Analysing Rice Seed Quality Using Machine Learning Algorithms

Dr. T. Avudaiappan[#1], S.Sangamithra*[2], A.Silpha roselin *[3], S.Sherin farhana *[4], KM.Visalakshi*[5]
[#1]*Assistant professor, Department of Computer Science, K.Ramakrishnan College of Technology, Samayapuram, Trichy-600013.*

Abstract

Rice is the real sustenance source in Southern India. It is the staple food for more than 80% of people around the globe. Many varieties of paddy crop are cultivated and exported. Detecting the defected grains and distinguishing rice variety is crucial in rice quality analysis. An automated system is introduced which can be used for rice grain type identification and classification where digital imaging is recognized as an efficient technique to extract the features from rice grains in a non-contact manner. Images are captured by using a camera. Image Pre-processing techniques, Filtering, Segmentation, and edge detection are performed on the acquired image. The morphological features that are extracted from the image are given to Machine Learning Algorithm and the output is displayed in LED.

I. INTRODUCTION

Analyzing the quality of rice is one of the vital roles in machine vision. Several researchers suggest that object shape is more information than its appearance properties such is more color vary between objects instances more than the shape. But it cannot give an accurate result. It can also identify the rice integrity problem. Rice integrity means touching of seeds while taking samples. The main purpose of this method is to offer an alternative way for quality control and to analyze the quality of rice which reduce the required effort, cost and time. The rice quality plays an important role in the production of agronomic and horticultural crops so identification of the quality of rice is very important. Image processing is a significant and advanced technological area where important developments have been made in tradition the farmers know about the rice seed quality by analyzing through the naked eye. They also have the experience to know about the various varieties. In the recent year's efforts to develop Machine vision system (MVS) for industrial application has been increased considerably owing to Availability of the low-cost electronic instruments and processing hardware. The advancement in computer technologies using digital image processing applications showed the way for analyzing the quality of food material. Rice is the major crop needed for the developing country. An automated machine that extracts information of rice using software system is more speedy, accurate, convenient, harmless and non-destructive and hence the result obtained is more precise. An algorithm can be developed to analyze the quality of the result along with the accuracy from the given samples.

II. LITERATURE SURVERY

The quality of rice is based on the two characteristics physical and chemical for evaluation and grading of rice grains. The quality assessment is done by finding the region of the boundary and end points by its measures. The average value of features is considered and they are implemented in Mat Lab [1].The image processing algorithm is applied over the sample grains through Mat Lab. The classification of the quality of rice is based on its color, size and shape. The accuracy and result are divided into good, bad and medium quality by Neural Network Classifiers [2]. An automated method for identifying different variety of rice seeds using machine vision technology and detection system that consist of inspector machine and an image processing unit. A back forward Neural Network was trained to Identify rice seed quality. From this The system was enough to provide Inspection of varieties of different rice seeds based on their appearance characteristics of seeds [3]. The solution for quality of evaluation and grading of Krishna kamod rice using image processing and soft computing techniques. The Feed Forward Neural Network techniques that are applied to find the high degree of quality. The long seeds and small seeds as well as unknown seed quality are found by the Trained Multilayer Feed Forward Neural Networks [4]. Clustering analysis that is mainly used in the area of data mining. As that K-means algorithm that mainly depends upon initial clustering center and its local optimum. The forward is used first by applying the selected initial cluster to the end. The hybrid optimization algorithm and clustering algorithm is used to find the optimal cluster center [5].The optimization algorithm is inspired by the breeding process of the three line hybrid rice. Where sterile , maintainer, and restorer line. The hybridization is evolutionary Process and self process is swarm search process. They are combined each other which appropriate proportion with respect to convergence and speed calculation that are involved as a whole [6]. The machine vision

system is used for the grain classification. The rice can be distinguished and analyzed using the different way according to its RGB color model, histogram, and edge detection. It helps to find the percentage of purity of rice grains based on its features such as shape, color[7]. The solution for quality assessment and grading of Indian Basmati Oryza Sativa(L) variety rice using machine vision and Image processing. It calculates the size of

Oryza Sativa (L) also with detection of chalky and broken rice with improved accuracy compared with human inspectors [8]. The normalization of food grains and eliminating the effect of orientation is done by using Image warping technique with proper scaling. The back propagation neural network based classification is developed to identify the unknown grain types. The color and textural features were Presented to the neural networks for training purpose that identifies the unknown grain type[9].The different models were developed for both the individual feature set and for the combined feature sets. The classification accuracy is given by textural features than morphological and color features. Thus Neural Network architecture tends to produce different accuracies for different feature sets.

