International Journal of Applied Information Systems (IJAIS) – ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 3– No.2, July 2012 – www.ijais.org

Design of an Expert System for Rice Kernel Identification using Optimal Morphological Features and Back Propagation Neural Network

S.J. MousaviRad
Department of Computer Engineering
University of Kurdistan
Sanandaj,Iran

F. Akhlaghian Tab
Department of Computer Engineering
University of Kurdistan
Sanandaj, Iran

K. Mollazade
Department of Mech. Eng. of Agri. Machinery
University of Kurdistan
Sanandaj, Iran

ABSTRACT
In this paper, an algorithm for identifying five different varieties of rice, using the morphological features is presented. The proposed algorithm consists of several steps: image acquisition, segmentation, feature extraction, feature selection, and classification. Eighteen morphological features were extracted from rice kernels. The Set of features contained redundant, noisy or even irrelevant information so features were examined by four different algorithms. Finally six features were selected as the superior ones. A back propagation neural network-based classifier was developed to classify rice varieties. The overall classification accuracy was achieved as 98.4%.

General Terms
Pattern Recognition, Data Mining

Keywords
Rice kernel Identification, feature extraction, feature selection, neural networks, morphological feature

1. INTRODUCTION
Rice is one of the most important cereal grain crops. It constitutes the world's principle source of food, being the basic grain for the planet's largest population. For tropical Asians it is the staple food and is the major source of dietary energy and protein. In Southeast Asia alone, rice is the staple food for 80% of the population[1].

In the current grain-handling systems, grains variety and quality are assessed by visual inspection. This evaluation process is boring and time consuming. The decision-making capabilities of a grain inspector can be affected by physical condition such as tiredness, mental state caused by work pressure, and working condition such as inappropriate lighting condition, etc. The evaluation problem can be lightened by automation of the process by developing an imaging system that should acquire, rectify, and analyze the grain images.

During the last decades several studies have been carried out related to the application of machine vision for quality evolution, Zhao-yan et al (2005) proposed a method of identification based on neural network to classify rice variety using color and shape features with accuracy of 88.3% [2]. Mousavirad et al (2012) extracted seventy eight color and texture features from rice kernel. The classification accuracy was achieved 96.67% using neural network [3].In another work, Researchers presented an algorithm for classifying five different varieties of bulk rice using texture features with 95.4% classification accuracy[4]. In another research, Van Dalan (2004) developed a method for determination of the rice size and the amount of broken rice kernels using image analysis [5].

In a primary study Zayas et al., 1989 used machine vision to identify different varieties of wheat and to discriminate wheat from non-wheat components [6]. In later research Zayas et al., 1996 found that wheat classification methods could be improved by combining morphometry (computer vision analysis) and hardness analysis. Hard and soft recognition rates of 94% were achieved for the examined seventeen varieties [7]. In another report, the discrimination power of size, shape, color and texture for the identification of seeds of fifty seven weed species was assessed. Size and shape characteristics had larger discriminating power than color and texture ones. However, all of these features are required to reach an acceptable identification performance for practical applications [8]. In the extension of this work, they used a much larger database and discussed the discriminating power of weed seed characteristics[9].

In another work, morphological features of wheat grains were used to train and test neural networks with different combinations of nodes and iterations. Classification accuracy was about 88% for all of the wheat grains [10]. Morphological, color and wavelet features were extracted using different classifiers for classifying cereal grain. These parameters were tested on different classification methods and best classification gave the accuracy rate of 98% [11]. Majumdar and Jayas (2000) developed classification models by combining two or three feature sets (morphological, color, texture) to classify individual kernels of Canada western Red Spring (CWRS) wheat, Canada western ambur durum (CWAD) wheat, barley, oat and rye[12]. Shape variation, based on grain morphology, was quantified in fifteen Indian wheat varieties by Shouche et al., 2001. Geometric features such as area, perimeter, compactness, major and minor axis length and their ratios and slenderness were computed on the binary images. Then five other shape factors were derived from these basic geometric features [13].

In this paper, we propose a simple, effective and high accuracy vision-based approach using pattern recognition techniques to identify rice kernel varieties. The specific goal was to generate the optimal morphological features for classifying the rice varieties with high accuracy using comparing the performance of the some feature selection algorithms during the classification process.

