



# A cascade network for the classification of rice grain based on single rice kernel

Ksh Robert Singh<sup>1</sup> · Saurabh Chaudhury<sup>2</sup>

Received: 3 October 2019 / Accepted: 28 January 2020 / Published online: 12 February 2020  
© The Author(s) 2020

## Abstract

This paper describes the classification of four different varieties of rice grain based on four sets of features, namely morphology, colour, texture and wavelet. The classification is carried out on single rice kernel using image pre-processing steps followed by a cascade network classifier. The performance of the classifiers based on the above feature sets is also compared. It is found that morphological feature is more suitable for the classification of rice kernels, as compared to other features. The number of input features is reduced by a feature selection process using statistical analysis system (SAS) software. The classification accuracy based on selected features is compared with that of original features using different classifiers. It is found that the selected features are able to provide classification accuracy very close to the original features. The performance of the proposed cascade classifier is also tested against standard datasets from the University of California, Irvine (UCI), and the results are compared with other classifiers. The results show that the proposed classifier provides better classification accuracy as compared to other classifiers.

**Keywords** Image · Rice grain · Morphological · Colour · Texture · Neural network

## Introduction

Digital image processing and computer vision plays a very important role in our day to day life. There is no area of technical endeavour that is not impacted in some way by digital image processing and computer vision. Computer vision system involves a hygienic and non-destructive way of mimicking the human visual system. This includes image analysis, understanding, identification, categorization and discrimination. The need of computer vision system arises in such environment where human visual system gets degraded with increasing number of observations and fatigue of eye sight due to prolonged recognition processes. Thus, a knowledge-based system is desirable for a reliable, quick and accurate recognition or identification of objects from noisy and partial input images [14]. Applications of computer vision system have significant impact in all domains of

Engineering including agriculture and food sector. The past few years marked the development of an automatic recognition and classification of agricultural products. Identification and grading of cereal grains has become an important part of grain handling process in agricultural sector. During this process, information on grain, mainly its types and quality, is required before the next course of operation is performed. When the crops are in the field, their yield becomes the prime importance of a farmer, but when the products of these crops come into the market, quality becomes the determining factor for its saleability. Also, it is important for the consumers to have proper information regarding the quality of the grains that they are going to purchase, as there may be chances of sub-standard grains being released into the market by some traders to gain profit unethically. Classifications of different varieties of seed were also carried out during the past few years using different classification models. Various classification models based on bulk grain samples were reported in [1, 2, 12, 18, 25, 37]. It is observed that the classification tasks involved in the above literature were not challenging. This is due to the fact that the classifications were performed based on grains belonging to different grain species such as rice, wheat, red gram, green gram, ground nut, barley, oats, rye, corn, pearl millet, peas and horse gram having distinguish-

✉ Ksh Robert Singh  
robert\_kits@yahoo.co.in

<sup>1</sup> Department of Electrical Engineering, Mizoram University, Aizawl 796004, India

<sup>2</sup> Department of Electrical Engineering, National Institute of Technology, Silchar 708810, India

able colour and texture attributes. Classification of cereal grain based on single grain kernel using different features, namely morphological, colour and texture, were carried out in [3, 5–7, 9, 15, 16, 19–24, 27–32]. The classification accuracies that were reported in the literature were mostly based on author's own data. It is observed that all the classification models proposed in the literature were not validated against publicly available data set. It is also observed that the accuracy differs based on number of classes, size of feature and the nature of the inputs (data set). Most of the cereal grain classifications carried out in the past few years were based on different varieties of wheat [1, 6, 7, 9, 18–23, 27–31, 37]. Rice is one of the major foods and it is widely consumed in almost all parts of the country. Varietal identification, classification and grading of rice are important not only for consumers and rice dealers, but also for plant breeders. It is found from the literature that very less work was carried out towards the classification of rice grain and the same are presented in [10, 12, 16, 32]. Most of the classification task reported in [2, 5, 7, 8, 12, 15, 16, 24, 27–32, 37] were based on artificial neural network. The final outputs of this network generally reach the threshold to achieve the desire target, which were set during training. It is learnt that too many redundant features jeopardise the classification accuracies [9, 16, 28, 29]. Classifications task based on different classifiers were also reported in [3, 7–9, 11, 17, 26, 29, 30, 32, 36]. It is learnt from the literature that in most cases the classification model based on artificial neural network outperforms other classification models [2, 7, 27, 29, 37]. Classification of grain based on bulk image does not require arranging the kernel in a non-touching fashion and does not involve image pre-processing steps. However, this may not be the best option for quality inspection and grading of seeds, where inspection at the kernel level is indeed important. The quality of seed has a direct impact on the growth and yield of that particular plant, so proper inspection of seed is important [1, 16]. Although computer vision system proved to be more efficient in terms of precision and rapidity, it is still a challenging task to identify and classify naturally varying appearance of biological entities such as cereal grain. This is because unlike other industrial products, the shape, size, colour and texture of agricultural products are not bound by a unique mathematical function [23, 27].

It is understood from the literature that the past few years have made significant works towards the classification of cereal grain using image processing and computer vision. However, the classification tasks were performed based on grains belonging to different classes of grain species having easily distinguishable features. Studies related to classification of different varieties of cereal grains belonging to a particular species are still less and rather challenging. Classification of different varieties of rice grains is rather challenging, as their appearance in terms of morphology,

colour and texture is almost similar. It is learnt from the literature that the classification task based on single grain kernel is tedious and involves various image pre-processing steps as compared to classification based on bulk grain images. Classification of cereal grain based on bulk images may be useful for the identification of grain variety in commercial markets. However, classification based on bulk grain images may not be suitable for varietal identification of grains to achieve better crop yield. The quality of rice seed or kernel has a distinct effect on the yield of rice. Thus, varietal identification and classification of rice kernel become important to breed key traits, which can be achieved only when the quality of the seeds are monitored individually at kernel level. Rice seed selection is a complex and sometimes misunderstood process. Some farmers assess individually entire plant's performance throughout the season when selecting seeds, whereas other farmers select seeds post-harvest from threshed crops in their storage. It is found from the literature that most of the cereal grain classification tasks were performed using different features and combination of features. As the feature size increases, the complexity of the classification task also increases. This paper deals with the classification of four varieties of rice grain based on single rice kernel using few selected features. The classification task is performed using a combination of feed-forward back-propagation neural network and a fuzzy inference system. As most of the work reported in the literature was based on the author's own data, the classification accuracies claimed by the authors are also restricted to their own data set. It will be unwise to claim the performance of a proposed classifier until it has been tested against a standard and publicly available data set. Thus, testing the proposed classification model against publicly available data set not only validates the classifier but also provides a room for other researchers to draw a comparison with the existing work. For better generalisation, the performance of the proposed classifier in this paper is also tested against publicly available data set from University of California, Irvine (UCI).

The steps involved in this work are arranged as follows: first, the colour images of the rice grain are acquired using a digital camera and it is converted to its corresponding HSI channels. Second, the rice kernel is segmented and its background is removed. The third step involves the extraction of features, namely morphological, colour, texture and wavelet features from the segmented rice kernel image. The fourth step involves classification based on the above features using different classifiers, namely back-propagation neural network (BPNN), fuzzy classifier, adaptive neuro-fuzzy inference system (ANFIS), K-nearest neighbour (KNN), linear discriminant analysis (LDA) classifier and naïve Bayes (NB) classifier. Finally, the classification accuracy of the proposed network model is compared with other classifiers using standard datasets.

