Rice Grain Identification and Quality Analysis using Image Processing based on Principal Component Analysis

Muhammad Junaid Asif [1], [4], Tayyab Shahbaz [2], Dr. Syed Tahir Hussain Rizvi [3], Sajid Iqbal[1]

[1] Department of Mechatronics & Control Engineering, University of Engineering & Technology, Lahore.

[2] Department of Electrical Engineering, University of Engineering & Technology, Lahore.

[3] Dipartimento di Automatica e Informatica, Politecnico di Torino, Turin, Italy.

[4] R&D Engineer, AMBER Capacitors Limited, Lahore, Pakistan.

muhammadjunaid94ee@gmail.com

***Abstract*— Different types of foods are available in grain form, but rice is one of the important and most used cereal grains of Pakistan and all over the world. Quality inspection of rice grain is also important for both local as well as export purpose. It is necessary to propose an automatic solution to perform the quality analysis as well as to distinguish between different classes of rice. Main purpose of this paper is to present an image processing-based solution to classify the different varieties of rice and its quality analysis. An approach based on the combination of principal component analysis and canny edge detection is used for the classification. Quality analysis of rice grain is determined by morphological features of rice grains. These morphological features include eccentricity, major axis length, minor axis length, perimeter, area and size of the grains. Six different varieties of rice are classified and analyzed in this paper. A database is trained by feeding the 100 images of each variety of rice grains. Classification and quality analysis is done by comparing the sample image with database image. Canny edge detector is applied to detect the edges of rice grains. Eigen values and Eigen vectors are calculated on the basis of morphological features. Then by applying the PCA, different varieties of rice are classified by comparing the sample image with a database. Results obtained in terms of classification and quality analysis are 92.3% and 89.5% respectively. Proposed system can work well within minimum time and low cost.**

Keywords—Canny edge detection, Principal component Analysis, Quality analysis and classification, Rice grain analysis, Rice grain axis length, rice grain area, Rice grain perimeter

I. INTRODUCTION

Rice is one of the most important and most used cereal grains of Pakistan as well as overall world. It is also most important for the human nutrition caloric intake. It generally provides 130 calories per 100 grams with 1% calcium, iron and 3% magnesium. It is the seed of grass specie Oryza sativa (Asian rice) and Oryza glaberrima (African rice) **[1]**.

Due to the blessed agro-climatic conditions, Pakistan is known for its production of three different types of rice which includes aromatic, medium and round grain rice. China, Pakistan, India, Indonesia and Vietnam are the main rice producing countries **[2]**. It is the second staple food of Pakistan first being wheat. According to the latest ranking, Pakistan is ranked among top twenty producers of rice with the annual production of ***9.935*** million metric tons [1]. It is also an important cash crop. Pakistan is the fifth largest exporter with annual export of ***38,00,000*** metric per ton and cultivating different varieties of rice including basmati rice, kernel rice, kainat rice, khushboo rice, super basmati rice, kainat saila rice, non-basmati long and short grain rice.

All of the rice producing countries are striving their best for improving the rice quality. Proper quality inspection of rice is very important. **[3]** Therefore, it is necessary to propose an automatic solution for the classification and quality analysis of different varieties of rice grains. In Pakistan, different types of software are used in different mills to automate all processes including ***rice server and Compute rice*** developed by ***AGsoft*** and ***Softronix*** respectively. But there is no locally developed software for the quality analysis and classification and most of the rice mills are using ***Australian developed software SATAKE RSQI10A Rice grain scanner*** for this purpose. **[4]**. Main objective of this system is to propose a cost effective automatic solution for the classification and QA of rice grains.

Quality analysis using IVP technique is a popular research domain and it is advantageous over traditional methods for analysis, because of its easy implementation, no human interference, cost effective and less time consumption **[5]**. An approach based on the combination of principal component analysis and canny edge detection is used for the variety classification of rice. However, Quality of rice will depend on the different morphological features of grains including major and minor axis length, eccentricity, perimeter and area of rice grains.

First of all, an image will be acquired with the help of color digital camera and perform different operations like pre-processing, background estimation and RGB to binary conversion. Second step is to build the database for the training of system. System is trained by feeding at least 100 images of each variety of rice with white background. Data in the form of morphological features, eigen values and vectors of all of data base images will be stored.

