Classification of Rice Grains Using Neural Networks

Chathurika Sewwandi Silva and Upul Sonnadara
Department of Physics, University of Colombo, Colombo 3
email: upul@phys.cmb.ac.lk

ABSTRACT

This paper presents a neural network approach for classification of rice varieties. A total of 9 different rice verities were considered for the study. Samples were drawn from each variety and images of seeds were captured. Algorithms were developed to extract thirteen morphological features, six colour features and fifteen texture features from colour images of individual seed samples. A different neural network models were developed for individual feature sets and for the combined feature set. High classification accuracy was given by textural features than morphological and colour features. An overall classification accuracy of 92% was obtained from combined feature model. Individual classification accuracies of AT307, BG250, BG358, BG450, BW262, BW267, W361, BW363 and BW364 were 94%, 98%, 84%, 100%, 94%, 68%, 98%, 94% and 94% respectively. It was noted that different neural network architectures tend to produce different accuracies.

1. INTRODUCTION

The accurate identification of rice seeds is very important when classifying rise verities. The identification of the level of purity of rice varieties makes the identification task more difficult and complicated [1]. Commercial value, genetic characteristics and quality depend on the rice variety type. The grade and price of rice is decided by these factors. However, even a trained human can perform quality examination only on a few known rice varieties [2].

Machine vision is an emerging technology for rapid identification of grains. With the use of digital images and digital video, machine vision is widely used in many different fields in recent years [3]. Due to the computational power and memory of modern computers, machine vision systems can be used for online inspection of agricultural products [4]. A machine vision system can be effectively used for physical quality parameters of the grain at terminal elevators (grain handling facility) [5]. However, the natural diversity of different varieties of cereal grains makes classification task by machine vision a complex work [6]. Image analysis based on texture, morphology and colour features of grains is essential in this area. The morphological characters of grains are heritable in nature [7] and play an important role in variety identification [8]. When both colour and morphological features are combined results are much more accurate. In addition, texture features can be added to improve the classification accuracies. Classification accuracies are very high when different features of the tested varieties are used [9]. Artificial neural networks have many advantages over fuzzy classifiers and statistical classifiers [5]. Back-propagation neural network is the most popular choice for classification of agricultural products [10].

The main objective of this work was to investigate the feasibility of classifying local rice varieties using different feature sets.

2. METHODOLOGY

2.1 Rice Seed Samples

Two hundred and fifty (250) grams of grain samples from 9 rice varieties were selected as experimental seed materials. These verities were grown in different zones in Sri Lanka. They were obtained from the Regional Agricultural Research and Development Center, Bombuwela. The verities considered in this research work are; BW 267-3 (Low Country Wet Zone), BW 262-6B (Low Country Wet Zone), BW 363 (General Cultivation), AT 307 (General Cultivation), AT 306 (General Cultivation), BG 450 (General Cultivation), BG 358 (General Cultivation), BW 364 (Wet zone), BW 361 (Dry and Intermediate zone) and BG 250 (Rain fed and flash flood areas). Figure 1 shows the images of rice varieties used for this study.

Figure1: Rice Varieties

First, the rice grains were cleaned to remove impurities and samples from each variety were drawn randomly for image acquisition. In order to access each and every seed in the image, non-touching seeds were photographed. About 50 grains from each verity were chosen for the analysis. Out of 450 grain sample, 70% seeds were taken for the training, 15% for testing and 15% for validation.

2.2 Image Acquisition

The image acquisition was carried out with a Sony DSC-W270 digital camera. The camera was mounted on a fix stand. The software development and the image processing were carried out on a personal computer (2.1 GHz i3 processor with a 4 GB of RAM). A desk lamp with fluorescent light source was used in all experiments for lighting the area to be photographed. A light beam from a desk lamp was directed on seeds to avoid formation of shadows.

A black surface was used as a background. Images were taken by spreading the seeds on the black background. The fluorescent bulb was switched on for few minutes before taking photographs for stability. The camera was white balanced according to the light source. Seeds were placed in random orientation at variety of position inside the field view. Grains were arranged in non-touching pattern and images were acquired and stored for later analysis.

2.3 Image Analysis and Feature Extraction

Initially, Gaussian filter was used to enhance the original colour image by reducing noise. Next the colour image was converted to gray scale image. Morphological opening was used to estimate the background. Image subtraction was applied to remove the background. In order to have complete contrast between seeds and the rest of the area of the image, contrast stretching was applied. The gray scale image was converted to a binary image by threshold technique. Next dilation and erosion were done on resultant image. In order to observe only the required grain objects, all connected components (objects) that have very few pixels were removed from the binary image, producing a new binary image. All seeds in the new image were labelled to give unique identification. Morphological features such as length, width, rectangular aspect ratio, convex area, eccentricity, extent, filled area, major axis length, minor axis length and solidity were extracted from the labelled image.