## Materials and methods

### Image acquisition

Four different varieties of rice were collected from the north-eastern region of India. The rice kernel images were taken using 12-megapixel Nikon digital camera (Coolpix S2500) with CCD image sensor having a sensor size of 28.0735 mm<sup>2</sup>, image size of 4000 × 3000 and resolution of 300 dpi. The images were taken outdoor under natural lighting condition during morning hours so that the shadow of the camera does not appear on the image. An adjustable steel camera stand was used for mounting the camera and the camera is mounted at a height of 10 cm above the target object (i.e. rice kernel). The camera was set to macro mode before capturing the picture so that the image can be taken from a close distance. Sample images of each of the four rice types are presented in Fig. 1. In this figure, the four different varieties of rice were denoted by A, B, C and D, respectively. The original rice kernel images of the four varieties are represented by A.1, B.1, C.1 and D.1. The segmented binary images are represented by A.2, B.2, C.2 and D.2. Similarly, the segmented colour images are represented by A.3, B.3, C.3 and D.3. The images related to hue, saturation and value of each of the rice kernel image are represented as A.4–A.6 for type ‘A’, B.4–B.6 for type ‘B’, C.4–C.6 for type ‘C’ and D.4–D.6 for type ‘D’, respectively. The hue image is considered for segmentation and feature extraction as it can be seen from Fig. 1 that hue image provides a clear distinction between background and the object of interest (i.e. rice kernel). The overall steps involved in the classification process of four varieties of rice grain are presented in Fig. 2, which involves image segmentation, feature extraction, feature selection and classification using different classifiers and comparing their classification accuracies based on the input data presented in this paper and also with that of standard data sets.

### Image segmentation

Image segmentation is an important phase in image analysis process. It is the process of partitioning a digital image into multiple regions and extracting meaningful region known as region of interest for further image analysis [35]. The process here involves three main steps. In the first step, the RGB image is converted into HSV and the hue channel is selected for segmentation as it provides better distinction between the rice kernel and the background. This grey-scale image is then converted into binary image using the function *im2bw* in MATLAB 2016a. The second step involves the removal of unwanted segmented area other than the rice kernel. This is performed by removing all connected components (objects) that have fewer than 30 pixels using the function *bwareaopen* in MATLAB 2016a. The last step involves filling up of the

holes, if any, in the segmented binary rice kernel image using the function *imfill* in MATLAB 2016a [15].

### Feature extraction

Features are the characteristics or attributes of an object that help in object classification. The purpose of an image feature is to describe the object in a meaningful manner to aid the recognition processes for better object discrimination. Feature extraction is a process of extraction and generation of features to assist the classification task. This phase is critical because the quality of the feature influences the classification task [35]. Four different approaches, namely morphological, colour, texture and wavelet features, were used for the classification process.

### Morphological features

Ten numbers of morphological features were extracted using the function *regionprops* in MATLAB 2016a which are as follows:

1. *Area* It is the number of pixels in the region of rice kernel.
2. *Convex area* It is the number of pixels in the convex image. A convex image is a binary (logical) image with all pixels within the convex hull set to on. A convex hull is defined as the smallest convex polygon that can contain the region. Convex means the polygon has no corner that is bent inwards.
3. *Filled area* It is the number of pixels in the filled image. A filled image is also a binary (logical) image of the same size as the bounding box of the region. The on pixels correspond to the region, with all holes filled in. A bounding box is the smallest rectangle that can contain the region.
4. *Solidity* It specifies the proportion of the pixels in the convex hull that are also in the region. It is the ratio of area to convex area:

$$\text{Solidity} = \frac{\text{Area}}{\text{Convex area}}. \quad (1)$$

5. *Extent* It is the ratio of pixels in the region to pixels in the area of bounding box:

$$\text{Extent} = \frac{\text{Area}}{\text{Area of bounding box}}. \quad (2)$$

6. *Perimeter* It is the number of pixels present on the boundary of the rice kernel.

