np.mean(f1))

[[123   0   0   0]
 [  0 121   1   1]
 [  0   2 121   0]
 [  0   0   0 123]]
RECALL :  0.9918699186991871
PRECISION :  0.878921568627451
F1 SCORE :  0.9918615891108097
```

Fig 5.5 Confusion Matrix

[2] Rice Variety Classification

Image Segmentation and feature extraction

Image segmentation is a process by which we partition images into different regions. Contours are the continuous lines or curves that bounds the full boundary of an object in an image. We use an image segmentation technique called contours to extract the parts of an image.

Also contours are very much important in

- Object detection
- Shape analysis

And they have a very broad field of application from real world image analysis.

Implement contours in opencv, by extracting contours of squares.

1. Let’s load a simple image with rice images

```
def imagepath():
    global img_path
    p = input("Enter the image path : ")
imagepath()

def show_images(images):
    for i, img in enumerate(images):
        cv2.imshow("image_" + str(i), img)
    cv2.waitKey(0)
    cv2.destroyAllWindows()
```

2. Convert image into Grayscale

```
gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
blur = cv2.GaussianBlur(gray, (9, 9), 0)
```

3. Find canny edges

```
edged = cv2.Canny(blur, 50, 100)
edged = cv2.dilate(edged, None, iterations=1)
edged = cv2.erode(edged, None, iterations=1)
```

4. Finding contours

```
cnts = cv2.findContours(edged.copy(), cv2.RETR_EXTERNAL, cv2.CHAIN_APPROX_SIMPLE)
cnts = imutils.grab_contours(cnts)
```

5. Draw all contours