Compute Rice, Software for Rice mills is available at: http://www.softronix.pk/products/software-for-rice-mills.htm

Rice Server, Details for software available at: https://www.iagsoft.com.pk/riceserver_detail.php

978-1-5386-8204-3/18/$31.00 ©2018 IEEE

Training needs to be done once, so that the system can identify the type of rice and analyze the quality of grains. For classification and quality analysis or rice grains, sample image will be compared with database. Acquired images are then subjected to preprocessing, smoothing and estimation of background. After estimating the background, RG B to binary conversion is done on the image of grain, dividing the grain into 0 (Black) and 1 (White).

Edges of grains are then detected by the use of canny edge detector too calculate the different morphological features. The quality of rice will depend on these morphological features including axis length, perimeter, eccentricity and area.

Calculated morphological features are used to calculate the eigen values and vectors of the grains. PCA is applied to classify the different varieties of rice on the basis of Eigen values and vectors. Same steps will be performed on the sample image automatically. Classification and quality analysis is done by comparing the sample image with database.

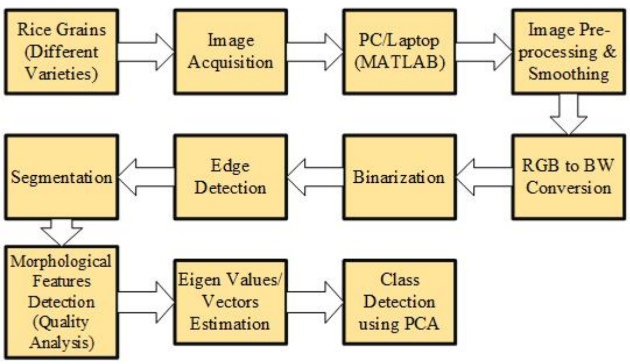


Fig 1 – Block Diagram of Proposed System

This paper is divided as follows: Brief literature review is discussed in Section II. Section III discusses the brief algorithm of principal component analysis and canny edge detection. The proposed methodology presents in section IV. Section V discusses the detailed description of image processing-based system. Section VI describes the implementation of proposed system. The results and discussion are in section VI.

II. LITERATURE REVIEW

There is a lot of research going in the field of image and video processing. Teams from many universities are entering in this field every year and contributing to the fields of AI, Neuro robotics, Computer Vision, Control and Communication Techniques and algorithms concerning detection, motion, sensing and video processing.

Optimal and image processing-based technique presented for the characterization and quality analysis of rice grains on the basis of chalky white part of rice grain. White chalky area of grains is detected by the use of extended maxima operator **[3]**.

Computer vision-based quality analysis is frequently used, because it can provide fast and reliable results. A technique based on PCA and K-mean cluster analysis also presented for the quality analysis of rice bran collected from rice mills. Rice bran can be defined as the wastage collected during the rice milling process and can be used for oil manufacturing. On the basis of oil contents there are different types of rice bran e.g. oil contents of boiled rice bran is ranging from (20~26%), and oil contents for raw rice bran are ranging from (16~18%) **[5]**.

A low-cost solution for the replacement of SATAKE RSQI10A presented for quality analysis of rice grains. Locally developed software minimizes all features and operations of SATAKE grain analyzer with overall efficiency of 95% **[4]**.

Quality analysis of grainy materials can also be done using depolarized Rayleigh-Scattering. A solution based on the vector network analyzer and two corrugate horn antennas was presented at 35GHz **[6]**. The efficiency of these systems is more than 70%.

But due to the increase in efficiency, cost and complexity of system also increase. In order to decrease the cost and reduce the complexity of system, a technique based on image and video processing is presented in this system.

A technique based on image processing and ANN is also presented to check the size and quality of basmati rice granules. The proposed system is implemented using Open CV is a library of functions that image processing in real time. The estimated morphological features are given to the ANN to classify the rice granules **[7]**. We are implementing our system based on PCA and canny edge detection using MATLAB, because of the fast processing and easy to use environment of MATLAB.

Manual quality analysis is time consuming and costly. An alternative solution proposed for quality analysis of rice on the basis of physical and chemical properties. Physical properties include size, shape, chalkiness, milling degree while chemical properties consist of gelatinization and temperature **[8]**.