In order to extract colour features, colour image was processed first by applying a filter and then background was subtracted to improve the clarity of the objects. Then, RGB pixel values from colour images of individual grains were chosen and R, G, B values were separated into 3 different arrays. Average of R, G, B, H, S and I were calculated as colour features.

Texture analysis refers to the characterization of regions in an image by their texture content. The images of individual grain were cropped from each colour channel (red, green, blue) using the bounding box coordinates specified from labelled image. Next the images of individual grain were given to the gray level co-occurrence matrix (GLCM). Number of gray levels in GLCM was set to the gray level value in the image used (In this study it was either 2 or 8). The spatial relationships of pixels that were defined by array of offsets, where d = 4 and 0^0, 45^0, 90^0, 135^0 directions were specified. Next properties of GLCM such as contrast, homogeneity, correlation and entropy were measured. Also the images of individual grain were given to the gray level run length matrix (GLRLM). Properties of GLRLM are; Shot Runs Emphasis, Long Runs Emphasis, Gray Level Non-uniformity, Run Length Non-uniformity, Run Percentage, Low Gray Level Run Emphasis, High Gray Level Run Emphasis, Short Run Low Gray-Level Emphasis, Short Run High Gray Level Emphasis, Long Run Low Gray Level Emphasis and Long Run High Gray Level Emphasis.

2.4 Classification Analyses

Neural network models were designed and developed using Matlab toolbox. Initially individual neural network models were created for each feature set (colour, morphology and texture) separately. Then combination of feature set model was implemented. In order to reduce the dimension of the input feature set, principal component analysis was applied.

Following the work carried out in a previous study [11], the number of nodes in the hidden layer was calculated using the following equation,

$$n = (I + O)/2 + Y\hat{}0.5$$

where *n* is the number of nodes in hidden layer, *I* is the number of input nodes, *O* is the number of output nodes, and *Y* is the number of input patterns in the training set.

Two hidden layers were used in all networks. Thus, the calculated number of nodes was equally divided between two hidden layers of the network. For all neural networks used in this work, a common structure was used (i.e., neural networks consisted of 4 layers). With this approach, the MLP (Multi Layer Perception) consisted of 4 layers (input layer, two hidden layers and output layer) with an output layer having 9 nodes and input layer having nodes equal to the number of features.

3. RESULTS AND DISCUSSION

3.1 Individual Feature Model

The morphological model of MLP with 13 input nodes, 2 hidden layers with 9 nodes each and output layer with 9 nodes (13-9-9-9) produced an overall classification accuracy of 33%. Many morphological features were highly correlated with one another and many of them did not contribute significantly to the morphology model. When the optimal features were given to the network (11-7-7-9), the overall accuracy was increased to 44%. Colour model of MLP (5-5-5-9) produced an overall classification accuracy of 51%.

For the texture, 3 separate models were built to test the features extracted from the red, green and blue channels. Results of the textural features extracted from the red, green and blue colour bands (at gray level value 8) with the model of MLP (15-10-10-9) were 61%, 59% and 55% respectively. Textural features extracted from images with red colour band gave the highest overall accuracy. Results of the optimal textural features (at gray level value 8) which were extracted from the red colour band with the model of

MLP (15-10-10-9) gave slightly increased overall classification accuracy of 63% (which was 61% previously).

Both binary and 8 bit gray level images were used to determine how gray level value influences the classification accuracy with the model of MLP (15-10-10-9). Overall classification accuracy was reduced to 34% for binary images.

Results obtained from different feature models could be summarized as shown in Figure 2. Here only the models with optimum feature set were considered. For example, in texture feature model, only the results for features obtained from 8 bit gray level red colour channel are shown.

