Classifying the Shape of Aggregate using Hybrid Multilayered Perceptron Network

ARIFFUDDIN JORET[1], NOR ASHIDI MAT ISA[2], AHMAD NAZRI ALI[3], KAMAL ZUHAIRI ZAMLI[4], KHAIRUN AZIZI AZIZLI[5], M. SUBHI M. AL BATAH[6]

[1,2,3,4,6]*School of Electrical and Electronic Engineering*
Universiti Sains Malaysia, Engineering Campus
14300 Nibong Tebal, Seberang Perai Selatan, Pulau Pinang, MALAYSIA
[1]*Tel: +604-5937788 ext 6072, Fax: +6045941023*
[2]*Tel: +604-5937788 ext 6051, Fax: +6045941023*
[3]*Tel: +604-5937788 ext 6020, Fax: +6045941023*
[4]*Tel: +604-5937788 ext 6790, Fax: +6045941023*
[6]*Tel: +604-5937788 ext 6072, Fax: +6045941023*

[5]*School of Mineral and Material Resources Engineering*
Universiti Sains Malaysia, Engineering Campus
14300 Nibong Tebal, Seberang Perai Selatan, Pulau Pinang, MALAYSIA
Tel: +604-5937788 ext 6100

Abstract
In concrete production, shape of aggregate reflects the quality of concrete produced. The well-shaped aggregates are said to produce high quality concrete by reducing water to cement ratio. On the contrary, poor-shaped aggregates often require higher water to cement ratio in concrete production. Conventionally, the quality of concrete is determined by calculating the ratio of well-shaped aggregate to poor-shaped aggregate contained in concrete. This procedure is slow, highly subjective and laborious, which is inefficient and expensive. In order to decrease the problems, this paper proposed an intelligent classification system for the aggregates using neural network. The system uses Zernike moments, Hu's moment invariants, area and perimeter of the aggregate's mass and boundary as input data for the neural network. The HMLP which is trained using MRPE algorithm, has been used as the classification system. The system produced 85.53% accuracy. This shows that the HMLP network has high capability to be used as intelligent shape classification system for aggregates.

Key-Words:- Neural Networks, HMLP, MRPE, Aggregate Classification, Artificial Intelligent System.

1 Introduction

Aggregates produced after crushing process can be divided into six types of shape, which are Angular, Cubical, Elongated, Flaky, Flaky&Elongated and Irregular as shown in Fig.1. From these six shapes, the aggregates can be divided further into two categories, the well-shaped aggregate (i.e. Angular and Cubical) and the poor-shaped aggregate (i.e. Elongated, Flak, Flaky&Elongated and Irregular).

The aggregate's shape has significant effects on the quality of the concrete produced. The used of well-shaped aggregates will enhance the overall quality of the concrete due to the reduction of water to cement ratio. The used of well-shaped aggregates can decrease the cost of production and placement of concrete and increase its workability [1].

As a major factor in the production of high quality concrete, the aggregates need to be classified before used. The conventional classification method of the aggregates is done manually. The method is slow, highly subjective and laborious. The method can be replaced by using digital image processing technique [2] and neural network, known as automatic classifier.

As neural networks provide better generalization, robustness and parallel implementation paradigm properties, it becomes a popular choice in object recognition [3][5][7][9]. In [4], Mashor has introduced the Hybrid Multilayered Perceptron network (HMLP) that has been proven to significantly improve the performance of multilayered perceptron network.