Efforts were made to develop the grain quality analyzer to replace the traditional human sensory panel by the development of quality analyzer based on physical attributes of grains **[9] [10].**

In this paper, a system based on image processing is presented to classify the different varieties of rice. The whole system is implemented using PCA and canny edge detector. Main reason to use these techniques is fast and easy processing. A single trained database can be used for analysis and classification. Another advanced feature is that the system is portable and easy to carry wherever needed.

III. ALOGORITHMS

3.1 Principal Component Analysis

A linear transformation method used to reduce the multidimensional database into a linear database to analyze and visualize the data. A dataset may be different patterns, different features, etc. We are using this technique in our project to find the components that maximize the variation of a database. Nowadays, the use of PCA is increasing in every field like Neuro robotics, IVP, Pattern recognition, and signal processing etc. **[11]**. PCA reduces the observed variables to the independent variables. Independent variables will be then used to describe the data economically.

PCA is widely used in the data compression. It is generally based on the calculation of Eigen value, decomposition of data and covariance matrix. The steps to implement PCA are following:

- First of all, standardize the data in the matrix form.
- Calculate the mean of data in the form of Column matrix.
- Subtract the calculated mean from each point of database.
- Calculate the variance and co-variance matrix.
- Calculate the Eigen value and Eigen vectors of the covariance matrix (C).

$$C\alpha = \lambda\alpha$$

- Where λ represents the Eigen value and α is representing the Eigen vector of above calculated co-variance matrix.
- Sort out the Eigen values so that $\lambda 1 > \lambda 2 > \lambda 3 > > > \lambda n$.
- Select the Eigen vectors to generate a compressed data.

3.2 Canny Edge Detection

Edge Detection is a technique used to detect edges of objects in an image. Edges are the location of rapid intensity of images and play very important role in human perception. For edge detection, we can apply different masks and different techniques, we are using canny edge detector in our proposed system. Canny edge detection is an optimal recognition process to detect the gray scale image boundaries. The edges are made by locating the local maxima and minima of the gradient of intensity function [9]. Main steps for canny edge detection are following:

- **Smoothing:** It defines as the blurring of the image to remove the noise.
- Find out the **gradients** and marked the edges where the gradients of an image are high.
- Apply the **non-maxima suppression** technique to mark the edges.
- Apply the **double thresholding technique** to find out the potential edges.

IV. PROPOSED METHODOLOGY

To enhance the quality and classify the different varieties of rice grains, image processing-based technique is proposed [12]. Focus of this system is using principal component analysis algorithm for classification and canny edge detection based morphological features for quality of rice grains. Using canny edge detection, morphological features are calculated and then based on these features quality of grains is predicted. Then these features are subject to PCA for the calculation of Eigen values and Eigen vectors.

- We have considered five different varieties of rice grains including ***Super Colonel, Khushboo, Basmati, Kainat Sailla, Old Awami.***