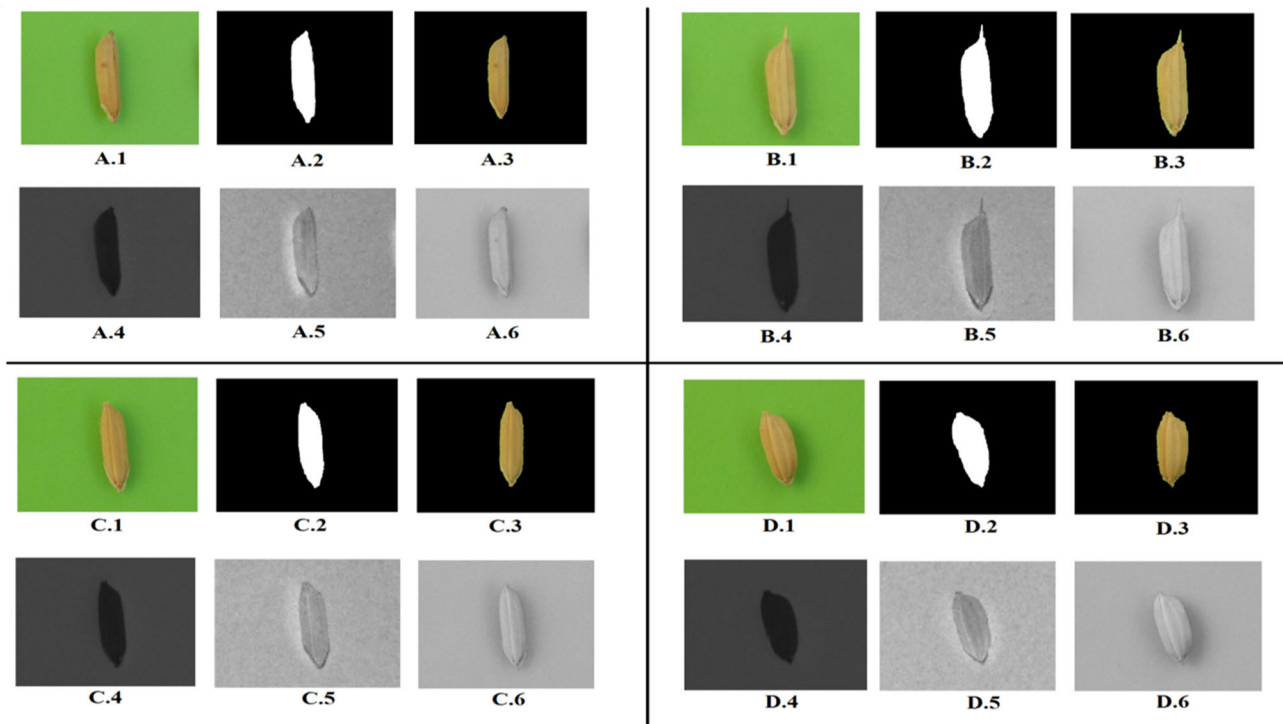


Fig. 1 Single rice kernel images

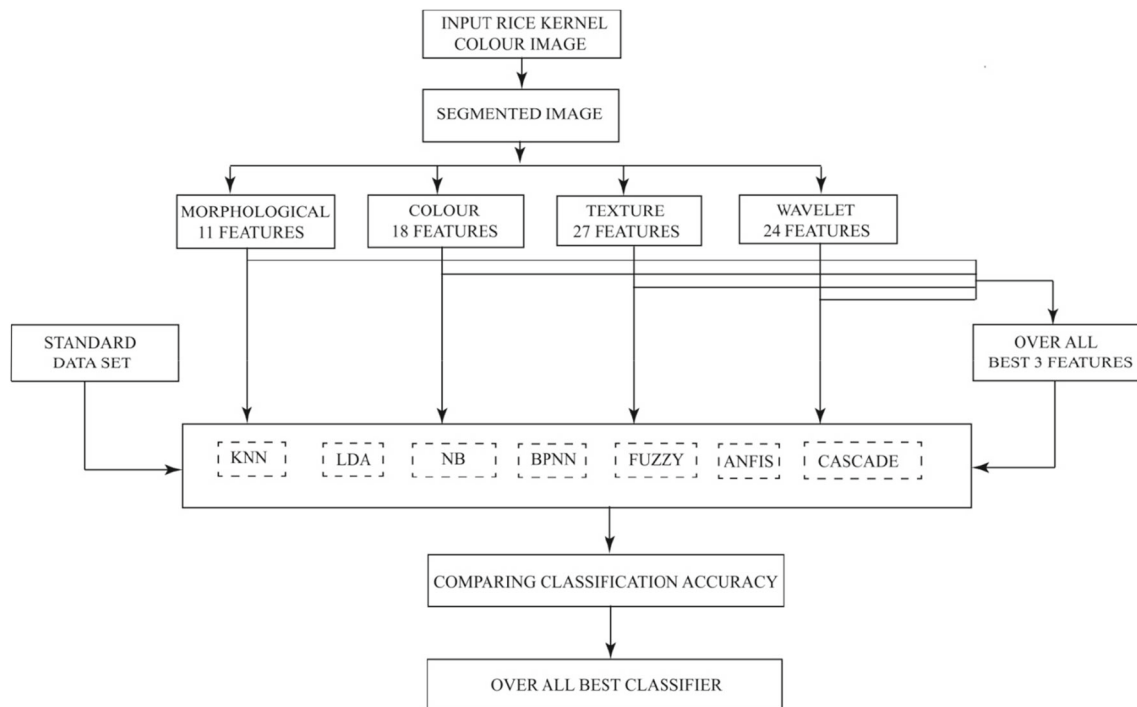


Fig. 2 Steps involved in the process of classification and performance comparison of different classifiers

7. *Eccentricity* It is the ratio of the distance between the foci of an ellipse (having the same second moment as that of the rice kernel region) to its major axis length.

8. *Equiv diameter* Diameter of a circle with the same area as the region of rice kernel:

**Table 1** 18 colour features

Sl. no.	Features	Sl. no.	Features
1.	Red mean	10.	Hue mean
2.	Red variance	11.	Hue variance
3.	Red range	12.	Hue range
4.	Green mean	13.	Saturation mean
5.	Green variance	14.	Saturation variance
6.	Green range	15.	Saturation range
7.	Blue mean	16.	Intensity mean
8.	Blue variance	17.	Intensity variance
9.	Blue range	18.	Intensity range

$$\text{Equiv\_diameter} = \sqrt{\frac{4 \times \text{Area}}{\pi}} \tag{3}$$

- 9. *Major axis* It is the length (in pixels) of the major axis of the ellipse that has the same second-moments as the rice kernel region.
- 10. *Minor axis* It is the length (in pixels) of the minor axis of the ellipse that has the same second-moments as the rice kernel region.

The eleventh feature is the number of pixels contained in the first 25 rows of the rice kernel, in a view to distinguish the tips of rice kernel, either pointed or blunt.

**Colour and texture features**

Eighteen numbers of colour features were extracted from the colour image of rice grain. The RGB (red, green and blue) colour planes were extracted from the colour image in MATLAB platform. The HSI (hue, saturation and intensity) image planes are derived from the RGB planes using Eqs. (4–7). Three statistical measures, namely mean, variance and range, were calculated for each of the six image planes (R, G, B, H, S and I). Thus, a total of eighteen numbers of colour features were extracted and are presented in Table 1.

$$H = \begin{cases} \theta & \text{if } B \leq G \\ 360 - \theta & \text{if } B > G \end{cases} \tag{4}$$

where

$$\theta = \cos^{-1} \left\{ \frac{\frac{1}{2}[(R - G) + (R - B)]}{[(R - G)^2 + (R - B)(G - B)]^{1/2}} \right\} \tag{5}$$

$$S = 1 - \frac{3}{(R + G + B)} \{\min(R, G, B)\}, \tag{6}$$

$$I = \frac{(R + G + B)}{3} \tag{7}$$

Texture features were extracted using a grey-level co-occurrence matrix (GLCM). This matrix provides information about the frequency of occurrence of pixel pairs in a particular direction. A grey-level co-occurrence matrix for a given input matrix of size 5 × 5, with maximum pixel intensity of 5 is given in Fig. 3. Four numbers of GLCM (for 4 different angles, 0°, 45°, 90° and 135°) were generated from each of the colour planes (R, G and B) and the average of these four is considered for extraction of texture features. Nine features were extracted from each of the average GLCMs using Eqs. (8–15). Here energy describes the uniformity of grey levels in the image. It is given by the sum of the square elements in the GLCM. The contrast feature provides information related to the local variation in an image.

This measures the intensity contrast between a pixel and its neighbour. Entropy measure the texture complexity. Complex textures have more entropy than simpler and uniform structure. Inverse difference moment feature indicates the amount of local uniformity in the image, whereas correlation feature indicates the statistical measure of how correlated a pixel is to its neighbour over the whole image. Homogeneity measures the closeness of the distribution of elements in the GLCM to GLCM diagonal. A total of 27 texture features were extracted from three colour planes (R, G and B) and are presented in Table 2 [2, 4, 13, 34].

$$\text{Mean} = \frac{\sum_{x,y} P(x, y)}{N^2}, \tag{8}$$

$$\text{Variance} = \frac{\sum_{x,y} [P(x, y) - \text{mean}]^2}{N^2}, \tag{9}$$

$$\text{Range} = \max(P) - \min(P), \tag{10}$$

$$\text{Energy} = \sum_{x,y} P^2(x, y), \tag{11}$$

$$\text{Entropy} = - \sum_{x,y} P(x, y) \log_2(P(x, y)), \tag{12}$$

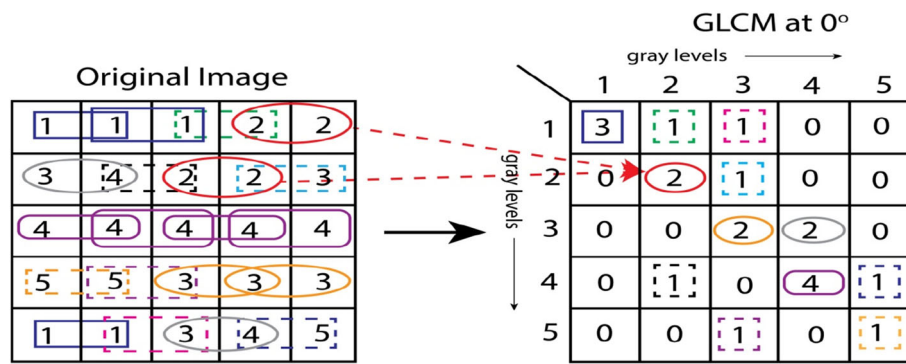
$$\text{Contrast} = \sum_{x,y} |x - y|^2 P(x, y), \tag{13}$$

$$\text{Inverse difference moment} = \sum_{x,y;x \neq y} \frac{P(x, y)}{|x - y|^2}, \tag{14}$$

$$\text{Correlation} = \frac{\sum_{x,y} [(xy)P(x, y)] - \mu_x \mu_y}{s_x s_y}, \tag{15}$$

where  $\mu_x$  are  $\mu_y$  are means and  $s_x$  are  $s_y$  are standard deviations defined by





