Grenze International Journal of Engineering and Technology, Jan Issue

Recognition of the Varieties of Rice using CNN

Hritika Jadhav[1], Rahul Sanap[2], Anuradha Kotgire[3], Sanchi Kamble[4] and Prof. Geetanjali Mate[5]
[1-5]JSPM'S Rajarshi Shahu College of Engineering, Department of Information Technology

***Abstract*—Rice is a high value crop that feeds more a lot of people around the world. Rice is a grain product rich in carbohydrates and proteins. It is important as the top 5 countries exporting rice had a network export around more than $19 billion in 2021.It is one of the staple food in most continents. Powerful rice grain analysis and classification system can efficiently improve both correctness and time. As there are many varieties of rice grains available in the market thus classification becomes an important and crucial. Classification is important as it has an influence on human health as we consume it. Classifying rice grains one by one manually is a monotonous task and doing it continuously can reduce work efficiency.To make work easy we need to automate the process. The automated system should be proper and intelligent so that this problem can be solved easily. System must be able to recognize and sort each rice grain with the respective species. It should be able to automatically identify and classify individual rice grains according to the respective species. Rice quality is determined by combination of physical and chemical properties. Physical quality of rice grain depends upon color, shape, and size of grain. Rice samples are collected and then the features are extracted. Various machine learning techniques are used for classification and quality determination.**

***Index Terms*— Machine Learning, CNN, Deep-Learning, Image processing, SVM, etc.**

I. INTRODUCTION

Agriculture is one of the earliest and most extensive and main industry since the beginning of mankind. Quality of rice grain is determined by morphological features by human beings. Neural networks, digital image processing, etc are the various methods used for classification of rice grains. Various image processing techniques are used to identify and study about rice grains. Manual classification techniques are inadequate and time consuming. Quality analysis is very crucial in terms of agriculture and agricultural industry. An expert in the field can visually analyze the rice grain. But this review is not authentic as it has chances of containing human error. Visual analysis done by a person can take a lot of time, which is not feasible. To get better of results and overcome the limitations of traditional methods, various image processing techniques are designed.Classification and quality evaluation is done by making an analogy between the test image and database image. In the first step we convert the RGB image into the binary image and greyscale image. Edge detection is done by using Canny edge detection Classification is done by using the algorithms like SVM, CNN and multilayer perceptron. The planned system works accurately with low cost and time.

II. REVIEW OF LITERATURE

A. Material and Methods

The rice classification system is based on computer vision image processing and machine learning algorithms.

Grenze ID: 01.GIJET.9.1.546_2

Above system classifies the rice grains according to the size and shape. The first step is pre-processing that includes entering of the input images for the detection of rice grains. Then the images are converted into grey scale and binary images. The noise and error present in the background is removed. These images are then used for training processes in various machine learning and deep learning algorithms.

B. Data Acquisition

The images are obtained by a mechanism showed in the above figure. Here we have a light source and a black background so that the shadow of the grains is not visible in the images.The background is selected black so that it doesn't take light from outside and hence is easier for processing the images. The captured images are then transferred into the database for further operations.

C. Data pre-processing

The images are acquired and are then processed further. The input image is a color image which is converted into a greyscale image by applying filter and then these images are converted into binary images. The noise in the image is removed and the image is further smoothen and sharpen. [5][12][19]

Morphological Features

Features which are relates to the external structure of the grain is known as morphological features. The features are as follows-

No	Feature	Definition
1.	Area	Number of pixels consisting within the edges of rice grain.
2.	Perimeter	Distance between the pixels at the edges of the rice grain.
3.	Major axis length	The line that joins the two farthest points present on the edge of the rice grain.
4.	Minor axis length	The line that joins the two nearest points on the edge of the rice grain.
5.	Aspect ratio	The ratio between the major and minor axis.