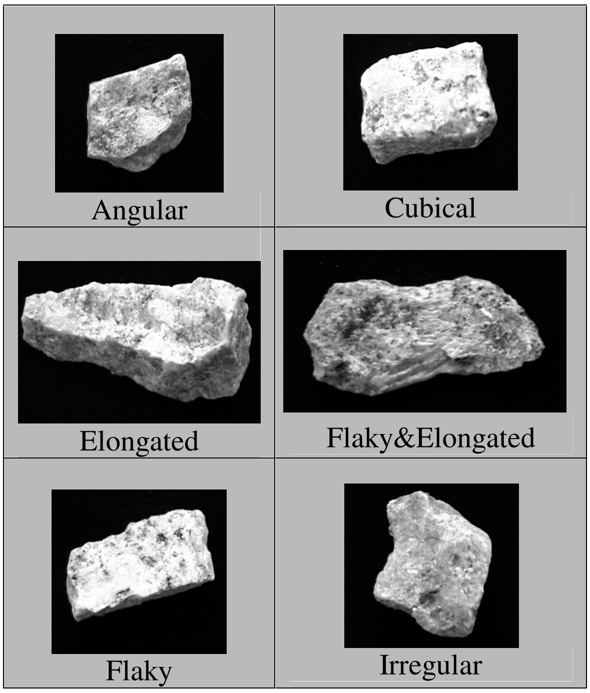

Fig.1 Aggregates Types

The HMLP network has been applied in pattern recognition as the classification system. In [3], Mashor *et al.* had used the HMLP network to classified 3D object using 2D moment. Their proposed system achieved the recognition accuracy of up to 100%. For the cervical cancer application with image processing, Mat-Isa *et al.* [5] had applied the HMLP network in cervical cancer classification. The HMLP had successfully determined each type of cervical cells correctly with high percentage in both training and testing phase.

2 Hybrid Multilayered Perceptron Network

In this study, the HMLP network is proposed as shape classification for aggregate. Consider HMLP network with one hidden layer as shown in Fig.2.

The equation for the HMLP network is given by [4]:

$$\hat{y}_k(t) = \sum_{j=1}^{n_h} w_{jk}^2 F\Big(\sum_{i=1}^{n_i} w_{ij}^1 x_i^0(t) + b_j^1\Big) + \sum_{i=0}^{n_i} w_{ik}^l x_i^0(t)$$

for $1 \le k \le m$ (1)

where w_{jk}^2 denotes the weights of the connections between the hidden and output layers. w_{ij}^1 denotes the weights that connect the input and the hidden layers. w_{ik}^l denotes the weights of extra connections between the input and output layers. x_i^0 and b_j^1 denote the inputs that are supplied to the input layer and thresholds in hidden nodes respectively. n_i , m and n_h are the number of input nodes, output nodes and hidden nodes respectively. $F(\bullet)$ is an activation function that is selected as sigmoidal function. The $w_{jk}^2, w_{ik}^l, w_{ij}^1$ and threshold b_j^1 are unknown and have been selected to minimize the prediction error, defined as:

$$\varepsilon_k(t) = y_k(t) - \hat{y}_k(t) \qquad (2)$$

where $y_k(t)$ and $\hat{y}_k(t)$ are the actual and network outputs respectively.

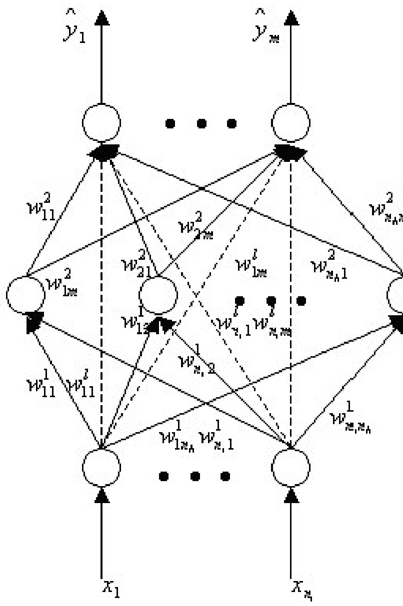
Fig.2 One-hidden layer HMLP network

3 Modified Recursive Prediction Error

Refer to Equation (1), the values of w_{ij}^1, w_{jk}^2, w_{ik}^l and b_j^1 must be determined using appropriate algorithm. Back Propagation (BP) algorithm is commonly used to find optimum values for those parameters. Although the algorithm is easy to be implemented and produces a good performance, but its convergence rate is slow. To overcome the problems, Chen *et al.* [6] proposed recursive prediction error (RPE) to replace the BP algorithm. The RPE algorithm provides a faster convergence rate and better final convergence values of weights and thresholds. In [4], Mashor proposed a modified version of RPE algorithm, known as modified recursive prediction error (MRPE). By optimising the way the momentum and the learning rate are assigned, the MRPE algorithm is able to improve the convergence rate of the RPE algorithm.