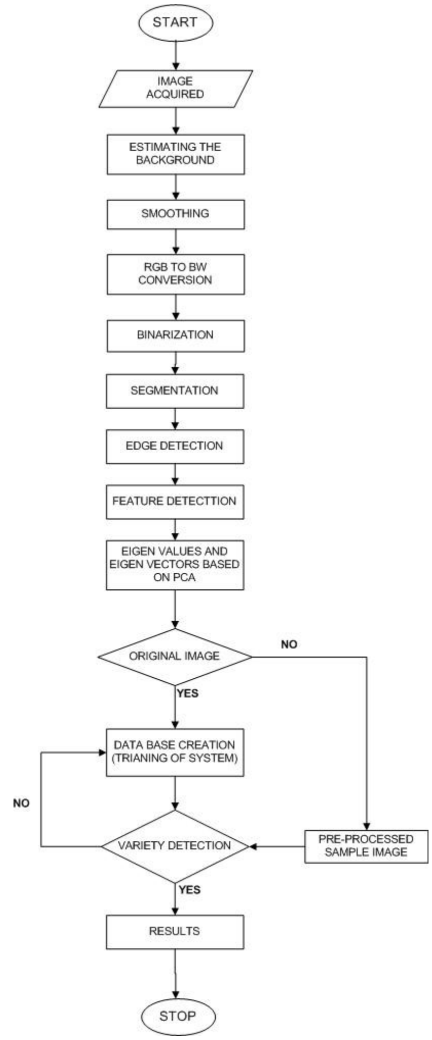


Fig 2 – Block Diagram of Proposed System

- 100 images of each type are captured by placing grains of rice on a black sheet of paper.
- Acquired images are then subjected to pre-processing operations (including background estimation, contrast adjustment, smoothing and RGB to BW conversion)
- Binarization and segmentation process is then applied to pre-adjusted images. Canny edge detector is used for detection of grain edge.
- Morphological features, Eigen values and Eigen vectors calculated for all images of each type. A database is created on the basis of images.
- Eigen values and vectors are calculated on the basis of morphological features and stored in data base.
- Same steps are then applied to sample image. Eigen values and Eigen vectors of sample image are compared by the database. Results are based on the least variation between the Eigen values and vectors of database and sample image.

V. DESCRIPTION OF SYSTEM

A. Images Acquisition:

Acquiring the images in any image processing based project is the first step because image processing cannot be done without acquiring the images. We captured six images of different types of rice grain with approximate black background.

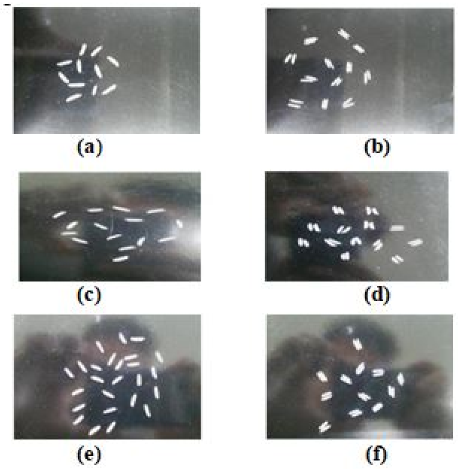


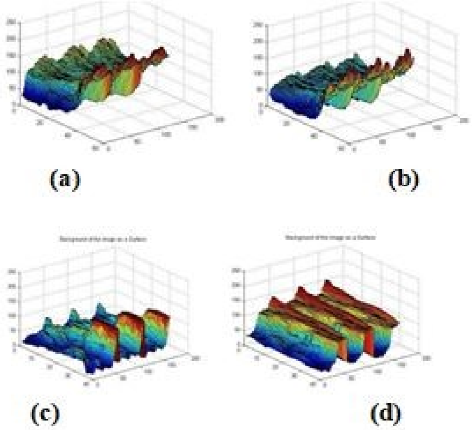
Fig 3 - (a) Sample 1- Kainat Sailla Rice Single Grain (b) Sample 2-Kainat Sailla Rice Double Grain (c) Sample 3 - Super Colonel Rice Single Grain (d) Sample 4 - Super Colonel Rice Double Grain (e) Sample 5 - Old Awami Rice Single Grain (f) Sample 6 - Old Awami Rice Double Grain

B. Image Pre-processing:

The images acquired may have irregular background and unwanted distortion. Image pre-processing is an important step to eliminate these unwanted sources, irregular background noise and blur **[13]**. The images acquired are stored on the hard disk of a PC.

C. Estimation of Background:

Background of the image is blurred, distorted and irregular. Background of images is estimated by using ***imopen().*** Estimated background is then plotted as surface as shown in fig 4. Estimated background is then subtracted from the original image to remove the unwanted noise and distortion in the background.



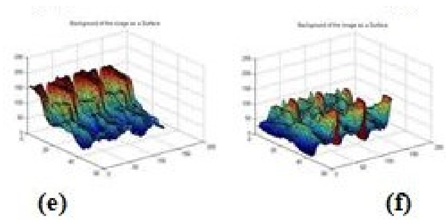


Fig 4 – Surface Plot of Estimated background (a) Sample 1 Kainat Sailla Rice Single Grain (b) Sample 2 Kainat Sailla Rice Double Grain (c) Sample 3 Super Colonel Rice Single Grain (d) Sample 4 Super Colonel Rice Double Grain (e) Sample 5 Old Awami Rice Single Grain (f) Sample 6 Old Awami Rice Double Grain