Texture Features

The low-level feature present in the image that is important to describe the contents of the image is called texture features. The features are as follows

No.	Feature	Definition
1.	Contrast	The difference in properties that makes the object more identifiable.
2.	Energy	The change in the quality of the image.
3.	Correlation	The compactness between the two objects in the image.
4.	Homogeneity	The change in the intensity of the regions in the image.

Color Features

The image is characterized by the intensity of colors present in the image.

No.	Feature	Definition
1.	HSV conversion	The parameters here are Hue, Saturation and Value.
2	YCbCr conversion	The parameters here are Y stands for brightness, Cb stands for blue difference and Cr stands for red difference of components.
3.	XYZ conversion	The parameters here are X stands for red, Y stands for brightness and Z stands for blue.

D. Feature Selection

The selection of relevant data from the original data to get rid of noise and to reduce input to the model is called Feature selection. Feature selection techniques are divided into three classes as supervised, unsupervised, and semi supervised.Supervised feature selection uses the output label class for feature selection whereas semi-supervised does not use the label classes. Further the supervised feature selection is classified into three methods.The supervised methods are filter method, wrapper method and intrinsic method. The unsupervised feature selection method is feature clustering. Principle component analysis and Linear discriminant analysis are commonly used for extraction of features.[17]

II. Performance Measures

These are the parameters that help us to talk about the success of the model. It tells us about the performance and the accuracy of the model. One of the ways to calculate the accuracy of our model is by using complexity matrix. The matrix tells us the details about real classes and the predicted classes after the classification performed after the model.

No.	Measure	Definition
1.	Accuracy	Regularity of the model to predict accurate output. It is ratio between number of correct predictions to all predictions.
2.	Precision	The prediction by the model is accurate and matches the real classes.
3.	Recall	Out of total positive class how many were correctly classified by our model.
4.	Specificity	Ability of the model to predict that the observation does not belong to a specific class.

Reference	Number of Varieties	Features extracted	Pre processing techniques	Classifier used	Performance	Training/Testing data	Year
[11]	7	Morphologica	Greyscale conversion	NN	-	Training-48 Testing-8	2019
[13]	3	Morphologica And color	Greyscale and Binary	SVM	92.22%	Testing-90	2019
[14]	13	Morphologica And texture	Otsu Method	Back-propaga ANN	98%	Training-3600	2019
[15]	5	Color and sha	Thresholding	M-SVM With ANN	97.67%	-	2019
[5]	4	Morphologica	-	SVM, LR, Random fores	93%-SVM 97%-Logistic regressior	-	2020
[1]	5	12 Morpholog 90 color	Otsu Method	k-NN, SVM Decision Tree LR	k-NN-97.15, SVM-97.02 LR-97.83	75000-total	2022
[10]	6	214 features	Color spaces			Training-5225 Testing-5253	2019
[12]	90	Morphologica	-	k-NN SVM LQV	Average-98.55%	Total-8640	2019
[16]	9	6 color 2 texture 4 shape	Thresholding method		Increased accuracy by 100%	Training-630 Testing-270	2019
[3]	15	Morphologica	RGB conversion	ANN with k-NN	79%	-	2021
[9]	20	Size of grain	Background subtraction	CNN with SVM & k-NN	89.75%-CNN 85.06%-SVM, 84.30%-	Total-2000	2019
[7]	4	Morphologica	Binary Conversion	SVM, CNN	-	Total-42	2020
[8]	2	7 morphologic	Binary Conversion	LR, SVM, DT,RF	92.02%-LR, 92.83%-SV 92.49%-DT	Total=3810	2019

III. Survey on Current Publications

Reference	Number of Varieties	Features extracted	Pre processing techniques	Classifier used	Performance	Training/Testing data	Year
[11]	7	Morphologica	Greyscale conversion	NN	-	Training-48 Testing-8	2019
[13]	3	Morphologica And color	Greyscale and Binary	SVM	92.22%	Testing-90	2019
[14]	13	Morphologica And texture	Otsu Method	Back-propaga ANN	98%	Training-3600	2019
[15]	5	Color and sha	Thresholding	M-SVM With ANN	97.67%	-	2019
[5]	4	Morphologica	-	SVM, LR, Random fores	93%-SVM 97%-Logistic regressior	-	2020
[1]	5	12 Morpholog 90 color	Otsu Method	k-NN, SVM Decision Tree LR	k-NN-97.15, SVM-97.02 LR-97.83	75000-total	2022

