Sprague Dawley Rat Sperm Classification Using Hybrid Multilayered Perceptron Network

MOHD FAUZI ALIAS [1], NOR ASHIDI MAT ISA [2], SITI AMRAH SULAIMAN [3] and MAHANEM MOHAMED [4]

[1, 2] School of Electrical & Electronic Engineering,
Universiti Sains Malaysia, Engineering Campus,
14300 Nibong Tebal, Penang,
MALAYSIA.
[1]Email: fauzialias@msi.unikl.edu.my.
[2]Email: ashidi@eng.usm.my.

[3, 4] School of Medical Science,
Universiti Sains Malaysia, Health Campus,
16150 Kota Bharu, Kelantan.
MALAYSIA.
[3]Email: sbsamrah@kb.usm.my
[4]Email: mahaneem@kck.usm.my

Abstract: - As of now, the analysis of sperm such as counting and detection processes are still operated manually which tends to produce errors. False detection in sperm analysis must be minimized as possible. Therefore, the current study focuses on developing a Sprague Dawley rat sperm classification system to assist the detection process by pathologist. The system has the ability to classify the Sprague Dawley rat sperm into two classes namely normal and abnormal based on the morphological characteristic of sperm's head. The proposed system employs the Hybrid Multilayered Perceptron (HMLP) neural network trained with the Modified Recursive Prediction Error (MRPE) algorithm as the intelligent classifier tool. This study will also investigate three significant morphological characteristics of rat sperm's head namely opened degree, width of the curve and template matching percentage as the input data for the HMLP network. Furthermore, this study further classifies the abnormal sperm into hookless and banana shape. Promising result with 100% and 94.62% of accuracy has been achieved for two classes and three classes classification of rat sperm respectively.

Key-words: - HMLP, MRPE, Sprague Dawley rat, Opened degree, Width of the curve, Matching percentage.

1 Introduction

A lot of researches have been done in medical field which focus on the detection and classification of specific organs and features. Previously, detection and classification processes were manually done by human expert which concise a lot of work and time. According to [1]-[4], errors are still produced by the pathologist in the diagnosis and classification results. Furthermore, second time false diagnosis and a death case of patient caused by the error diagnosis from pathologist results have also been recorded [5]. Therefore, the reability of the medical analysis is fully depended on the accuracy of these processes.

Rat sperm cell has always been used in the previous researches. One of the objectives in using rat sperm cell as sample analysis is to observe the medicine and chemical effects on the rat body before it can be implemented into human body [6]-[8]. Thus, numerous drug discoveries will be tested on rat before they can be approved and tested on human. The same procedure is applied for determination and classification of normal and abnormal sperms. One of the main issues in sperm classification is difficulty in determining the type of sperm due to its tiny size. Therefore, this paper proposes a system to detect and classify rat sperm into two main classes namely normal and abnormal. Furthermore, this paper will also further classify the abnormal sperm into hookless abnormal and banana shape abnormal.

For the system development, the Hybrid MultiLayared Perceptron (HMLP) network trained by the Modified Recursive Prediction Error (MRPE) algorithm has been proposed as the rat sperm

classifier. As inputs for the HMLP network, three features from rat sperm have been recognized and proposed in this paper named as opened degree, width of the curve and the template matching percentage.

2 Sprague Dawley Rat Sperm Image

Rat sperm image has been widely used as the research sample in medical imaging analysis for many years. The main reason is the anatomy similarity between rat and human body. Numerous researches have been done using various sperms especially rat sperm such as male fertility, content of protein, calcium response etc. [6]. For human safety, rat sperm has been used in the experimentation to avoid unexpected results on human body.

Sperm is classified as a reproductive cell in male. The head of sperm contains genetic information and influent the entire system of the sperm. In this paper, Sprague Dawley rat species has been selected for the research samples. The Sprague Dawley rat sperm is divided into three parts namely head, mid piece and tail as shown in figure 1 [9]. The sperm's head is approximately 2.5µm long [9].