```
# Draw contours
for cnt in cnts:
    box = cv2.minAreaRect(cnt)
    box = cv2.boxPoints(box)
    box = np.array(box, dtype="int")
    box = perspective.order_points(box)
    (tl, tr, br, bl) = box
    cv2.drawContours(image, [box.astype("int")], -1, (0, 0, 255), 2)
```

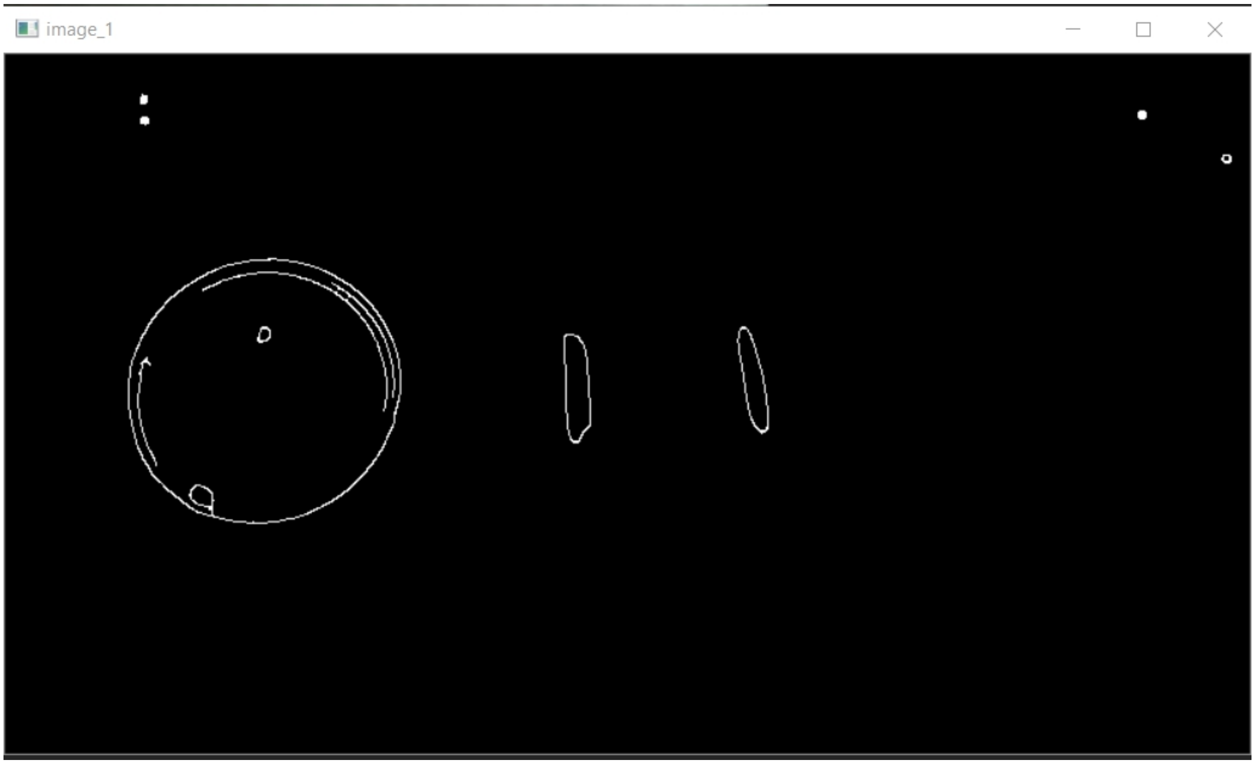


Fig 5.6 : Pre-processed image of rice grain

Algorithms used

A) Logistic Regression

Logistic Regression is an eminent machine learning algorithm which is beneficial for predicting probabilities involved in various classification challenges. It is nothing but a type of classification algorithm which assigns observations to a distinct and unique set of classes. Examples of classification problems include classifying high risk and low risk patients, good credit and bad credit, pass and fail, etc. Logistic Regression takes some input and transforms it into the required output in the form of probabilistic decisions. There are two main types of Logistic Regression: Binary and Multi Linear functions.

B) Decision Tree Classifiers

A decision tree classifier is a predictive model having applications in a number of areas. It consists of a graph consisting of nodes which represents attributes and asks questions. On the other hand, edges in a graph represent the answers to the questions asked earlier. Lastly, the leaves are used to denote the actual result obtained after following a path down the tree. This type of modeling tool can be implemented in non-linear structures to get the desired output. Every node in the tree represents a test case for a distinct attribute, whereas, every edge depicts some possible response. This flow is recursive and the same process is repeated for individual sub trees.

C) Naive Bayes Classifier

Naive Bayes classifier is a machine learning model that is used for classification tasks based on probability. The classifier is based on the Bayes theorem.

$$P(A|B) = P(B|A)P(A) / P(B)$$

Using Bayes theorem, we can find the probability of A's happening, given that B has occurred. Here, A is the hypothesis and B is proof. The assumption made here is that the indicators/highlights are autonomous. That is, the presence of one specific element doesn't influence the other. Consequently it is called naive.

| **Algorithm Used** | **Accuracy** |
|---|---|
| Logistic Regression | 60.86% |
| Decision Tree Classifier | 91.30% |
| Naive Bayes Classifier | 88.57% |

Table 5.1 : Comparison of different algorithms in rice variety classification

The system is successfully able to capture various parameters. The system is successfully able to efficiently extract features and parameters from the rice image and store them for further processing.

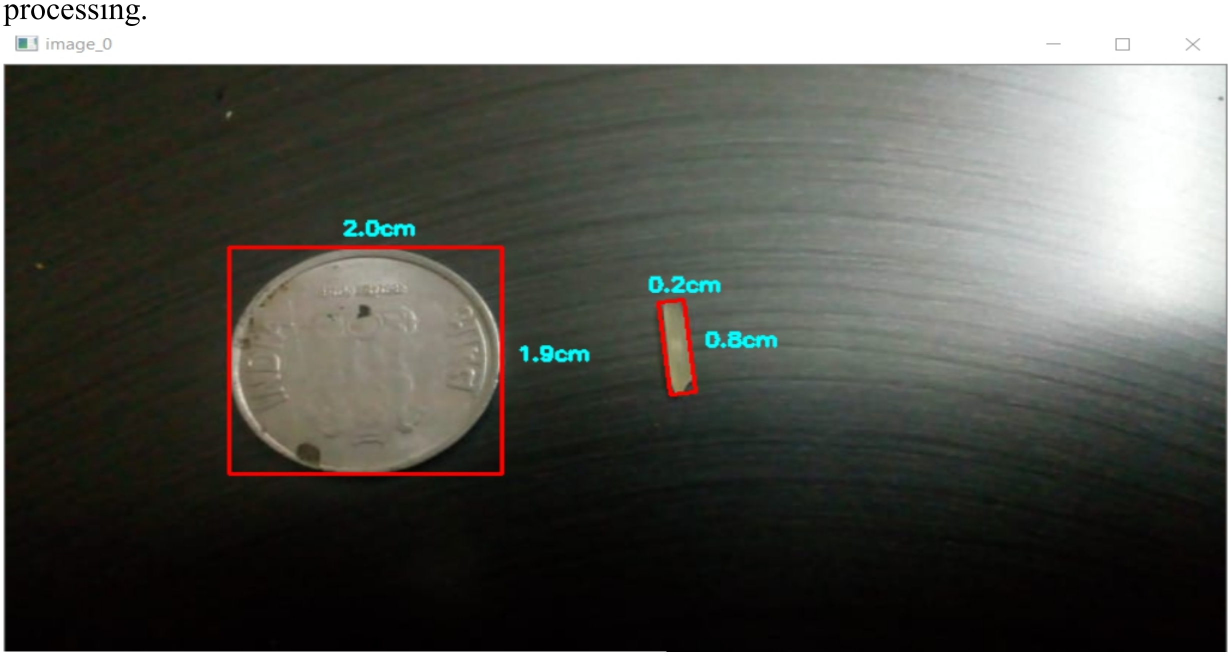


Fig 5.7 : Calculating length and width of rice image

This image highlights an individual rice grain and specifies its dimensions in millimeters. These dimensions are the two parameters namely length and breadth of rice grain.