III. SIZE AND APPEARANCE

CHARACTERISTICS FEATURES	KAR RICE	KURUVAI RICE	NAVARAI	LATE THALADI
Variety	Adt-43	Adt-45	Adt-36	Adt-36
Grain lib rate	2.18	2.98	3.1	3.1
Grain type	Medium slender	Medium slender	Medium	Medium
Habit	Semi dwarf	Semi dwarf , erect	Erect	Erect
Length	5.46	8.00	7.8	7.8
Breadth	1.94	2.16	2.5	2.5
Thickness	1.63	1.97	2	2

Table1: Rice characteristics

IV. ARCHITECTURAL DIAGRAM

Existing System

In 2017, an automatic evaluation method for the determination of the quality of rice granules. The quality of rice samples is to be determined with the help of geometric features. A model of quality grade testing and identification is built which is based on appearance features such as area, major axis length, minor axis length, aspect ratio, the morphological and color with technology of computer image processing and neural network. The color and morphological features are presented to the neural network for training purposes. The trained network is then used to identify the unknown grain types and its quality. This proposed method gives good results in evaluation of rice quality. Machine learning is part of a broader family of learning methods based on learning data representation, as opposed to task-specific algorithms. Learning can be supervised, supervised or unsupervised. However, this approach introduces a number of privacy and efficiency challenges, as the cloud operator can perform secondary inferences on the available data. Recently, advances in edge processing have paved the way for more efficient, and private, data processing at the source for simple tasks and lighter models, though they remain a challenge for larger and more complicated models.