For the classification analysis, rat sperms are divided into two main classes namely normal and abnormal. For the abnormal class, rat sperms can further be divided into sub-classes. Based on CASA [9], rat sperms abnormalities are caused either by its head or tail. To be focus, this paper only covers the abnormalities caused by the head of rat sperm. For a normal rat sperm, the head resembles a hook as shown in figure 2. For abnormal rat sperm, there are several shapes such as banana shape, hookless, double headed and others.

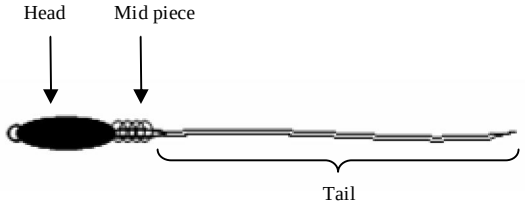

Fig. 1. The morphological of rat sperm.

Fig. 2. The normal sperm image captured under 400x magnification using digital microscope.

For the analysis purpose, the rat sperm images have been taken randomly from the samples and captured manually under 400x magnification. Based on the image analysis, each image has been captured in the best quality in term of the clear distinguishable between sperm and its background. The reason is because the shape characteristic is the only way to define the abnormalities of the rat sperm. Therefore, any missing information from rat sperm shape could lead to the wrong decision in the sperm classification.

3 Hybrid Multilayered Perceptron (HMLP) Network

Hybrid MultiLayered Perceptron (HMLP) network was introduced by [10] in 1999 to improve the performance of the conventional MultiLayered Perceptron (MLP) network. The conventional MLP is the nonlinear model. The HMLP network is a modified version of the MLP network and it covers both the linear and nonlinear models. Several researches have applied the HMLP network in their researches and proved to obtain better performance compared to other neural networks such as the Radial Basis Function (RBF) and the MLP [11], [12].

According to [11], the HMLP network is capable to improve the MLP performance as well as generate Mean Squared Error (MSE) as good as RBF network at the small numbers of hidden nodes. [14] applied the HMLP network for breast cancer classification as well as compared the HMLP result with the MLP and RBF networks. In [14], the HMLP and RBF networks were trained using Modified Recursive Prediction Error (MRPE) and Linear Least Square (LLS) respectively. The MLP network was trained by five algorithms namely Quasi-Newton, Gradient Descent with Momentum and Adaptive, Resilient Back Propagation, Levenberg Marquardt and Recursive Prediction Error (RPE). According to the research, the HMLP network trained by the MRPE algorithm produced the highest performance with 100% accuracy, sensitivity and specificity compared to other neural networks.

The HMLP network has also been implemented by [15] on cervical cell classification. The HMLP network was trained with the MRPE algorithm. The research concluded that the HMLP network produced the best performance by producing the high accuracy, sensitivity and specificity percentages. Besides, the HMLP network produced

low false negative and false positive. Furthermore, the HMLP network has also been used by [16] in the cervical cancer diagnosis application. The research proved the HMLP network is capable in producing high accuracy with 89.23%. [16] classified the cervical cell into three classes namely normal cell, low-grade squamos intraepithelial lesion (LSIL) and high-grade squamos intraepithelial lesion (HSIL).

Other researches which applied the HMLP network as intelligent classifier tool are [17] and [18]. [18] applied the HMLP network to recognize 3D images from two groups of images. One of the groups contains cylinder, cubic, round and other images and the other groups contains free shape images. As a result, the research successfully produced 100% of accuracy for the 3D images classification.

3.1 The HMLP Network Architecture

Based on the conventional MLP architecture, [10] proposed new linear connections to the architecture in creating linear and nonlinear model. This model was named as the HMLP network as shown in figure 3. Based on figure 3, the HMLP network contains there layers namely input, output and hidden layers. As compared to the conventional MLP network, a linear connection between the input nodes and output nodes (as represented by dotted lines) are introduced. The direct propagation of input data from input layer to output layer will form a linear model.