**Fig. 3** A  $5 \times 5$  GLCM at zero degree direction

**Table 2** 27 texture features

Sl. no.	Features	Sl. no.	Features	Sl. no.	Features
1.	Red GLCM mean	10.	Green GLCM mean	19.	Blue GLCM mean
2.	Red GLCM variance	11.	Green GLCM variance	20.	Blue GLCM variance
3.	Red GLCM range	12.	Green GLCM range	21.	Blue GLCM range
4.	Red GLCM energy	13.	Green GLCM energy	22.	Blue GLCM energy
5.	Red GLCM entropy	14.	Green GLCM entropy	23.	Blue GLCM entropy
6.	Red GLCM contrast	15.	Green GLCM contrast	24.	Blue GLCM contrast
7.	Red GLCM I.D.M	16.	Green GLCM I.D.M	25.	Blue GLCM I.D.M
8.	Red GLCM correlation	17.	Green GLCM correlation	26.	Blue GLCM correlation
9.	Red GLCM homogeneity	18.	Green GLCM homogeneity	27.	Blue GLCM homogeneity

$$\mu_x = \sum_x x \sum_y P(x, y),$$

$$\mu_y = \sum_y y \sum_x P(x, y),$$

$$s_x = \sum_x (x - \mu_x)^2 \sum_y P(x, y),$$

$$s_y = \sum_y (y - \mu_y)^2 \sum_x P(x, y).$$

### Wavelet features

Wavelet transform is useful to analyse image at multiple resolutions. When we look at images, generally, we see connected regions of similar texture and intensity levels that combine to form objects. If the objects are small in size or low in contrast, we normally examine them at high resolution; on the other hand, if they are large in size or high in contrast, a coarse view is all that is required. If both small and large objects are present simultaneously, it can be advantageous to study and analyse them at several resolutions. Thus, features that might go undetected at one resolution may easily be detected at another [33].

In the discrete wavelet transform, an image signal can be analysed by passing it through an analysis filter bank followed by decimation operation. The analysis filter bank consists of low-pass filter and high-pass filter at each decomposition stage. When the signal passes through these filters, it splits into two bands. The low-pass filter, which corresponds to an averaging operation, extracts the coarse information of the signal. The high-pass filter, which corresponds to a differencing operation, extracts the detailed information of the signal. The output of the filtering operation is then decimated by two to retain the size of the decomposed image to that of the original image. A 2-D transform is accompanied by performing two separate one-dimensional transforms. First, the image is filtered along the row and decimated by two. It is then followed by filtering the sub-image along the column and decimated by two [14]. This operation splits the image into four bands, namely low–low (LL), low–high (LH), high–low (HL) and high–high (HH), respectively, as shown in Fig. 4a. Two-level wavelet decomposition of an input image is shown in Fig. 4b. Two-level wavelet decomposition is performed on all the three colour planes, namely red, green and blue. From each of the decomposed image, mean and variance were calculated for LLLL, LLLH, LLHL and LLHH sub-bands. A total of 24 features were extracted and the same is presented in Table 3.

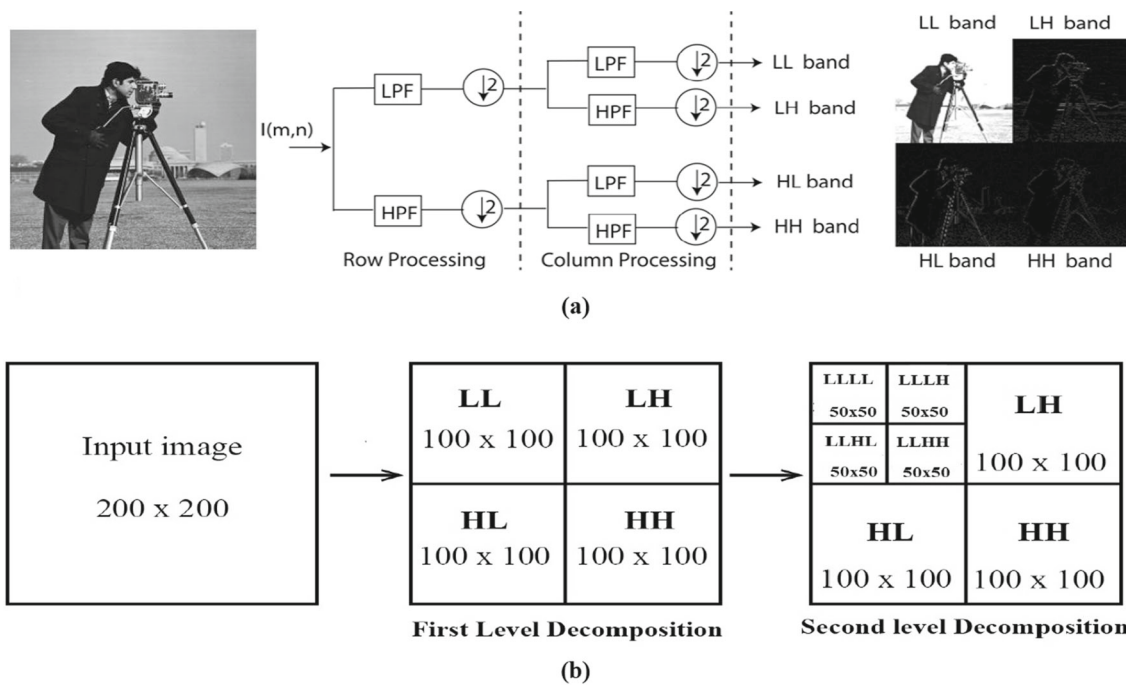


Fig. 4 a 2-D wavelet decomposition at level 1 and b a two-level wavelet decomposition

Table 3 24 wavelet features

Sl. no.	Features	Sl. no.	Features	Sl. no.	Features
1.	Red LLLL mean	9.	Green LLLL mean	17.	Blue LLLL mean
2.	Red LLLL variance	10.	Green LLLL variance	18.	Blue LLLL variance
3.	Red LLLH mean	11.	Green LLLH mean	19.	Blue LLLH mean
4.	Red LLLH variance	12.	Green LLLH variance	20.	Blue LLLH variance
5.	Red LLHL mean	13.	Green LLHL mean	21.	Blue LLHL mean
6.	Red LLHL variance	14.	Green LLHL variance	22.	Blue LLHL variance
7.	Red LLHH mean	15.	Green LLHH mean	23.	Blue LLHH mean
8.	Red LLHH variance	16.	Green LLHH variance	24.	Blue LLHH variance

### Classification model

Four varieties of rice grain were considered for classification. A dataset of 400 images was taken (100 images for each rice type). The classification was performed based on four different feature sets, namely morphological, colour, texture and wavelet. The classification accuracy was also evaluated based on selected features using SAS software (University Edition). Three best features were selected based on the selection strategy presented in [28]. To check the performance of the proposed classifier and also for generalisation, the classifier was tested with standard datasets from University of California, Irvine (UCI).