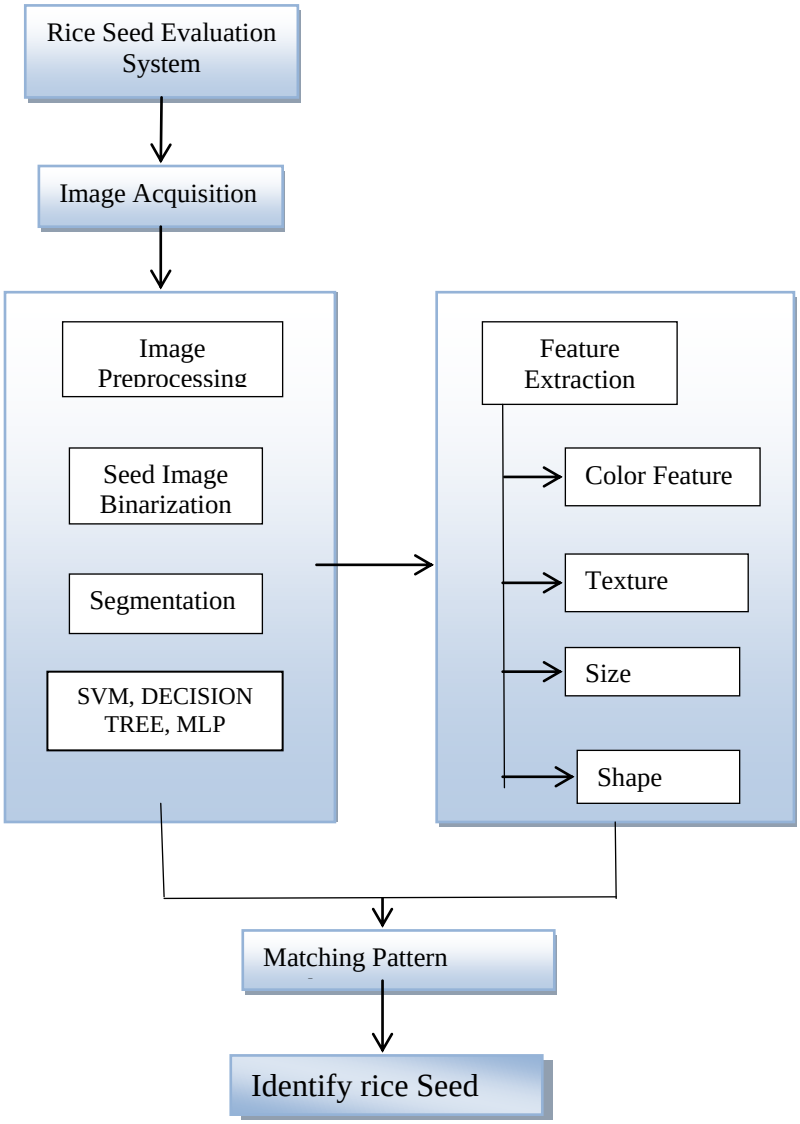


Fig:1.1

V. ALGORITHMS

Multi Layer Perceptron:

Artificial Neural Network (ANN) types and it is used for classification purpose. It is one type of Supervised Learning. MLP consists of three layers: an input layer, a hidden layer, and an output layer. It uses the Back Propagation algorithm for classification of milled rice samples. It is also known as a feed-forward method because in this all the information is passed only in the forward directions through the nodes. It calculates neurons weight by its activation function that means a linear function calculates the weight of the inputs to the output. In MLP some neurons may be nonlinear. The activation function consists of two common functions. The first one is hyperbolic tangent ranges from -1 to 1 and another one is logistic function ranges from 0 to 1. Each node in one layer connects with the weight of the other nodes within that layer. The input layers are visible layers. Each input is taken as a neuron. In this, the inputs are passed through the next layer. After the input layer, a hidden layer is placed. Hidden layers are known as invisible layer. The input passed through the hidden layer and the hidden layer provides output value directly to the output layer.

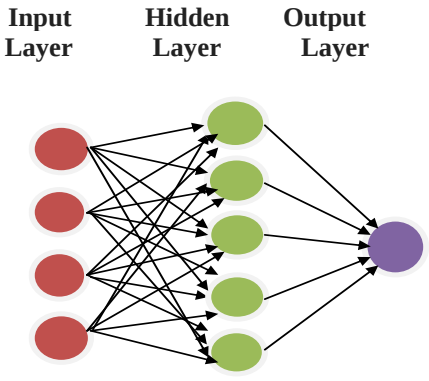


Fig: 1.2

Support Vector Machine:

Support vector machine is one of the supervised learning models. That is used for twopurposes they are classificationand regression. But mostly used in classification problems. Here the sample is represented as points.The optimal boundary is also known as hyper-plane. And the hyper-plane for two sets in a vector space are obtained independently that is based on the probabilistic distribution of training vectors in the set. Hyper-plane locates boundary which is far away from the nearest vectors to the boundary in two sets. The vectors located near the boundary are called support vectors. Suppose if space is not linear at that point hyper-plane can't be used to distinguish. The kernel function is used to solve those problems. Kernel trick is one of the approaches to solve nonlinear solvable problems. This approach is based on the inner product of input data and with the definition of the suitable kernel function. The kernel function enables to perform the operations in input space rather than high dimensional feature space. kernel functions are known as a class of algorithms used for pattern analysis. Four kernel functions are there. They are polynomial, normalized polynomial, RBF and universal Pearson VII. The polynomialkernel function is used to represent the similarity of vectors in feature space. The normalized polynomial gives better results when compared to polynomial kernel.RBF stands for radial basis function. It is used in various kernelized learning algorithms. Universal Pearson VII is capable to serve as a generic alternative to the common linear, polynomial and RBF kernel functions.

VI. WORKING MODEL

Computing mlp:

A multilayer perceptron (MLP) is a class of feed forward artificial neural network. A MLP consists of, at least, three layers of nodes: an input layer, a hidden layer and an output layer. Except for the input nodes, each node is a neuron that uses a nonlinear activation function. MLP utilizes a supervised learning technique called back propagation for training. Its multiple layers and non-linear activation distinguish MLP from a linear perceptron. It can distinguish data that is not linearly separable.

1)Activation function

If a multilayer perceptron has a linear activation function in all neurons, that is, a linear function that maps the weighted inputs to the output of each neuron, then linear algebra shows that any number of layers can be reduced to a two-layer input-output model. In MLPs some neurons use a nonlinear activation function that was developed to model the frequency of action potentials, or firing, of biological neurons.