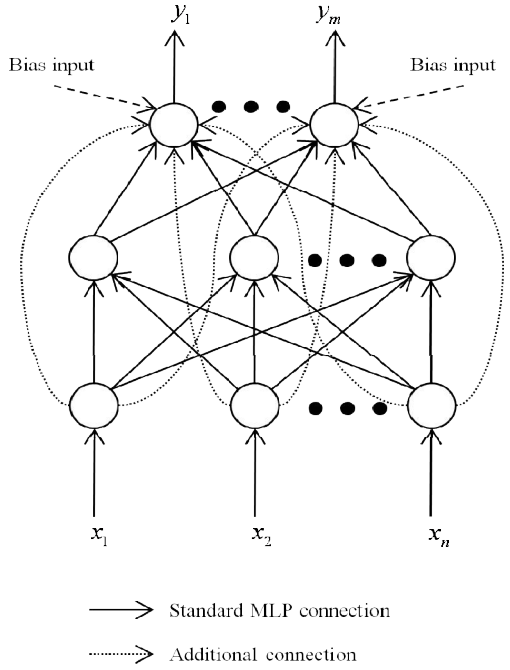

Fig. 3. HMLP network architecture.

Based on figure 3, all input data, x_1 are directed to the hidden nodes via weighted connection. Each node operates as perceptron model activated by the sigmoid function. Output of the *j*-th hidden node, u_j is represented by:

$$u_j(t) = \varphi\left(\sum_{i=1}^{n_j} w_{ij}^1 x_i(t) + b_j^1\right); \quad for\ 1 \le j \le n_h \tag{1}$$

where;

$\varphi(\bullet)$ represents the activation function.

Output from hidden node is directed to each output's nodes via weight's connection.

The HMLP network output $\hat{y}_k$ is given by.

$$\hat{y}_k(t) = \sum_{j=1}^{n_h} w_{jk}^2 \left(\sum_{i=1}^{n_i} w_{ij}^1 x_i^0(t) + b_j^1\right) + \sum_{i=1}^{n_i} w_{ik}^l x_i^0(t) \tag{2}$$
$$for\ 1 \le k \le m$$

where;

w_{ik}^l represents the weights of the connection between input and output layers.

w_{jk}^2 represents the weights of the connection between hidden and output layers.

w_{ij}^1 represents the weights of the connection between input and hidden layers.

b_j^1 represents the thresholds in the hidden nodes.

x_i represents as inputs to the input layer.

$F(\bullet)$ is an activation function and is commonly be selected as sigmoid function. Based on the equation 2, the right hand side of the equation is the addition part to the conventional MLP equation. The addition part represents the direct connection between input and output nodes. The weights of this connection is given as $w^l{}_{ik}$. According to the equation 2, when $i = 0$, $w^l{}_{ik}$ and x_i represent connection weight and input for the bias input network respectively.

3.2 The Weights of The Network Connection

In this paper, the HMLP network is trained using the Modified Recursive Predection Error (MRPE) algorithm for the classification purpose. The MRPE is a modified version of the Recursive Prediction Error (RPE) algorithm. According to [12], the RPE training algorithm is not very compatible for the linear system model. [12] proved that the implementation of RPE algorithm in the HMLP network could not improve the network performance. Therefore, the RPE algorithm has been modified to optimize its momentum and learning rate. The modification of the RPE

algorithm is called the MRPE. The implementation of the MRPE algorithm in this paper is based on the references [10], [12] and [15].

Generally, the RPE algorithm was implemented by [20] to minimize the cost function as shown in equation 3.

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

By updating the estimated parameter vector, $\hat{\Theta}$ (consists of *ws* and *bs*), recursively using Gauss-Newton algorithms as shown in equations 4 and 5.

$$\hat{\Theta}(t)=\hat{\Theta}(t-1)+P(t)\Lambda(t) \quad (4)$$

and

$$\Lambda(t)=a_m(t)\Lambda(t-1)+a_g(t)\psi(t)\varepsilon(t) \quad (5)$$

where;

$\varepsilon(t)$ represents prediction error,

Λ represents the $m\times m$ symmetric positive definite matrix which m is the number of output nodes,

$a_m(t)$ represents momentum,

$a_g(t)$ represents the learning rate.