```
print("RICE VARIETY DETAILS : ")
print("Length     : ",l)
print("Width      : ",w)
print("L/W Ratio : ",z)
print("Shape      : ",z1)
print("Color      : ",text)
```

```
RICE VARIETY DETAILS :
Length     :  8.0
Width      :  2.0
L/W Ratio :  4.16
Shape      :  Slender
Color      :  Shadow
```

Fig 5.8 : Features extracted after image processing

The above image is a sample of the features extracted by performing image processing on the Basmati Rice Images. These values are used for further classification.

```
filename = 'finalized_model.sav'
pickle.dump(classifier, open(filename, 'wb'))

# some time later...

# load the model from disk
loaded_model = pickle.load(open(filename, 'rb'))
#r = classifier.predict(sc.transform([[4.0,2,2.40,0.0,26.0]]))
r = classifier.predict(sc.transform([[float(l),w,float(round(z,2)),float(z1),float(text)]]))
r = np.array(r)
r = r.reshape(1,-1)
print(encoder.inverse_transform(r))
```

```
[['Basmathi']]
```

Fig 5.9 : Predicting rice variety code

# RESULTS

[1] Rice Disease Prediction

```
Detected Disease is  LeafBlast
Probability of this class is  98.20059537887573
PREVENTION MEASURES

 1.Remove the alternate host.

 2.Split application of nitrogen

CHEMICAL CONTROL

 1.Treat the seeds with Captan or Thiram or Carbendazim or Carboxin or Tricyclazole at 2 g/ kg.

 2.Spray nursery with Carbendazim 25 g.

 3.Main field with Edifenphos at 0.1% or Carbendazim at 0.1% or Tricyclazole.
```


Fig 6.1 : Predicted result of rice disease prediction

[2] Rice Variety Classification

[15]