### Cascade network classifier

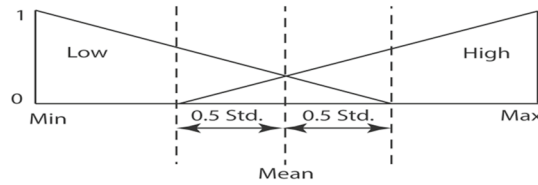
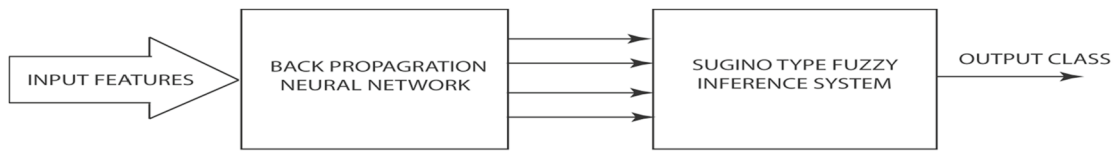
A neuro-fuzzy cascade network with back-propagation algorithm was implemented using MATLAB 2016a. A neural

network model with four outputs and two hidden layers is used, where the number of hidden nodes is given by the following equation [37]:

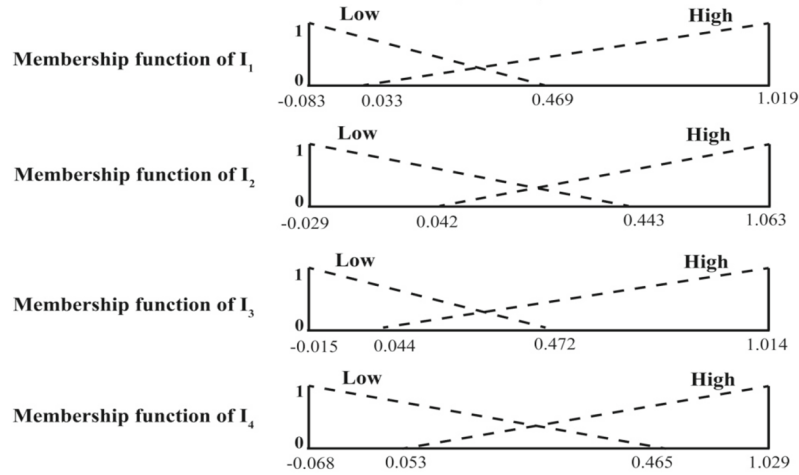
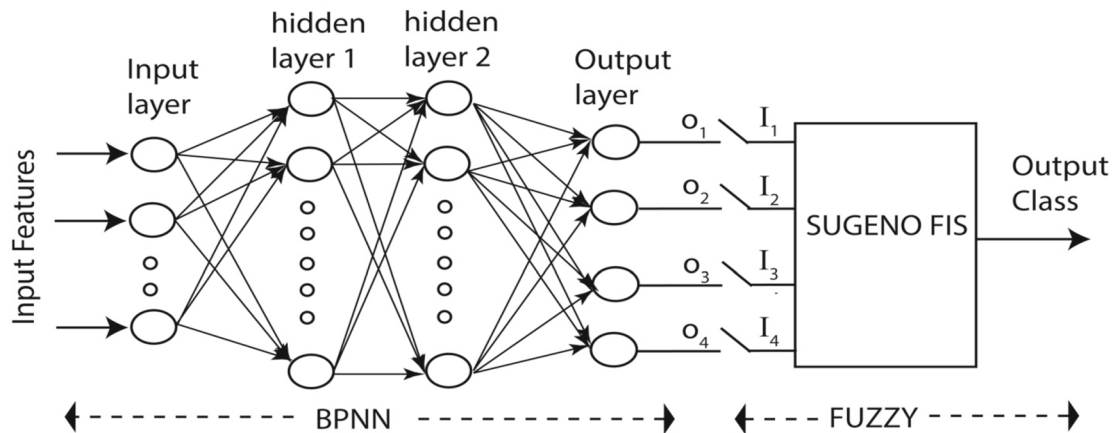
$$N = \frac{I - O}{2} + Y^{0.5}, \tag{16}$$

where  $N$  total numbers of hidden nodes,  $I$  numbers of input nodes,  $O$  numbers of output nodes, and  $Y$  numbers of training patterns.

The targets for each of the four classes were set to  $[0 \ 0 \ 0 \ 1]^T$ ,  $[0 \ 0 \ 1 \ 0]^T$ ,  $[0 \ 1 \ 0 \ 0]^T$  and  $[1 \ 0 \ 0 \ 0]^T$ . If the outputs of the network were simply threshold at a value (say 0.5), the output may sometimes provide a pattern which does not belong to those predefined target values. To avoid obtaining an output which is not at all belonging to any of the target value set during training, a cascade network with an aim to introduce a little bit of reasoning at the output portion of



(a)



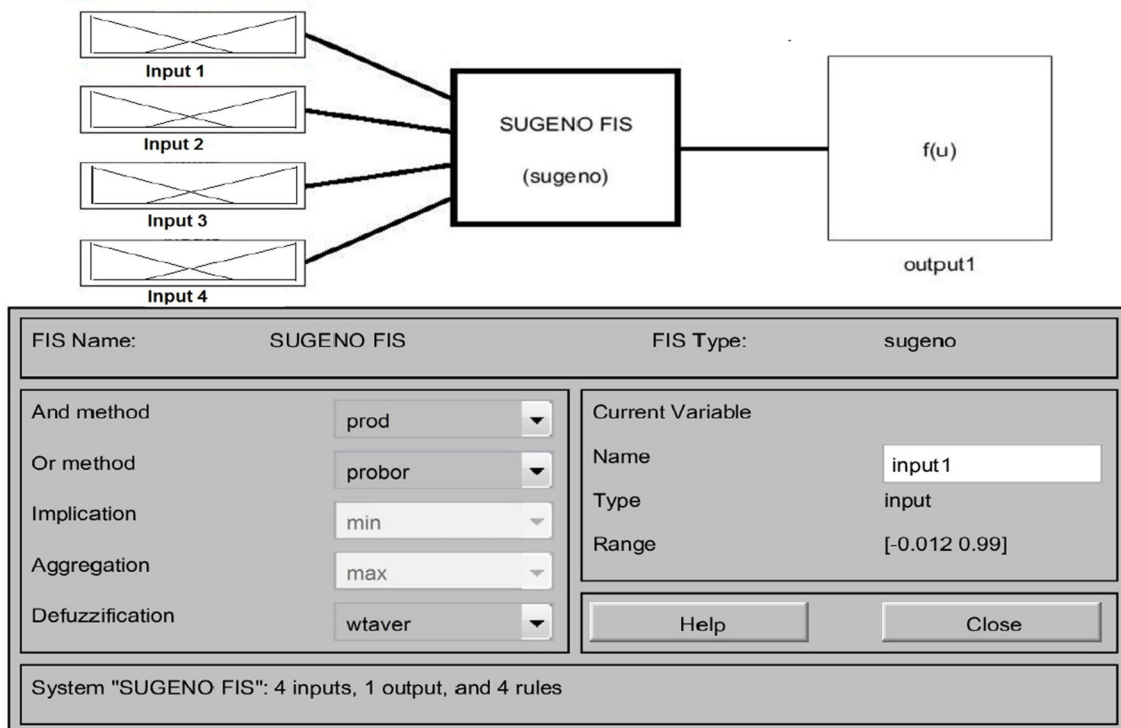
(b)