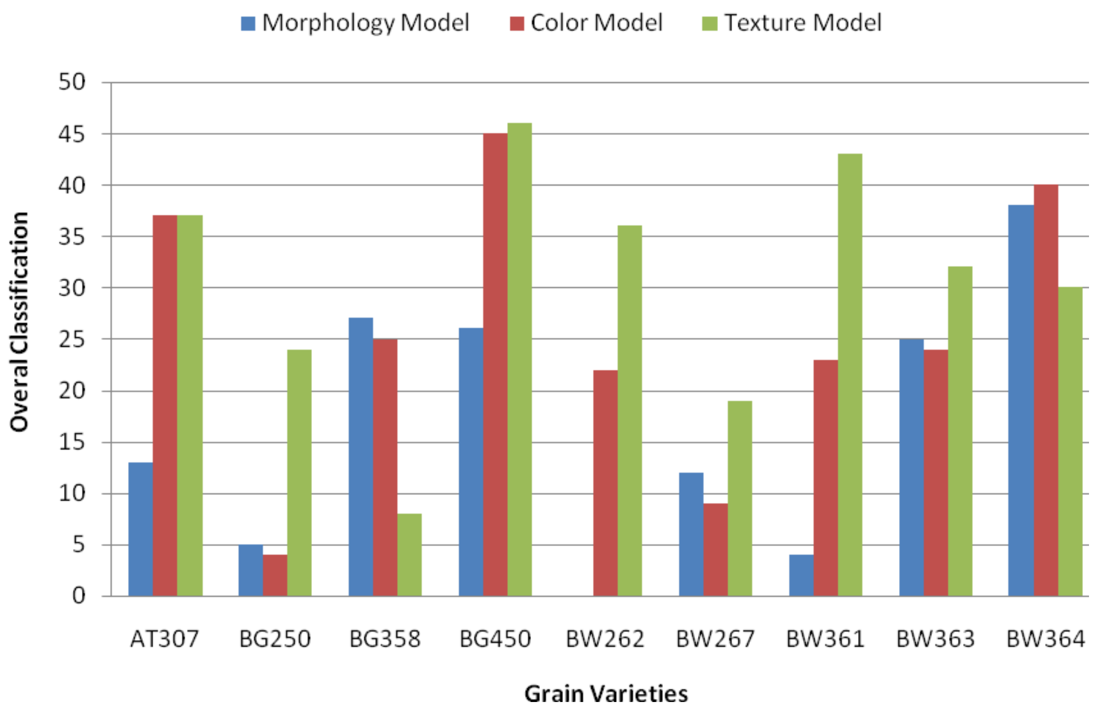

Figure 2: Comparison of morphology, colour and textural models for each rice verity

3.2 Combined Feature Model

Although overall classification of colour model was greater than that of morphology model, for some varieties (BG250, BG358 and BW 267), texture model or the morphology model outperforms the colour model. Similarly, the texture model which produced the highest classification accuracy for BG450, underperform for some varieties (BG358 and BW364) compared to both morphology and colour models. A variation of performance is clearly seen for three models build on separate features morphology, colour and texture. Therefore, a combined feature model was developed to extract the best performance to identify rice varieties.

In the combined feature model, all 3 feature sets (morphology, colour and texture) were combined together. Principal component analysis was used to eliminate input features which gave minimum contribution to the combined feature model. The combined feature model produced classification accuracies of 94%, 98%, 84%, 100%, 94%, 68%, 98%, 94% and 94% for AT 307, BG 250, BG 358, BG 450, BW 262, BW 267, W 361, BW 363 and BW 364 respectively with the overall classification accuracy of 92%.

4. CONCLUSIONS

This work presents classification accuracies for rice seeds obtained through a machine vision combined with neural network architecture. Out of three feature sets, texture features produced high classification accuracy. Especially texture features obtained from red colour band produced better predictions. Improved classification accuracies were obtained when the network trained with optimal data set. The combined feature model produced the overall classification accuracy of 92%.

The preliminary work presented in this paper could be further enhanced by focusing on different sampling methods, sample sizes, sample pre-processing techniques, different features and different neural network models to match the requirements of the rice industry.

REFERENCES

[1] Fang, C., Xiu-qin, R., Zhao-yan, L. and Yi-bin, Y., *Identification of rice seed verities using neural network*, J. Zhejiang University Science B. 6 11, (2005) 1095-1100

[2] Guzman, J.D. and Peralta, E.K., *Classification of philippine rice grains using machine vision and artificial neural networks, in world conference on agricultural information and IT*, IAALD AFITA WCCA, Tokyo, Japan, (2008) 41-48.

[3] Bebis, G., Egbert, D. and Sham, M., *Review of computer vision education*, IEEE Transactions on Education 46, (2003) 2-21

[4] Choudhary, R., Paliwal, J. and Jayas, D.S., *Classification of cereal grains using wavelet, morphological, colour and textural features of non touching kernels image*, Biosystems Engineering. 99 3, (2008) 330-337

[5] Majumdar, S. and Jayas, D.S., *Classification of cereal grains using machine* vision: Morphology models, Transactions of ASABE. 43 6, (2000) 1669-1675

[6] Douik, A. and Abdellaoui, M., *Cereal verities classification using wavelet techniques combined to multi-layer neural networks*, Mediterranean Conference on Control and Automation. 16, (2008) 1822-1827

[7] Harper, J.L. and Lovell, K.G., *The shapes and sizes of seeds*, Annual Review of Ecology and Systematics. 1, (1970) 327-356.

[8] Shouche, S.P., Rastogi, R., and Bhagwat, S.G., *Shape analysis of grains of Indian wheat verities*, Computers and Electronics in Agriculture. 33 1, (2001) 55-76

[9] Paliwal, J., Visen, N.S. and Jayas, D.S., *Evaluation of Neural Network Architectures for Cereal Grain Classification Using Morphological Features*, J. Agricultural Engineering Research. 79 4, (2001) 361-370

[10] Jayas, D.S., Paliwal, J. and Visen, N.S., *Multi-layer Neural Networks for Image Analysis of Agricultural Products*, J. Agricultural Engineering Research. 77 2, (2000) 119-128

[11] Kannur, A.I., Kannur, A. and Rajpurohit, V., *Classification and grading of bulk seeds using artificial neural network*, Int. J. Machine Intelligence. 3 2, (2011) 62-73