2. MATERIALS AND METHODS

In this paper, a new approach for classification of rice kernels variety using Feed-Forward Neural network is presented. The block diagram shown in Fig. 1 illustrates the procedure for recognition and classification of rice grains. First, 8 bit images were obtained by a flatbed scanner. Then images were segmented by thresholding. In the next step, 18 morphological features were extracted from segmented images. After that superior features were determined through feature selection process. Finally these features were fed to a neural network classifier.

2.1 Image Acquisition

Image of rice kernels was acquired by a flatbed scanner used 8-bit grayscale and a resolution of 600 dpi. Images were stored in jpg format for further analysis (Fig. 2). For each variety, the measurements were conducted for 300 kernels of rice. Using a black background it was possible to easily extract the kernels from the background by standard segmentation routines.

2.2 Segmentation

Image segmentation subdivides an image into different parts or objects and is the first step in image analysis. The image is usually subdivided until the objects of interest are isolated from their background. There are generally two approaches for segmentation algorithms. One is based on the discontinuity of gray level values; the other is based on the similarity of the gray-level values. The first approach is to partition an image based on abrupt changes in gray levels. The second approach uses thresholding and region growing[2]. Region thresholding is an important part of image segmentation step. The threshold value is generated according to the results of the histogram analysis. This value is supposed to be constant in all images with the same lighting conditions.

In this study, it was determined that the gray values between the background and the rice kernels are very different. Therefore, a fixed threshold value defined from the histogram of the gray plane could separate the rice kernels from its background. A typical histogram of the gray plane of a rice kernels image is shown in Fig. 3. The rice kernels always had the area of gray levels higher than 110, while background pixels had the area of gray levels less than 110 with different shapes. The threshold value for this research was set at 110. All the pixels with gray value less than 110 were assigned the value 0, and all pixels with gray value equal or greater than 110 were not processed is further operations. The level 0 area was set for the background, and the unchanged area was set for the rice kernels region. Fig. 4 shows a typical segmented image.

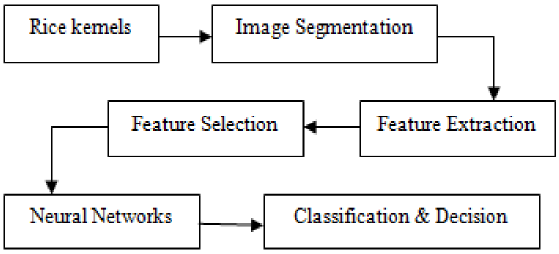

Fig 1. The procedure for classification of rice grain

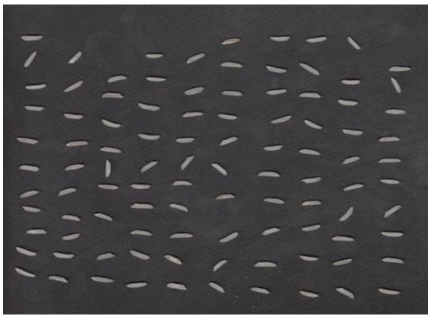

Fig 2. A sample of acquires images of rice kernels

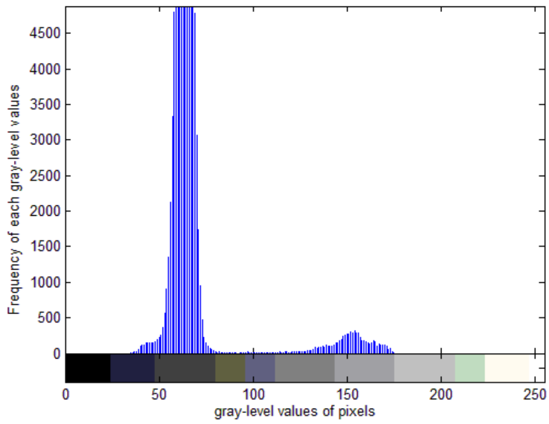

Fig 3. Histogram of gray level

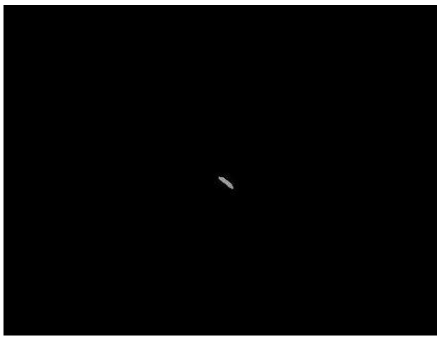

Fig 4. Image after segmentation

2.3 Morphological feature extraction

Algorithms were developed in windows environment using Matlab 7.10 programming languages to extract morphological features. The following morphological features were extracted from images of individual rice kernels:

- Area: the algorithm calculates the number of pixel inside and including the rice kernel boundary.
- Perimeter: the algorithm calculates the distance around the boundary of the rice kernels.
- The ratio of area to the sum of area and perimeter.
- The ratio of perimeter to the sum of the area and perimeter values.
- The ratio of area to perimeter.
- Major axis length: it is the distance between the end points of the longest line that could be drawn through the rice kernel.
- Minor axis length: it is the distance between the end points of the longest line that could be drawn through the rice kernels while maintaining perpendicularity with a major axis.
- Aspect ratio: the ratio of major axis length to the minor axis length.
- Eccentricity: It is the ratio of the distance between the foci of the ellipse and its major axis length.
- Solidity: it describes the extent to which the shape is convex or concave.
- Rectangularity: this represents how much the kernels have the rectangular shape, i.e. how much fills its minimum bounding rectangle:

$$\text{Rectangularity} = \frac{A_S}{A_R} \quad (1)$$

  Where A_s is the area of a shape and A_R is the area of the minimum bounding rectangle.

- Circularity: It is the ratio of a shape to the area of a circle having the same perimeter.
- Compactness: It is the ratio of the area of the shape to the area of a circle having the same perimeter.
- Equivalent Diameter: It specifies the diameter of a circle with the same area as the region.
- Also three other shape factors were derived from the basic features:
- Shape factor 1:the ratio of major axis length to area
- Shape factor 2: the ratio of minor axis length to area
- Shape factor 3: Area/(Major axis length*Minor axis length)

3. FEATURE SELECTION

From each rice kernels image, we measured 18 morphological features which can be used for classification stage. Of course this set of features contains redundant, noisy or even irrelevant information for classification purpose. To optimize the number of features that contributed significantly to the classification, we implemented four feature selection algorithms namely: branch and bound[14], standard sequential forward (SFS)[15], standard backward sequential (SBS)[15] and plus-l-takeaway-r [16] algorithms. The selection algorithm reduced the parameters to nearly optimal set of six morphological features which were finally used to build the classifier. The final six selected parameters, ranked based on their influence, are area by perimeter ratio, the ratio of area to the total of area and perimeter, aspect ratio, rectangularity, equivalent diameter and shape factor 3, respectively. Additionally, a study was performed to find the best method to classify the kernels of rice with the lowest classification error.

4. BACK PROPAGATION NEURAL NETWORK

A feed-forward neural network was trained for classification of the rice kernels. The number of neurons in the first layer was six which is equal to the number of input patterns vector. The number of neurons in the output layer was five, which is equal to the number of pattern classes the neural network has been trained to recognize them. In general, the neural network with too few hidden neurons will not have a sufficient capability to represent the input-output relationship accurately. Contrarily, the network with too many hidden neurons may leads to a problem of data over fitting, and affect the system’s generalization capability[17] .Hence, determining the optimal number of hidden neurons is a crucial step in designing classifiers. However, such a determination is not an easy task since it depends largely on experience and trial and error methods.

The validation process of dataset has been carried out using the mean square error of the classification result of different neural network structures. Structures of one hidden layer with 1 to 20 hidden nodes and a few structures of two hidden layers were tested. We used a structure of two hidden layers with 2 neurons in the first hidden layer and 5 neurons in the second hidden layers.

About the 70% of the samples (210 kernels for each rice type) were randomly selected as training set, while the rest of the samples were used as test set for classification. At the end of the training process, the network was tested with test dataset, and the classification accuracies were 100%, 100%, 96%, 100%, and 96% for Mahli, Neda, Gerde, Fajr, and Hashemi varieties, respectively

5. RESULTS AND DISCUSSIONS

Rice kernel of five Iranian varieties namely; Fajr, Neda, Hashemi, Mahali, and Gerde were taken up for classification. The objective of the presented analysis is to check morphological features which can classify rice kernels. In the first step, eighteen morphological features extracted. Then, selection of features was carried out.