The RPE algorithm modified by Chen *et al.* [6] minimizes the following cost function:

$$J\left(\hat{\Theta}\right)=\frac{1}{2N}\sum \varepsilon^{T}\left(t,\hat{\Theta}\right)\Lambda^{-1}\varepsilon\left(t,\hat{\Theta}\right) \tag{3}$$

by updating the estimated parameter vector, $\hat{\Theta}$ (consists of *w*s and *b*s), recursively using the Gauss-Newton algorithm:

$$\hat{\Theta}(t)=\hat{\Theta}(t-1)+\mathrm{P}(t)\Delta(t) \tag{4}$$

and

$$\Delta(t)=\alpha_m(t)\Delta(t-1)+\alpha_g(t)\psi(t)\varepsilon(t) \tag{5}$$

where $\varepsilon(t)$ and Λ are the prediction error and an $m\times m$ symmetric positive definite matrix respectively, and m is the number of output nodes; $\alpha_m(t)$ and $\alpha_g(t)$ are the momentum and the learning rate respectively. $\alpha_m(t)$ and $\alpha_g(t)$ can be arbitrarily assigned to some values between 0 and 1, and the typical values of $\alpha_m(t)$ and $\alpha_g(t)$ are closed to 1 and 0 respectively. In the present study, $\alpha_m(t)$ and $\alpha_g(t)$ are varied to improve further the convergence rate of the RPE algorithm according to:

$$\alpha_m(t)=\alpha_m(t-1)+a \tag{6}$$

and

$$\alpha_g(t)=\alpha_m(t)\left(1-\alpha_m(t)\right) \tag{7}$$

where a is a small constant (typically $a=0.01$); $\psi(t)$ represents the gradient of the one-step-ahead predicted output, $\hat{y}$ with respect to the network parameters:

$$\psi(t,\Theta)=\left[\frac{d\,\hat{y}(t,\Theta)}{d\Theta}\right] \tag{8}$$

P(t) in equation (4) is updated recursively according to:

$$P(t)=\frac{1}{\lambda(t)}[P(t-1)-P(t-1)\psi(t)\left(\lambda(t)I+\psi^{T}(t)P(t-1)\psi(t)\right)^{-1}\psi^{T}(t)P(t-1)] \tag{9}$$

where $\lambda(t)$ is the forgetting factor, $0<\lambda(t)<1$, and has been updated using the following scheme:

$$\lambda(t)=\lambda_0\lambda(t-1)+(1-\lambda_0) \tag{10}$$

where λ_0 and the initial forgetting factor, $\lambda(0)$ are the design values. The initial value of the *P(t)* matrix, *P(0)* is set to αI where I is the identity matrix and α is a constant, typically between 100 and 10000.

The gradient matrix, $\psi(t)$ can be modified to accommodate the extra linear connections for a one-hidden-layer HMLP network model by differentiating equation (1) with respect to the parameters, θ_c, to yield:

$$\psi_k(k)=\frac{d\hat{y}_k(t)}{d\theta_c}=\begin{cases} u_j & \text{if} \quad \theta_c=w^2_{jk} \quad 1\le j\le n_h \\ x_i & \text{if} \quad \theta_c=w^l_{ik} \quad 0\le i\le n_i \\ u_j(1-u_j)w^2_{jk} & \text{if} \quad \theta_c=b^1_j \quad 1\le j\le n_h \\ u_j(1-u_j)w^2_{jk}x_i & \text{if} \quad \theta_c=w^1_{ij} \quad 1\le j\le n_h, 1\le i\le n_i \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

The modified RPE algorithm for a one-hidden-layer HMLP network can be implemented as follows [4]:

1. Initialize weights, thresholds, *P(0)*, *a*, *b*, $\alpha_m(0)$, λ_0 and $\lambda(0)$. (*b* is a design parameter that has a typical value between 0.8 and 0.9).
2. Present inputs to the network and compute the network outputs according to equation (1).
3. Calculate the prediction error according to equation (2).
4. Compute matrix $\psi(t)$ according to equation (11). Note that, elements of $\psi(t)$ should be calculated from the output layer down to the hidden layer.
5. Compute matrix *P(t)* and $\lambda(t)$ according to equations (9) and (10) respectively.
6. If $\alpha_m(t)<b$, update $\alpha_m(t)$ according to equation (6).
7. Update $\alpha_g(t)$ and $\Delta(t)$ according to equations (7) and (5) respectively.
8. Update parameter vector $\hat{\Theta}(t)$ according to equation (4).
9. Repeat steps (3) to (9) for each training data sample.

4 Methodology and Data Samples

To determine the suitability of the HMLP network for classifying the shapes of aggregates, the HMLP network needs to go through training and testing phases. During both phases, the optimum structure and the performance of the HMLP network were determined. The performance analysis of the HMLP network is based on three important characteristics, which are overall performance, accuracy of correct determination of well-shaped aggregates and accuracy of correct determination of poor-shaped aggregates. During the training phase, the weighs and the bias of the network are calculated according to the MRPE algorithm. Based on these values, the performance of the network are analysed to obtain the suitable values for the testing phase.

In search of the optimum structure, the HMLP performance has been analysed for the optimum number of hidden nodes and epochs. In this case, the hidden nodes analysis was conducted to maximize performance of the HMLP network by making the epochs constant. The number of hidden nodes giving the best performance is then selected as the optimum number of hidden nodes. Then, by fixing the optimum number of hidden nodes based on the aforementioned process, the epochs analysis was conducted in a similar fashion to obtain the optimum number of epochs.

In this study, 1666 images of aggregate were use to be classified into well-shaped or poor-shaped aggregates. 900 images were used as training data while the remaining 766 images were used as testing data. From 900 images for training data, 457 were the well-shaped aggregates while 443 were the poor-shaped aggregates. As for the testing data, 334 data were the well-shaped and 332 data were the poor-shaped aggregates.

The HMLP network has been trained using 17 hidden nodes and 28 epochs. The output nodes of the network are 2, which represent the 2 types of aggregate. The inputs of network are 8, which are Zernike moment calculated for mass and boundary, first and second Hu's moment invariants calculated for mass and boundary, area and perimeter. The Hu's moment invariants were selected for first and second order since the higher order is more sensitive to noise [3]. For the Zernike moment, the value used was the sum of the moment calculated from order 0 to 4. The Zernike values have been combined since it gave the better clustering results.

The designing parameters for MRPE were selected as $\alpha_m(0) = 0.00$, $\alpha(t)_g = \alpha_m(t)(1-\alpha_m(t))$, $a = 0.01$, $b = 0.85$, $\lambda_0 = 0.99$, $\lambda(0) = 0.95$ and $P(0) = 1000I$.

5 Results and Discussions

Table 1 and 2 show the recognition performance of the proposed system for training and testing phases respectively. The result was produced after the HMLP network was trained using 28 epochs and 17 hidden nodes.

Table 1 Results for aggregate's shape classification in training phase.

Classification	True	False	Total	Performance
well-shaped	388	69	457	84.90%
poor-shaped	413	30	443	93.23%
Overall	801	99	900	89.00%

Table 2 Results for aggregate's shape classification in testing phase.

Classification	True	False	Total	Performance
well-shaped	341	43	384	88.80%
poor-shaped	283	99	382	74.08%
Overall	624	142	766	81.46%

From table 1, the results show that the HMLP network classified the poor-shaped aggregates better than the well-shaped aggregates in the training phase, with 93.23% and 84.90% accuracy respectively. For overall performance, the HMLP network produced 89.00% of accuracy.

The results obtained in Table 2 show that the HMLP network successfully classified 88.80% of well-shaped aggregate correctly as compared to 74.08% of poor-shaped aggregate during the testing phase. The different is about 14.72%. For overall performance, the HMLP network classified 81.46% of the aggregate correctly. The promising results obtained show that the HMLP network has high capability to classify the aggregates into two classes, namely well-shaped and poor-shaped aggregates.