[10]	6	214 features	Color spaces			Training-5225 Testing-5253	2019
[12]	90	Morphological	-	k-NN SVM	Average-98.55%	Total-8640	2019
[16]	9	6 color 2 texture 4 shape	Thresholding method	LQV	Increased accuracy by 100%	Training-630 Testing-270	2019
[3]	15	Morphological	RGB conversion	ANN with k-NN	79%	-	2021
[9]	20	Size of grain	Background subtraction	CNN with SVM & k-NN	89.75%-CNN 85.06%-SVM, 84.30%-kNN	Total-2000	2019
[7]	4	Morphological	Binary Conversion	SVM, CNN	-	Total-42	2020
[8]	2	7 morphological	Binary Conversion	LR, SVM, DT,RF	92.02%-LR, 92.83%-SVM 92.49%-DT	Total=3810	2019

IV. MOTIVATION

To segment and identify rice grains image processing algorithms are used. Systematic use of image processing algorithm is a method to analyze grains quality by its morphological features. Edge detection algorithm is used to find out boundaries of each grain. End points of each gains are found and then are measured by using calliper to get estimated length and breadth of rice..The traditional methods used for grain shape and size measurement are grain shape tester, dial micrometer and graphical method, but these are very long and require more amount of time and hard work. At a single point of time only length and breadth of a single grain can be measured. The end point of this method is expensive and takes more time and has chances of human error. However, the morphological features of the different varieties are diversified to specify a universal formula for the categorization of all types of rice.

V. PROBLEM DOMAIN

Machine Learning :- An evolving technology called machine learning makes it easier for computers to adapt naturally from the data. Machine learning uses a various algorithms to build mathematical models and make predictions based on already available information. At present , it is used for various things, including recommender systems, email filtering, Facebook auto-tagging, image recognition, and speech recognition.

VI. PROBLEM STATEMENT

To recognize the varieties of rice grains by the use of Machine and Deep learning Techniques.The objective of the project is to recognize the rice grains to their respective varieties and give the count of the grains present in the input image. Future work is under consideration to improve the identification of rice grain to purity measurement at different condition with the help of different algorithms.

VII. SOLUTION METHODOLOGY

A. Training Models

1 Artificial Neural Network (ANN)

Neural network comprises of many neurons which are inter connected to each other and is used for finding the solution for a problem. Neural network consists of three layers that is input, hidden layer and output layer. The input layer would consist number of neurons depending upon the characteristics of the input dataset. .[14][15][18]

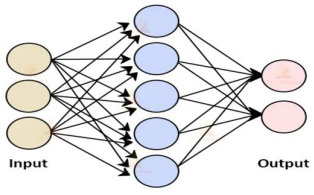


Fig.1: Artificial Neural Network (ANN)

2 Convolution Neural Network

The network is designed in such a way that the model can classify the rice grains efficiently. Here the training time is huge and depends upon the number of layers present in the network of the system requirement the network training parameters are adjusted so that the model would work properly in less amount of time thus increasing the training accuracy of the model. [4][7][9]