Fig. 5 Proposed cascade network classifier with triangular member function for the fuzzy inference system

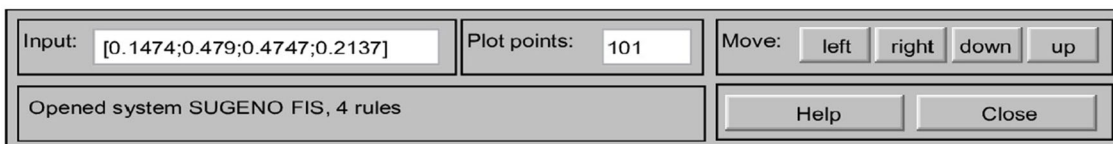
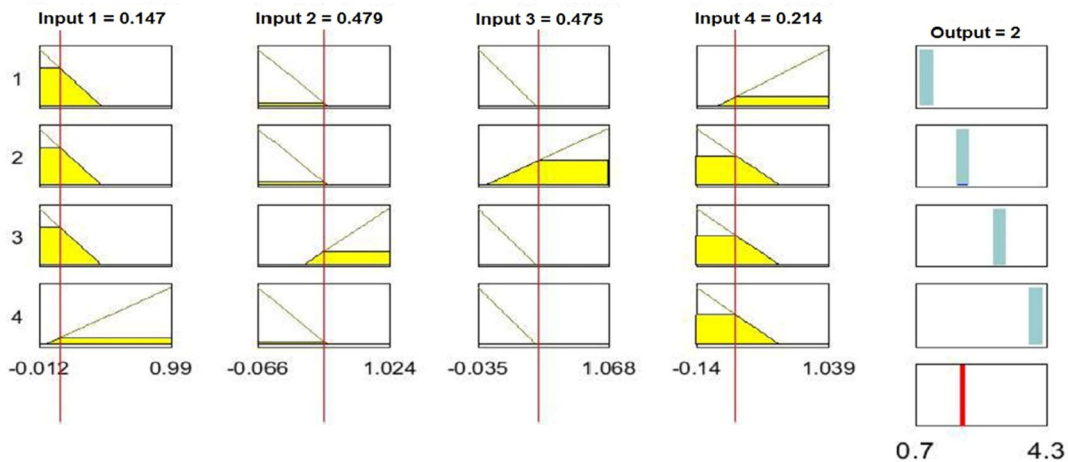
the BPNN network is proposed in this work. A Sugeno-type fuzzy inference system (with constant output) is introduced at the output of the back-propagation neural network (BPNN) which is shown in Fig. 5. The output of the BPNN is fed to a fuzzy inference system (FIS), where each of the four inputs to FIS is assigned with triangular membership function. The

output of fuzzy logic designer (in MATLAB platform) is shown in Fig. 6a and its corresponding rules are shown in Fig. 6b, c. For instance, if an input pattern  $I_1 = 0.147, I_2 = 0.479, I_3 = 0.475, I_4 = 0.2147$  is applied to FIS, the output will be class 2 as shown in Fig. 6b. The membership function of the fuzzy interface system is designed based on the out-





(a)



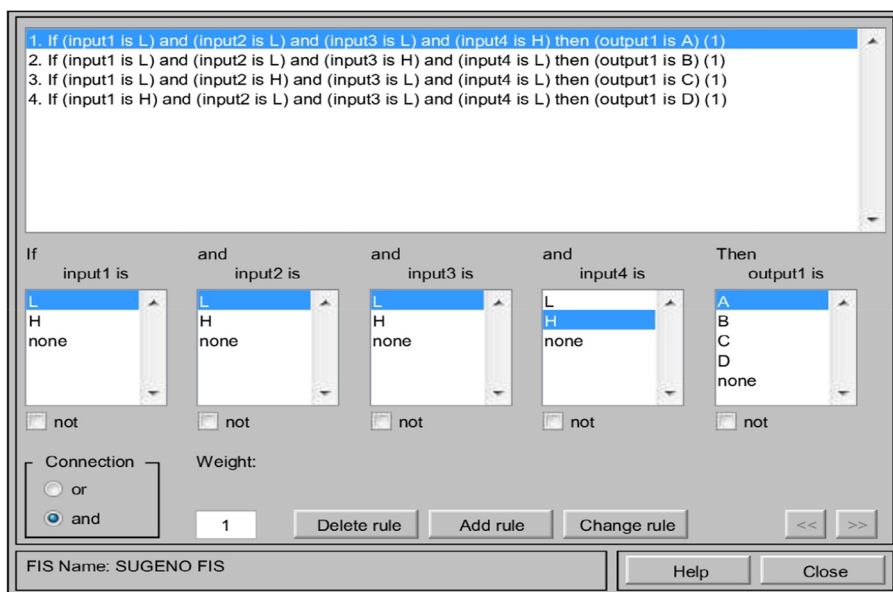
(b)

**Fig. 6** **a** Output of the MATLAB fuzzy rule designer. **b** output of the MATLAB fuzzy rule viewer and **c** output of the MATLAB fuzzy rule editor

puts of the BPNN neural network. To plot the membership function, 70% of the BPNN output was considered randomly. From each of the four outputs, mean and standard deviation were obtained for assigning the membership grades as shown

in Fig. 5a. Four simple rules were generated based on four numbers of input and four classes of rice, which are shown in Fig. 6c.

Fig. 6 continued



(c)

Table 4 Percentage classification accuracies based on morphology, colour and texture features

Rice type	BPNN				FUZZY				ANFIS				Cascade network			
	Different features				Different features				Different features				Different features			
	11	18	27	24	11	18	27	24	11	18	27	24	11	18	27	24
A	94	83	74	76	98	82	80	61	61	52	48	57	94	87	85	91
B	100	91	94	92	100	89	79	93	74	74	84	86	100	91	94	93
C	95	59	75	42	89	46	21	47	66	68	66	60	97	67	78	56
D	100	80	90	75	100	35	77	45	100	60	57	67	100	85	95	83
Rice type	KNN				LDA				NB							
	Different features				Different features				Different features							
	11	18	27	24	11	18	27	24	11	18	27	24				
A	91	82	67	66	90	81	81	71	93	80	53	57				
B	99	79	86	86	100	86	88	84	100	77	80	87				
C	94	68	59	50	95	72	79	57	92	64	62	33				
D	100	51	56	68	100	88	94	84	100	44	47	80				

### Results and discussion

The whole process of feature extraction and classification was carried out in MATLAB environment using MALTAB 2016a software. A dataset of 400 numbers (100 for each rice type) of feature vectors was created for each of the four feature sets, namely morphological, colour, texture and wavelet. The network is fed with 200 numbers of input patterns (50 patterns for each rice type) and the training parameters related to data division in MATLAB was set to 70:15:15 (training:testing:validation). The stopping criterion of the network was set based on the number of validation fail which is set to a default value of 6. This means the training stops when the validation fails continuously six times. The proposed cas-