```
dataset1[dataset1.rice == s]
```

| | Calories | Carbs | Fiber | Fat | Protein | Iron | Advantages | rice |
|---|---|---|---|---|---|---|---|---|
| 1 | 210 | 46 grams | 0.7 grams | 0.5 grams | 4.4 grams | 11% of DV | Good for diet | Basmathi |

Fig 6.2 : Predicted result of rice variety classification

# CONCLUSION

[1] Rice Disease Prediction

We used a batch size of 32 that is a hyper-parameter to adjust in machine learning. Data augmentation is also used to create artificially large training set from the existing dataset. We have used an image generator function which performs random rotations, shifts, flips, crops, and sheers on our image dataset. It helped us to achieve high results with a smaller dataset. In this model Our CONV layer has 32 filters with a 5 x 5 and 3 x 3 kernels and RELU activation (Rectified Linear Unit), a pooling layer of the size 3x3 and dense layer with activation function like RELU and Softmax. After that we have applied 25% (0.25) dropout layer. Visualization is done after every layer and ADAM optimizer is used with an epoch's value of 50 and achieved a classification accuracy of 97.17 % and r1 score of 99.18 %

[2]Rice Variety Classification

We focused on analyzing the rice variety based on visual features of rice grain images such as color, shape, texture. It can be applied to different classification models using these types of features. We observed that image processing techniques can combine with classification techniques such as Logistic Regression, Decision tree and Bayesian classifier to identify rice seeds. All the methods using simple features proved the best capability and accuracy of classification; on average it achieved 91.30% with Decision tree classifier respectively. The performance can be improved by using other types of features and further investigation of classification models.

# REFERENCES


[1 Rice Disease Prediction]

[1]I.S. Suraksha, B. Sushma, R.G. Sushma, K. Susmitha, S.V. Uday Shankar, “Disease Prediction of Paddy Crops using Data Mining and Image Processing Techniques”, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering,Vol. 5, Issue. 6, May 2016.

[2]R. Rajmohan, M. Pajany, R. Rajesh, D. Raghuraman, U. Prabu, “Crop Disease Identification using Deep Convolutional Neural Networks and SVM Classifier”, International Journal of Pure and Applied Mathematics Vol. 118, No. 15(2018)

[3]L. Barik, “Survey on Region Identification of Rice Disease Using Image Processing”, International Journal of Research and Scientific Innovation Vol. 5 Issue 1(2018)

[4]A Badage. “Crop Disease Detection using Machine Learning: Indian Agriculture”, International Research Journal of Engineering and Technology Vol. 5 Issue. 9 (2018)

[5]K. Jagan Mohan, M Balasubramanian, S Palanivel , “Detection and Recognition of Diseases from Paddy Plant Leaf”, International Journal of Computer Applications Vol. 144 No. 12(2016).

[6]S.B Dhaygude, N.P Kumbhar, “Agricultural Plant Leaf DiseaseDetection Using Image Processing”, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering,Vol. 2 Issue 1(2013).

[7]G Jayanthi,., K.S Archana., A Saritha, “Analysis of Automatic Rice Disease Classification Using Image Processing Techniques”,International Journal of Engineering and Advanced Technology

(IJEAT), Vol-8 Issue-3S pp. 15

[2 Rice Variety Classification]

[1] Machine Vision based Quality Analysis of Rice Grains VidyaPatil, V. S. Malemath, QualityAnalysis and Grading of Rice Grain Images, International Journal of Innovative Research in Computer & Communication Engineering (IJIRCCE) ISSN 2320-9801, Vol. 3,Issue 6, pp5672-5678, June 2015.
[2] Grain Quality Detection by using Image Processing for public distributionQing Yao, Jianhua Chen, Zexin Guan, Chengxiao Sun, Zhiwei Zhu“Inspection of rice appearance quality using machine vision”, Global Congress on Intelligent Systems, 978-0-7695-3571-5/09, IEEE 2009.

[3] An Automatic Method for Identifying Different Variety of Rice Seeds Using Machine Vision Technology P. Shatadal, “An identifying damaged soybeans by color image analysis,” Applied Engineering in Agriculture, Vol. 19, PP. 65-69, 2003.

[4] Non-Destructive Quality Analysis Of Kamod Oryza Sativa Ssp Indica (Indian Rice) Using Machine Learning Technique Abdullah MZ, Fathinul-Syahir AS,MohdAzemi BMN,“Automated inspection system for color and shape grading of star fruit (Averrhoacarambola L.) using machine vision sensor,” Transactions of the Institute of Measurement and Control, 27 (2), 65-87,2005.

[5] Improved K-Means Algorithm Based on Hybrid Rice Optimization Algorithm T. Kanungo, D. M. Mount, N. S. Netanyahu, et al., “An efficient K-means clustering algorithm: analysis and implementation,” IEEE Transactions on Pattern Analysis & Machine Intelligence,vol. 24, issue 7, pp. 881-892, 2002.

[6] A Hybrid Rice Optimization Algorithm Zhiwei Ye, Lie ma, Hongwei Chen, The 11th international conference on Computer science and education(ICCSE) PP. 169-174, 2016.

[7] Rice Grains Classification Using Image Processing Technics Herath H.M.K.K.M.B and Eng.de Mel W.R Department of Mechanical Engineering,The open university of Sri Lanka, PP.1-6, 2016.

[8] An Efficient Method For Quality Analysis Of Rice Using Machine Vision System, Rahul Birla, Ajay Pal Singh Chauhan, Electronics and communication engineering Department, Journal of advances in Information technology,Volume 6,No.3, PP.140-145, 2015.

[9] Identification Of Food Grains And Its Quality Using Pattern Classification, Sanjivani Shantaiya, Mrs.Uzma Ansari, Special Issue of IJCCT 2010 for International Conference[ICCT-2010], Volume 2,Issue 2,3,4;pp.70-74, 2010.

[10] Classification Of Rice Grains Using Neural Networks, Chathurika Sewwandi Silva, Upul Sonnadara, Proceedings of Technical Sessions, Volume 29, PP. 9-14, 2013.