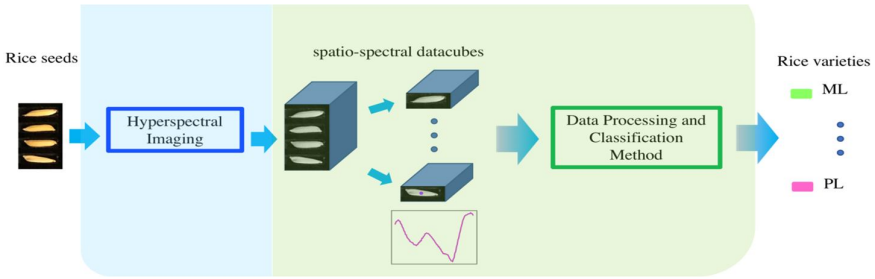


Fig.2: CNN

VIII. Classification Model

A. Support Vector Machine (SVM)

Most classification problems are solved using SVM. It can also be used in regression problems. Here we take each datapoint and plot it in n-dimension plane. SVM classifies the data into linear, planar, and hyper plane in 2-D, 3-D and multi-dimensional plane respectively. The hyperplane helps us to classify the datapoints into different classes.[8][7][9][18][5][13]

B. Logistic Regression

This is one of the most used and important machine learning algorithms. It is most helpful in speculate probabilities in classification problems. Predicted values are bounded between 0 and 1 as the algorithm gives the probability instead of results. There are two types of logistic regression that is Binary and Multi-linear regression.[8][5]

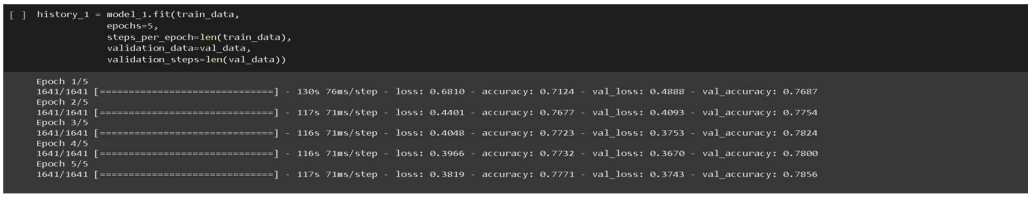
C. K- nearest neighbor (k-NN)

This is a type of non-parametric learning algorithm. Euclidean distance is used to classify the dataset. Here k stands for number of neighbors. This algorithm is best used where the data set class is not known .This algorithm is more accurate than the others and is more flexible to the input dataset.[9][3][12][1]

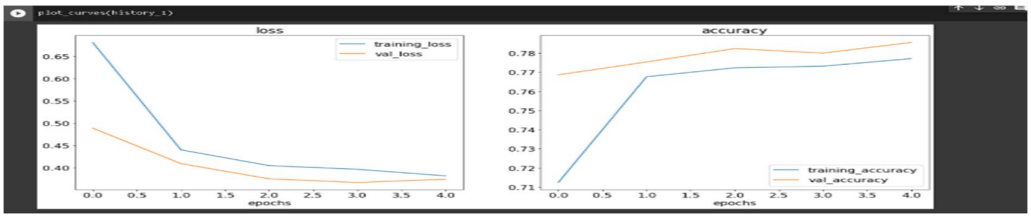
IX. Implementation

Dataset--https://www.muratkoklu.com/datasets/

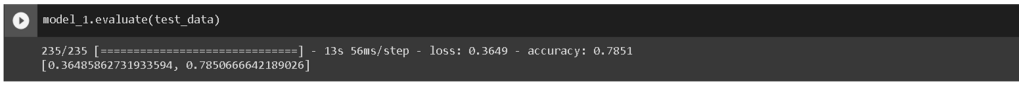
Model Definition



Model Training



Loss/Training Accuracy



Result

X. CONCLUSION

Above implementation and algorithms approaches to the successful classification of rice grains into corresponding varieties. The implementation of the process starts with data collection that is the images of the rice grain are collected. The system was able to identify different type of features that are morphological color and texture. Here the Accuracy of the partial implemented system is 75%. We can see further after implementing that the accuracy of the models are good when we do a combination of CNN and SVM. Studying other Deep learning techniques would help more to achieve extra accuracy with the dataset having more variety of rice grains.

FUTURE WORK

It can be applied to commercial grading of rice is done according to the size of the grain (full, half or broken).It can be used to obtain the percentage quality of rice grains from physical and chemical characteristics. Used by farmers to check the quality of their goods so to demand a proper rate for their goods while selling them in market.