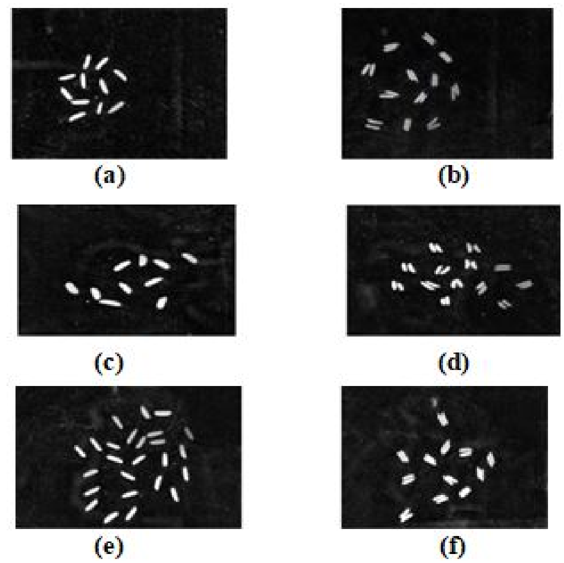


Fig 5 – Adjusted Images (a) Sample 1 Kainat Sailla Rice Single Grain (b) Sample 2 Kainat Sailla Rice Double Grain (c) Sample 3 Super Colonel Rice Single Grain (d) Sample 4 Super Colonel Rice Double Grain (e) Sample 5 Old Awami Rice Single Grain (f) Sample 6 Old Awami Rice Double Grain

D. RGB to Binary Conversion:

Binary image is defined as the image having only two levels 0 (for black) and 1 (for white). Binary version of images is created after setting the contrast level.

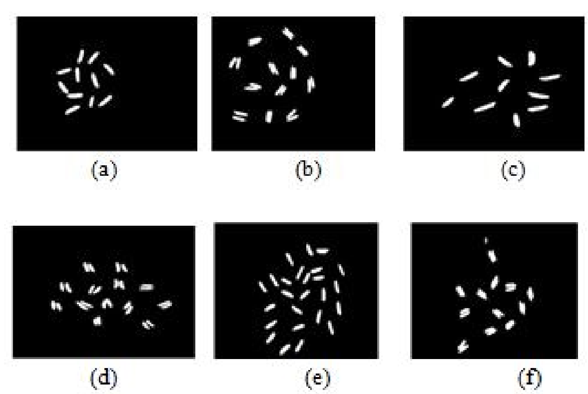


Fig 6 – Binarized Images (a) Sample 1 Kainat Sailla Rice Single Grain (b) Sample 2 Kainat Sailla Rice Double Grain (c) Sample 3 Super Colonel Rice Single Grain (d) Sample 4 Super Colonel Rice Double Grain (e) Sample 5 Old Awami Rice Single Grain (f) Sample 6 Old Awami Rice Double Grain)

E. Edge Detection:

In our proposed system canny edge detector is used for edge detection. Edge detection is a very important task because the quality of rice is directly depends on the edges. Canny edge detector is an edge detection operator developed by John F. Canny used to detect the edge in an image. These edges can be used to measure the different parameters of different objects in image **[15].** Segmented images are then subjected to canny edge detector for edge detection of rice grains. It reduces the error rate, amount of processing data. Edge detected images are shown in fig 7.

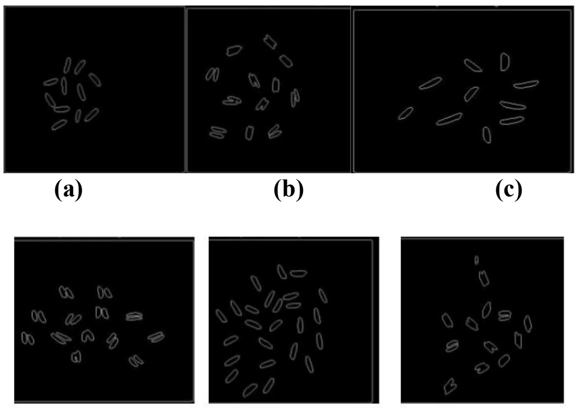


Fig 7 – Edge Detected Images (a) Sample 1 Kainat Sailla Rice Single Grain (b) Sample 2 Kainat Sailla Rice Double Grain (c) Sample 3 Super Colonel Rice Single Grain (d) Sample 4 Super Colonel Rice Double Grain (e) Sample 5 Old Awami Rice Single Grain (f) Sample 6 Old Awami Rice Double Grain)

F. Segmentation: Binarized images are then subjected to segmentation. Segmentation divides the image in to multiple segments so that they could analyze easily. Segmentation is an important step to calculate the morphological features [14]. Segmented images are shown in fig 8

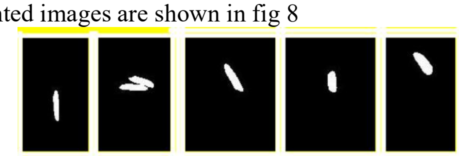


Fig 8 – Segmentation of Images

G. Identifying the objects:

Bwconncomp() is used to identify the all connected objects in an image. It will tell the image size, image pixels, connectivity and number of objects in an image.