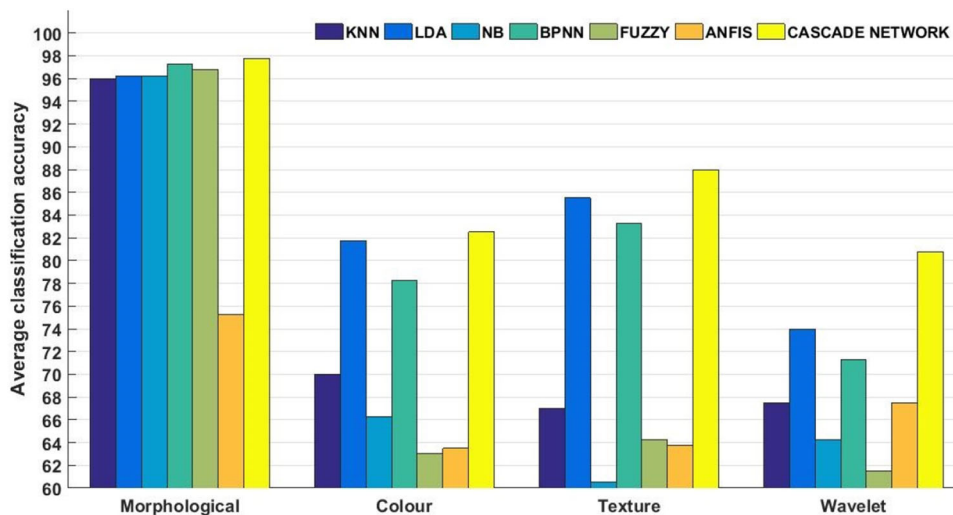
cade network is used for the classification of four varieties of rice grain based on four different features sets (11 morphological, 18 colour, 27 texture and 24 wavelet) and their classification accuracies based on different classifiers are presented in Table 4. It is found that in all the four classes, the classification accuracy based on cascade network is able to achieve better classification results as compared to other classifiers. The average classification accuracies based on all four feature sets are presented in Fig. 7. It is found that the proposed cascade classifier is able to yield better classification accuracy of 97.75% using morphological feature as compared to other features. The performance of the proposed classifier is also tested on selected features. The feature selection is performed using *stepdisc proc* in SAS software

**Table 5** Summary of feature selection using step disc procedure in SAS software

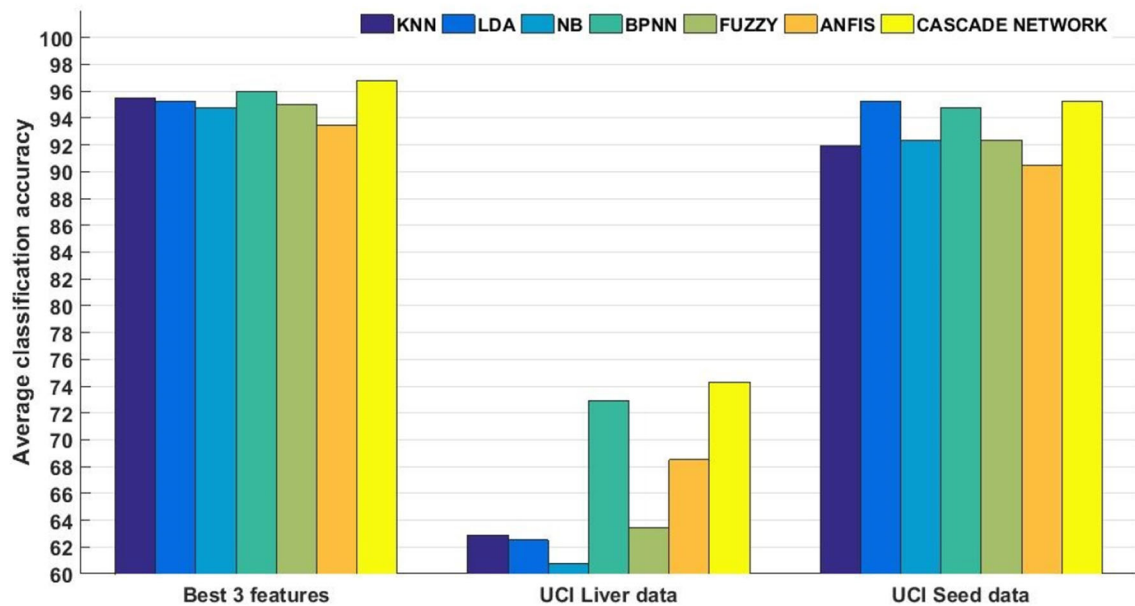
Feature no.	Name of features	F value	Average square canonical correlation (ASCC)
<b>Best 3 morphological features</b>			
F04	Solidity	1067.91	0.3069
F10	Minor axis	370.50	0.5723
F02	Convex area	55.77	0.6385
<b>Best 3 colour features</b>			
F2	Red—variance	157.82	0.2106
F7	Blue—mean	96.10	0.3812
F10	Hue—mean	21.72	0.4242
<b>BEST 3 texture features</b>			
F6	Red—contrast	46.97	0.2730
F3	Red—range	52.36	0.3784
F2	Red—variance	26.35	0.4107
<b>BEST 3wavelet features</b>			
F21	Blue—LLHL mean	129.91	0.1951
F17	Blue—LLLL mean	52.43	0.2964
F18	Blue—LLLL variance	7.44	0.3149
<b>Over all best 3 features</b>			
F1	Solidity	1093.24	0.3074
F2	Minor axis	463.78	0.5829
F3	Convex area	50.05	0.6469

**Table 6** Percentage classification accuracies based on overall best 3 features

Rice type	Classification based on 3 best features						
	KNN	LDA	NB	BPNN	FUZZY	ANFIS	Cascade network
A	89	90	87	92	90	87	94
B	99	96	99	99	99	100	100
C	94	95	93	93	91	87	93
D	100	100	100	100	100	100	100



**Fig. 7** Average classification accuracies based on different texture features



**Fig. 8** Average classification accuracies based on three selected features and UCI data

**Table 7** Percentage classification accuracies based on standard data set, University of California, Irvine (UCI)

Class	KNN	LDA	NB	BPNN	FUZZY	ANFIS	Cascade network
Classification based on UCI liver dataset							
A	54.28	38.57	85	64.82	52.41	51.03	67.58
B	71.50	86.50	36.50	81	74.40	86	81
Classification based on UCI seed dataset							
A	84.28	91.42	85.71	91.42	87.14	84.28	92.85
B	95.71	98.57	95.71	100	95.71	98.57	100
C	95.71	95.71	95.71	92.85	94.28	88.57	92.85

and the summary of feature selection process is presented in Table 5. Three best features for each of the four feature sets were selected (i.e. a total of 12 best features). The selection process is again applied to these 12 features and 3 final best features are selected. The classification accuracies based on over all best three features is presented in Table 6 and the average classification accuracies are presented in Fig. 8. It is found that the average classification accuracy of the proposed cascade classifier on three selected features is able to reach up to 96.75%. In most of the earlier research related to cereal grain classification, the performance of the classifier is judged according to the author's own dataset. The classifier so far reported in earlier works related to cereal grain classification was not tested against a standard data set. For generalisation of the proposed cascade classifier, the classification accuracy is tested against standard datasets from the University of California (UCI). The results of classification are presented in Table 7. Figure 8 shows the average classification accuracies using UCI database with different techniques. It can be seen from the figures that the proposed

classifier is able to provide better classification accuracy as compared to other classifiers on standard UCI data sets.