REFERENCES

[1] Identification of Rice Varieties Using Machine Learning Algorithms Ilkay CINAR ,Murat KOKLU(2022)
[2] On the Frontiers of Rice Grain Analysis, Classification and Quality Grading: A Review. SHEIKH BILAL AHMED, SYED FAROOQ ALI , AND AADIL ZIA KHAN(2021).
[3] Computer Vision and Machine Learning Analysis of Commercial Rice Grains : A potential Digital Approach for Consumer Perception Studies Aimi Aznan,Claudia Gonzalez Viejo, Alexis Pang,Sigfredo(2021)
[4] Localization and Classification of Rice-grain Images Using Region Proposals-based Convolutional Neural Network Kittinun Aukkapinyo Suchakree Sawangwong, Parintorn Pooyoi Worapan Kusakunniran, (2020)
[5] Rice Grain Classification using Image Processing & Machine Learning Techniques. Biren Arora, Nimisha Bhagat, Sonali Arcot, Saritha L R.(2020)
[6] A cascade network for the classification of rice grain based on single rice kernel Ksh Robert Singh, Saurabh Chaudhury(2020)
[7] Classification of Rice Grains using Machine Learning Techniques. Srinivasa Rao, Murali Krishna Talari, Rajendra Kumar Pedapudi, Sai Charan Akula(2020)
[8] Classification of Rice Varieties Using Artificial Intelligence Methods. Ilkay CINAR1, Murat KOKLU(2019)
[9] Deep Learning for Rice Quality Classification Nguyen Hong Son and Nguyen Thai-Nghe(2019).
[10] Dimensionality Reduction Based on Feature Selection for Rice Varieties Recognition Huu-Thanh Duong, Vinh Truong Hoang(2019)
[11] Survey on Classification of Rice Grains Using Neural Network A.H. Bhensjaliya1 H.D.Vasava (2019).
[12] Varietal Classification of Rice Seeds Using RGB and Hyperspectral Images SAMSON DAMILOLA FABIYI, HAI VU, CHRISTOS TACHTATZIS 1,PAUL MURRAY 1, DAVID HARLE , TRUNG KIEN DAO 2,IVAN ANDONOVIC , JINCHANG REN 1,AND STEPHEN MARSHALL.(2019)
[13] Rice grain classification using multi-class support vector machine (SVM) Shafaf Ibrahim, Nurul Amirah Zulkifli, Nurbaity Sabri, Anis Amilah Shari, Mohd Rahmat Mohd Noordin(2019)
[14] A Combined Method of Image Processing and Artificial Neural Network for the Identification of 13 Iranian Rice Cultivars Yousef Abbaspour-Gilandeh , Amir Molaee 1, Sajad Sabzi , Narjes Nabipur , ,Shahaboddin Shamshirband, and Amir Mosavi (2019)

[15] An Efficient Rice Variety Identification Scheme Using Shape, Harlick & Color Feature Extraction and Multiclass SVM M.Senthil Kumar, Md.Javeed(2019)
[16] Identification of Rice Variety Using Geometric Features and Neural Network Wahyu Srimulyani, Aina Musdholifah(2019)
[17] Rice Grain Identification and Quality Analysis using Image Processing based on Principal Component Analysis Muhammad Junaid Asif , Tayyab Shahbaz, Dr. Syed Tahir Hussain Rizvi, Sajid Iqbal(2018).
[18] Classification of Milled Rice Using Image Processing Rexce J, and Usha Kingsly Devi K,(2017)
[19] Image analysis based on color, shape and texture for rice seed (Oryza sativa L.) germination evaluation Benjamaporn Lurstwut, Chomtip Pornpanomchai*(2017).
[20] Assessment of Quality of Rice Grain using Optical and Image Processing Technique. Engr. Zahida Parveen, Dr. Muhammad Anzar Alam, Engr. Hina Shakir(20)