The two common activation functions are both sigmoid, and are described by

$$y(v_i) = \tanh(v_i) \text{ and } y(v_i) = (1 + e^{-v_i})^{-1}.$$

The first is a hyperbolic tangent that ranges from -1 to 1, while the other is the logistic function, which is similar in shape but ranges from 0 to 1. Here y_i is the output of the i-th node (neuron) and vi is the weighted sum of the input connections. Alternative activation functions have been proposed, including the rectifier and soft plus functions. More specialized activation functions include radial basis functions (used in radial basis networks, another class of supervised neural network models).

Computing SVM Classifier

Computing the (soft-margin) SVM classifier amounts to minimizing an expression of the form

$$\left[\frac{1}{n}\sum_{i=1}^{n}\max\left(0, 1 - y_i(w \cdot x_i - b)\right)\right] + \lambda\|w\|^2. \qquad (2)$$

We focus on the soft-margin classifier since, as noted above , choosing a sufficiently small value for λ yields the hard-margin classifier for linearly classifiable input data. The classical approach, which involves reducing (2) to a quadratic programming problem, is detailed below. Then, more recent approaches such as sub-gradient descent and coordinate descent will be discussed.

1)Primal:

Minimizing (2) can be rewritten as a constrained optimization problem with a differentiable objective function in the following way.

For each $i \in \{1, \ldots, n\}$ we introduce a variable $\zeta_i = \max\left(0, 1 - y_i(w \cdot x_i - b)\right)$. Note that ζ_i is the smallest nonnegative number satisfying $y_i(w \cdot x_i - b) \geq 1 - \zeta_i$. Thus we can rewrite the optimization problem as follows

$$\text{minimize } \frac{1}{n}\sum_{i=1}^{n}\zeta_i + \lambda\|w\|^2$$

$$\text{subject to } y_i(w \cdot x_i - b) \geq 1 - \zeta_i \text{ and } \zeta_i \geq 0, \text{ for all } i.$$

2) Dual:

By solving for the Lagrangian dual of the above problem, one obtains the simplified problem

$$\text{maximize } f(c_1 \ldots c_n) = \sum_{i=1}^{n} c_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} y_i c_i (x_i \cdot x_j) y_j c_j,$$

$$\text{subject to } \sum_{i=1}^{n} c_i y_i = 0, \text{ and } 0 \leq c_i \leq \frac{1}{2n\lambda} \text{ for all } i.$$

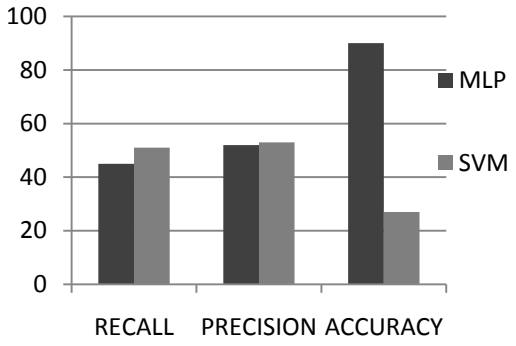
This is called the dual problem. Since the dual maximization problem is a quadratic function of

the c_i subject to linear constraints, it is efficiently solvable by quadratic programming algorithms.

Here, the variables c_i are defined such that

$$\vec{w} = \sum_{i=1}^{n} c_i y_i \vec{x}_i .$$

VII.RESULT AND CONCLUSION



This study focused on analysing visual features of rice seed images such as colour, shape, texture. It can be applied different classification models using these types of features. This research indicated that image processing techniques can combine with classification techniques such as MLP, SVM, and Decision tree and Bayesian network to identify rice seeds in mixed samples. All the methods using simple features proved the best capability and accuracy of classification; on average it achieved 90, 27%, 90.54% respectively. The performance can be improved by using other types of features and further investigation of classification models. It attempted to highlight the basic problems of rice industry to analyse the quality of rice grains and also highlighted the related work of researchers to eradicate the problem related to quality analysis of rice grains,

Future Work: The future work will be to correct the effects of the Non-Uniform Illumination and apply TopHat Transformation on rice grains and thus calculates the various parameters for the quality analysis of Indian Basmati rice grains so as to classify them into Normal, Small and Long rice seeds.

.

REFERENCES

[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,Mohd-Azemi 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.