In addition, this study also analyses the effect of the MRPE parameters by conducting several different set of those parameters. The different sets used are as follows.

Set 1:
$\alpha_m(0) = 0.00$, $\alpha(t)_g = \alpha_m(t)(1-\alpha_m(t))$, $a = 0.01$, $b = 0.85$, $\lambda_0 = 0.99$, $\lambda(0) = 0.95$ and $P(0) = 1000I$.

Set 2:
$\alpha_m(0) = 1.00$, $\alpha(t)_g = \alpha_m(t)(1-\alpha_m(t))$, $a = 0.01$, $b = 0.90$, $\lambda_0 = 0.99$, $\lambda(0) = 0.95$ and $P(0) = 10000I$.

Set 3:
$\alpha_m(0) = 0.50$, $\alpha(t)_g = \alpha_m(t)(1-\alpha_m(t))$, $a = 0.01$, $b = 0.80$, $\lambda_0 = 0.99$, $\lambda(0) = 0.95$ and $P(0) = 100I$.

The results showed that those set of parameters produced the same results as shown in Table 1 and Table 2. This shows that the MRPE algorithm is stable and does not depend much on its parameters.

6 Conclusion

In this paper, the intelligent shape classification for aggregate was proposed and evaluated. The system used the HMLP network with the MRPE algorithm as the training algorithm. The features used as input data for the HMLP network were the combination of two types of moments, i.e. Hu's moment invariants and Zernike moments, area and perimeter. Using these features, the system achieved accuracy as high as 85.53%. The system's accuracy however can be improved. In order to improve the performance, the features selected can be considered. Since the features are the important part in pattern recognition, in future, this study will seek the suitable features for the system.

References

[1] B. Hudson, The Effect of Manufactured Aggregate and Sand Shape on Concrete Production and Placement, Svedala New Zealand Limited, 1995, pp. 1-15.

[2] R. C Gonzalez, R. E. Woods, *Digital Image Processing,* New Jersey: Prentice Hall, 2002.

[3] M. Y. Mashor, M. K. Osman, and M. R. Arshad, 3D Object Recognition Using 2D Moments and HMLP Network, *In IEEE* Proceeding of the International Conference *on Computer Graphics, Imaging and, Visualization*, 2004, pp. 126-130.

[4] M. Y. Mashor, Hybrid multilayered perceptron networks. *International Journal of Systems Science* Vol. 31 No.6, 2000, pp. 771-785.

[5] N. A. Mat-Isa, M. Y. Mashor, and N. H. Othman, Classification of Cervical Cancer Cells Using HMLP Network with Confidence Percentage and Confidence Level Analysis, *International Journal of The Computer, The Internet and Management,* Vol.11, No.1, 2004, pp. 17-29.

[6] S. Chen, C. F. N. Cowan, S. A. Billings, and P. M. Grant, A parallel recursive prediction error algorithm for training layered neural networks. *International Journal of Control,* Vol.51, No.6, 1990, pp. 1215-1228.

[7] M. K. Osman, M. Y Mashor, and M. R. Arshad, 3D Object Recognition System Using Multiple Views and Cascaded Multilayered Perceptron Network, *In IEEE Proceedings of the conference on Cybernatics and Intelligent Systems*, *2004, pp.1010-1014.*

[8] M. K. Osman, M. Y. Mashor, and M. R. Arshad, An Optimised Camera-Object Setup for 3D Object Recognition System, *In Proceedings of the Second International Conference on Artificial Intelligence in Engineering & Technology*,2004, pp.267-273.

[9] Y. Zhu, L. C. De Silva, and C. C. Ko, Using moment invariant and HMM in facial expression recognition, *Pattern Recognition letters,* Vol.23, pp.83-91.

[10] Y. Bin, and P. Jia-Xiong, Invariance analysis of improved Zernike moments, *Journal of Optics A: Pure and Applied Optics* Vol.4, pp. 606-614.