Table 1. Result of Classification

Variety of Rice	Classification error (%)				
	Original Feature (n=18)	Feature selection(n=6)			
		SFS	SBS	B&B	plus-l-takeaway-r
Mahali	1.60	0.00	0.00	0.00	0.00
Neda	20.00	3.34	0.00	3.34	37.77
Gerde	33.33	0.00	4.45	39.00	39.00
Fajr	10.00	3.34	3.34	7.33	3.34
Hashemi	20.00	00.00	0.00	0.34	14.00

Six features characterized from the eighteen features. The selection was carrying out using four methods based on: Branch and bound, standard forward sequential, standard backward sequential and plus-l-takeaway-r algorithm. Classification error calculated for original data ranged from 1.6% to 33.33% .It was shown that data selection can distinguish varieties of rice with lower classification error. The result of feature selection showed that in the most analyzed cases the classification error was almost zero. The highest error of four methods was 40%. It also appears that the least suitable method for rice kernel classification is SFS and SBS. The SFS and SBS feature selection gives an error of classification equal with 0 to 4.45%. Table 1 shows the result of classification for five varieties of Iranian rice.

6. CONCLUSION

An algorithm was developed to identify varieties of rice kernels based on morphological features. Eighteen features were extracted and six superior features were selected with standard sequential forward and backward selection algorithm. A neural network was used to classify the rice kernels. In the test dataset, the total classification accuracies was 98.4%.

7. FUTURE WORKS

The present work can be extended for other food grains also color and texture features can be extracted to increase the accuracy rate. Also detection of various defects of rice kernels, like fissures, can be investigated.

8. REFERENCES

[1] V. Nguyen, 2000, "FAO Rice Information, Vol. 2. Food and Agriculture Organization," International Rice Commission, New York.

[2] Z. Liu, F. Cheng, Y. Ying, and X. Rao, 2005, "Identification of rice seed varieties using neural network," Journal of Zhejiang University. Science. B, vol. 6, p. 1095.

[3] S.J. Mousavirad, 2011, F. Akhlaghian Tab and K. Mollazade "Classification of rice varieties using optimal color and texture features and BP neural networks," presented at the The 7th Iranian Conferences on Machine Vision and Image Processing, 2011.

[4] S. MousaviRad, F. Akhlaghian Tab, and K. Mollazade, 2012, "Application of Imperialist Competitive Algorithm for Feature Selection: A Case Study on Bulk Rice Classification," International Journal of Computer Applications, vol. 40, pp. 41-48.

[5] G. Van Dalen, 2004 "Determination of the size distribution and percentage of broken kernels of rice using flatbed scanning and image analysis," Food research international, vol. 37, pp. 51-58.

[6] I. Zayas, Y. Pomeranz, and F. Lai, 1989, "Discrimination of wheat and nonwheat components in grain samples by image analysis," Cereal Chem, vol. 66, pp. 233-237.

[7] I. Y. Zayas, C. Martin, J. Steele, and A. Katsevich, 1996, "Wheat classification using image analysis and crush-force parameters," Trans. ASAE, vol. 39, pp. 2199-2204.

[8] P. M. Granitto, H. D. Navone, P. F. Verdes, and H. Ceccatto, 2002, "Weed seeds identification by machine vision," Computers and Electronics in Agriculture, vol. 33, pp. 91-103.

[9] P. M. Granitto, P. F. Verdes, and H. A. Ceccatto, "Large-scale investigation of weed seed identification by machine vision," Computers and Electronics in Agriculture, vol. 47, pp. 15-24, 2005.

[10] B. Dubey, S. Bhagwat, S. Shouche, and J. Sainis, 2006, "Potential of artificial neural networks in varietal identification using morphometry of wheat grains," Biosystems engineering, vol. 95, pp. 61-67.

[11] A. Douik, M. Abdellaoui, and E. N. I. de Monastir, 2010, "Cereal Grain Classification by Optimal Features and Intelligent Classifiers," International Journal of Computers Communications & Control, vol. 5, pp. 506-516.

[12] S. Majumdar, 1997, "Classification of cereal grains using machine vision,".

[13] S. Shouche, R. Rastogi, S. Bhagwat, and J. K. Sainis, 2001, "Shape analysis of grains of Indian wheat varieties," Computers and Electronics in Agriculture, vol. 33, pp. 55-76.

[14] P. M. Narendra and K. Fukunaga, 1977, "A branch and bound algorithm for feature subset selection," Computers, IEEE Transactions on, vol. 100, pp. 917-922.

[15] P. A. Devijver and J. Kittler, Pattern recognition: A statistical approach: Prentice/Hall International, 1982.

[16] S. D. Strearns, 1976, "On selecting features for pattern classifiers," presented at the The Third International Conference of Pattern recognition.

[17] R. O. Duda, P. E. Hart, and D. G. Stork, "Pattern classification," New York: John Wiley, Section, vol. 10, p. l, 2001.