## Conclusion

Classification of four varieties of rice grains using different feature vectors, namely morphological, colour, texture and wavelet is carried out successfully in this paper. Out of these feature vectors, morphological feature yields better result (with an average classification accuracy of 97.75%) as compared to other features for the classification of rice grain based on single kernel. The performance of the proposed cascade classifier also holds well even with just three selected features with an average classification accuracy of 96.75%. It can also be concluded that the proposed cascade network classifier outperforms other classifiers in all feature data sets discussed in this work including standard data sets from the University of California, Irvine (UCI). Most of the cereal grain classifications tasks reported in the litera-

ture were based on colour image using a digital camera and the quality of the grains were determined based of physical factor, namely kernel size, kernel hardness, damaged kernel. The major challenge, however, is to integrate such a system that can explain the internal grain quality attributes which include sanitary factors (fungal infection, insect infestation, odour, etc.) and intrinsic factors (oil content, starch content, protein content, etc.). The scope of this work may be extended towards quality grading and evaluation of cereal grain using hyper-spectral imaging.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Pourreza A, Pourreza H, Hbbaspour-Fard MH et al (2012) Identification of nine Iranian wheat seed varieties by textural analysis with image processing. *Comput Electron Agric* 83:102–108
- Anami BS, Savaka DG, Makandar A et al (2005) A neural network model for classification of bulk grain samples based on colour and texture. *Proc Int Conf Cogn Recogn* 2005:359–368
- Gullu A, Ozan AKI, Ucar E (2015) Classification of rice grain using image processing and machine learning techniques. In: *International Scientific Conference, UNITECH 2015-Gabrovo*, pp II352–354
- Anami BS, Savakar DG (2011) Suitability of feature extraction methods in recognition and classification of grains, fruits and flower. *Int J Food Eng* 7(1):1–27
- Silva CS, Sonnadara U (2013) Classification of rice grains using neural networks. *Proc Tech Sess Inst Phys Sri Lanka* 29:9–14
- Choudhary R, Paliwal J, Jayas DS (2008) Classification of cereal grain using wavelet, morphological, colour and texture features of non-touching kernel. *Biosyst Eng* 99:330–337
- Douik A, Abdellaoui M (2010) Cereal grain classification by optimum features and intelligent classifiers. *Int J Comput Commun Control* 5(4):506–516
- Marini F, Bucci R, Magri AL et al (2008) Classification of 6 durum wheat cultivars from Sicily (Italy) using artificial neural network. *Chemometrics Intell Lab Syst* 90:1–7
- Guevara-Hernandez F, Gomez-Gil J (2011) A machine vision system for classification of wheat and barley grain kernels. *Span J Agric Res* 9(3):672–680
- Kaur H, Singh B (2013) Classification and grading rice using multi-class SVM. *Int J Sci Res Publ* 3(4):1–5
- Zariefo Roush H, Sacid M, Alizadesh MR et al (2015) A hybrid intelligent approach based on computer vision and fuzzy logic for quality measurement of milled rice. *Measurement* 66:26–34
- Olpour I, Parian JA, Chayjan RA (2014) Identification and classification of bulk paddy, brown, and white rice cultivars with colour features extraction using image analysis and neural network. *Czech J Food Sci* 32(3):280–287
- Pujari JD, Yakkundimath R, Byadgi AS (2014) Recognition and classification of normal and affected agriculture produce using reduced color and texture features. *Int J Comput Appl* 93(11):17–23
- Jayaraman S, Esakkirajan S, Veerakumar T (2009) *Digital image processing*. McGraw Hill, New Delhi
- Kilic K, Boyaci IH, KoKsel H et al (2007) A classification system for beans using computer vision system and artificial neural networks. *J Food Eng* 78:897–904
- Liu Z-Y, Cheng F, Ying Y-B et al (2005) Identification of rice seed varieties using neural network. *J Zhejiang Univ Sci* 6B(11):1095–1100
- Omid MD, Soltani M, Dehrouyeh MH, Mohtasebi SS, Ahmadi H (2013) An expert egg grading system based on machine vision and artificial intelligence techniques. *J Food Eng* 118:70–77
- Majundar S, Jayas DS (1999) Classification of bulk samples of cereal grain using machine vision. *J Agric Eng Res* 73:35–47
- Majundar S, Jayas DS (2000) Classification of cereal grain using machine vision. I. Morphology model. *Trans ASAE* 43(6):1669–1675
- Majundar S, Jayas DS (2000) Classification of cereal grain using machine vision. II. Colour model. *Trans ASAE* 43(6):1677–1680
- Majundar S, Jayas DS (2000) Classification of cereal grain using machine vision. III. Texture model. *Trans ASAE* 43(6):1681–1687
- Majundar S, Jayas DS (2000) Classification of cereal grain using machine vision. IV. Combined morphology, colour and texture model. *Trans ASAE* 43(6):1689–1694
- Mebatsion HK, Paliwal J, Jayas DS (2013) Automatic classification of non-touching cereal grain in digital image using limited morphological and colour features. *Comput Electron Agric* 90:99–105
- Siddagangappa Megha R, Kulkarni AH (2014) Classification and quality analysis of food grains. *IOSR J Comput Eng* 16(4):1–10
- Patil NK, Malemath VS, Yadahalli RM (2011) Colour and texture based identification and classification of food grains using different colour models and Haralick features. *Int J Comput Sci Eng* 3(12):3669–3679
- Castillo O, Melin P, Ramirez E et al (2012) Hybrid intelligent system for cardiac arrhythmia classification with fuzzy K-Nearest Neighbors and neural network combined with fuzzy system. *Expert Syst Appl* 39:2947–2955
- Paliwal J, Visen NS, Jayas DS (2001) Evaluation of neural network architecture for cereal grain classification using morphological features. *J Argic Eng Res* 79(4):361–370
- Paliwal J, Visen NS, Jayas DS et al (2003) Cereal grain and dockage identification using machine vision. *Biosyst Eng* 85(1):51–57
- Paliwal J, Visen NS, Jayas DS et al (2003) Comparison of a neural network and a non-parametric classifier for grain kernel identification. *Biosyst Eng* 85(4):405–413
- Melin P, Olivas F, Castillo O et al (2013) Optimal design of fuzzy classification system using POS with dynamic parameter adaptation through fuzzy logic. *Expert Syst Appl* 40:3196–3206
- Pazokia AR, Pazokia Z (2011) Classification system for rain fed wheat grain cultivars using artificial neural network. *African J Biotechnol* 10(41):8031–8038
- Pazoki AR, Farokhi F, Pazoki Z (2014) Classification of rice grain varieties using two artificial neural networks (MLP and Neuro-Fuzzy). *J Anim Plant Sci* 24(1):336–343
- Gonzalez RC, Woods RE (2012) *Digital image processing*. Pearson Prentice Hall, New Delhi



34. Grassi S, Casiraghi E, Alamprese C (2018) Fish fillet authentication by image analysis. *J Food Eng* 234:16–23
35. Sridhar S (2016) *Digital image processing*. Oxford University Press, New Delhi
36. Wah TN, San PE, Hlaing T (2018) Analysis on feature extraction and classification of rice kernels for Myanmar rice using image processing techniques. *Int J Sci Res Publ* 8(8):603–606
37. Visen NS, Paliwal J, Jayas DS et al (2004) Image analysis of bulk grain samples using neural network. *Can Biosyst Eng* 46:7.11–7.15

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.