Table 1: Number of Grains

Sample	**1**	**2**	**3**	**4**	**5**	**6**
Number of Grains	11	24	10	26	25	24

H. Feature Detection: Based on the identification of objects in previous step, different morphological features are calculated by using ***regionprops(cc, features).*** These calculated morphological features include eccentricity, major axis length, minor axis length, perimeter and area of each rice grain in an image. Then this collected data is compared by the trained database. Different varieties of rice grains will have different axis lengths, perimeter and area. This data will be calculated for different varieties and fed in the database. Quality of rice is analyzed by comparing the calculated data of sample image with the pre-defined data of different varieties of rice grains.

Table 2: Morphological Features of Sample 1

Area	Eccentricity	Perimeter	Major Axis Length	Minor Axis Length
262	0.97	78.52	36.20	11.14
308	0.95	79.74	35.88	10.64
326	0.96	88.08	39.89	10.54
262	0.96	83.21	36.20	11.14
327	0.97	79.79	37.23	10.03
280	0.96	77.11	35.80	10.14
276	0.96	80.56	37.53	11.04
292	0.96	62.62	28.72	10.21

Table 3: Morphological Features of Sample 2

Area	Eccentricity	Perimeter	Major Axis Length	Minor Axis Length
218	0.94	65.84	29.40	89.65
201	0.97	68.87	32.23	8.11
210	0.97	74.28	34.80	7.94
201	0.95	65.45	29.59	8.11
402	0.90	122.42	32.23	16.69
416	0.87	91.82	33.47	16.42
325	0.88	82.52	30.17	14.30
346	0.90	98.91	32.98	14.61
342	0.80	95.39	30.84	14.50

Table 4: Morphological Features of Sample 3

Area	Eccentricity	Perimeter	Major Axis Length	Minor Axis Length
277	0.96	78.56	34.89	10.31
385	0.98	89.15	48.46	10.38
354	0.80	82.56	56.66	10.60
337	0.76	76.87	40.36	11.44
350	0.98	89.94	36.53	12.16
397	0.98	106.52	32.78	13.77

Table 5: Morphological Features of Sample 4

Area	Eccentricity	Perimeter	Major Axis Length	Minor Axis Length
250	0.802	73.84	34.29	9.53
278	0.955	79.49	32.11	11.24
292	0.969	77.98	36.04	10.47
304	0.982	84.08	35.59	11.11
284	0.920	83.39	38.48	9.59
355	0.992	79.84	37.20	12.36
310	0.925	81.25	35.61	11.27
228	0.855	76.08	37.02	9.94
261	0.892	65.94	34.25	9.26
273	0.882	81.25	33.95	10.37

Variety Detection based on PCA:

Based on calculated morphological features, mean value of each feature is calculated. Then by using mean and average values, Eigen values and vectors of each image are calculated. The calculated Eigen values and vectors of sample image are then compared with the Eigen values and vectors of images fed in database. If the calculated Eigen values and Eigen vectors of sample image are exactly matched with the Eigen values and vectors of any image in trained database, then the variety of rice grain is verified, otherwise the image is then fed to database for training.

VI. IMPLEMENTATION OF THE SYSTEM

The first step is to create a database. 100 images of each variety of rice grain are captured from the webcam. Creating the database of different varieties of rice is important to analyze the quality and classification of rice. Training needs to be done once, so that the system can analyze and classify the rice grain.

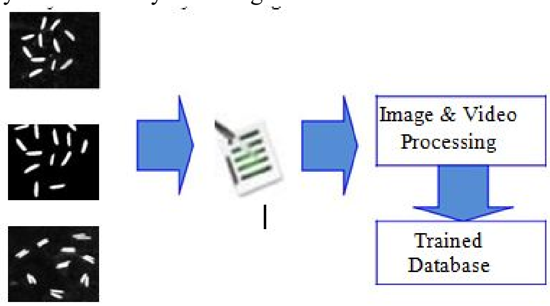


Fig 9 System Training

After creating the database, now capture the image of any sample of rice grain and upload it into the system. The system will then pre-process the sample image. Then after pre-processing, the system will determine the different morphological features of sample image and compare them to the features of database images. After analyzing the quality of grains, it applies the principal component analysis to detect the variety either it belongs to kainat sailla or old awami etc.

VII. RESULTS & DISCUSSION

In this research work, a simple, portable and efficient system is designed to analyze the quality and classify the different varieties of rice grains. Purpose of this research work is to help the industrialists to verify the quality of rice grains for import as well as export purposes. The whole system is implemented using PCA and canny edge detector. Results obtained in terms of classification and quality analysis are 92.3% and 89.5% respectively. Proposed system can work well within minimum time and low cost. In Future, efficiency of proposed system can be increased by using the algorithms of General Hough transform (GHT).

VIII. REFERENCES

[1] Abedullah, Shahzad Kouser And Khalid Mushtaq, "Analysis Of Technical Efficiency Of Rice Production In Punjab (Pakistan) Implications For Future Investment Strategies" Pakistan Economic and Social Review Volume 45, No. 2, pp. 231-244.

[2] A.R.Pazoki, F.Farokhi, Z.Pazoki, "Classification of Rice grain varieities using tow artifacial neural networks (MLP and Neuro Fuzzy)", The Journal of Animal and Plant Sciences", pp. 336-343, 2014.

[3] Engr. Zahida Parveen, Dr. Muhammad Anzar Alam, Engr. Hina Shakir, "Assessment of Quality of Rice Grain using Optimal and image processing technique", International Conference on Communication, Computing and Digital Systems (C-CODE), March 2017.

[4] Syed Farooq Ali, Halima Jamil, Razia Jamil, Iqra Torij, Saira Naz, "Low Cost Solution for Rice quality analysis using Morphological parameters, and its comparison with Standard measurements", International Multitopic Conference (INMIC'17), November 2017.

[5] Devraj Vishnu, Gunjan Mukherjee, Arpitam Chatterjee, "A Compuetr Vision approach for grade identification of Rice Bran", Third International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), November 2017.

[6] Anssi P. Toropainen, "A Method for measuring properties grainy materials using Depolarized Rayleigh Scattering", 22nd Microwave conference, vol. 2, pp 1041-1045, sep 1982.

[7] P. Neelamegam, S. Abirami ; K. Vishnu Priya ; S. Rubalya Valantina, "Analysis of rice granules using image processing and neural network", Conference on Communication & Information technologies, Jeju Island, pp. 879-884, Apr 2013.

[8] Bhagyashree Mahale, Prof. Sapana Korde, "Rice Quality Analysis using Image Processing Technique", International conference for convergence for technology, April, 2014.

[9] Abirami S., Neelamegam. P, Kala. H, "Analysis of Rice granules using image processingand neural netwrok pattern recognition tool", International Journal of Computer Applications (0975-8887), pp. 20-24, Volume 96-7, June 2014.

[10] Namita Patel. Hardik Jayswal, Amit Thakkar, "Rice Quality Analysis based on physical attributes using image processing technique", International conference on recent trends in Electronics Information Communication technology, May 2017.

[11] Rubi Kambo, Amit Yerpude, "Classification of Basmati Rice Grain Variety using Image Processing and Principal Component Analysis", International Journal of Computer Trends and Technology (IJCTT) – vol. 11 number 2, May 2014.

[12] Chetna Vasudevbhai Maheshwari, Kavindra R. Jain ; Chintan K. Modi, "Non-Destructive Quality analysis of Basmati Oryza Sativa (SSP) Indica using image Processing", International Conference on Communication Systems and Network Technologies (CSNT), pp. 189-193, May 2012.

[13] Wan Putri N. W. M. Tahir, Norhaida Hussin, Zaw Zaw Htike and Wai Yan Nyein Naing, "Rice grading using image processing", ARPN Journal of Engineering and Applied Sciences, Vol. 10, No 21, Nov, 2015

[14] Sneha S. Kausal , S.V.More, " Review on Identification and Classification of Grains Using Image Processing", IJSR, Volume 4 Issue 4, April 2015.