$a_m(t)$ and $a_g(t)$ are arbitrarily assigned to some values between 0 and 1. The typical values of $a_m(t)$ and $a_g(t)$ are closed to 1 and 0 respectively. $a_m(t)$ and $a_g(t)$ values are varied to improve the convergence rate of RPE algorithm. $a_m(t)$ and $a_g(t)$ equations are shown as equations 6 and 7 respectively.

$$a_m(t)=a_m(t-1)+a \quad (6)$$

and

$$a_g(t)=a_m(t)\left(1-a_m(t)\right) \quad (7)$$

where a represents a small constant typically 0.01. $\psi(t)$ represents the gradient of the one-step-ahead predicted output, $\hat{y}$ with respect to the network parameters as shown in equation 8.

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

$P(t)$ in equation 4 is updated recursively according to equation 9.

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

where $\lambda(t)$ represents the forgetting factor and its value range is between 0 and 1. The value of $\lambda(t)$ is updated using equation 10.

$$\lambda(t)=\lambda_0\lambda(t-1)+(1-\lambda_0) \quad (10)$$

where;

λ_0 represents the initial forgetting factor.

The initial value of the *P(t)* matrix, *P(0)* is set to αI. I and α represent the identity matrix and a constant respectively. The typical value of α is between 100 and 10000. The value of α is usually 1000. The gradient matrix $\psi(t)$ is calculated using equation 11.

$$\psi_k(k)=\frac{dy_k(t)}{d\theta_c}$$

$$=\begin{cases}u_j & if\quad \theta_c=w_{jk}^2;\ 1\le j\le n_h\\ x_i & if\quad \theta_c=w_{ik}^1;\ 0\le i\le n_i\\ u_j(1-u_j)w_{jk}^2 & if\quad \theta_c=b_j^1;\ 1\le j\le n_h\\ u_j(1-u_j)w_{jk}^2x_i & if\quad \theta_c=w_{ij}^1;\ 1\le j\le n_h,1\le i\le n_i\\ 0 & \text{otherwise}\end{cases} \quad (11)$$

The steps of the MRPE algorithm implementation into HMLP network can be described as follows [12];

1. Initialize weights, thresholds, $P(0),a,b,\alpha_m(0),\lambda_0$ and $\lambda(0)$. (b is a design parameter that has a typical value between 0.8 dan 0.9).
2. Present inputs to the network and compute the network outputs according to equation 2.
3. Calculate the prediction error according to equation 12.

$$\varepsilon_k(t)=y_k(t)-\hat{y}_k(t) \quad (12)$$

where;

$y_k(t)$ represents the actual output.

4. Compute matrix $\psi(t)$ using equation 11. (Elements of $\psi(t)$ should be calculated from the output layer down to the hidden layer).
5. Compute matrixs $P(t)$ and $\lambda(t)$ according to equations 9 and 10 respectively.
6. If $a_m(t)<b$, update $a_m(t)$ using equation 6.
7. Update $a_g(t)$ and $\Delta(t)$ using equations 7 and 5 respectively.
8. Update parameter vector $\hat{\Theta}(t)$ according to equation 4.
9. Repeat steps (2) to (8) for each training data sample.

Design parameter, b in the step 6 is the upper limit of the momentum. The momentum value will increase for each data sample (from smaller value)

until it reachs the value of *b*. The small value is always closed to 0.

4 The Proposed Features for Rat Sperm Classification

Based on the Sprague Dawley rat sperm image, this paper proposes three features extracted from the rat sperm as the neural network inputs. Based on the observation, the head's sperm shape characterictic plays an important role in classifying the different types of Sprague Dawley rat sperm. Shape characteristic that appear in the object contain a lot of information and important for object recognition process [21].

4.1 Opened Degree Feature

Based on the obvious shape differences between normal and abnormal sperm's head, first feature namely opened degree is proposed. The opened degree feature is refered to the deep value of the sperm head's bending. This value is calculated based on the degree unit as shown in figure 4. According to the sperm's head shape, the normal sperm has deeper curve compared to the abnormal sperm. Theoretically, the deeper curve generates smaller opened degree value. Therefore, the normal sperm's head should generate smaller opened degree value compared to the abnormal sperm's head.

Figures 4 and 5 show the image sample Dawley01 and the feature extraction algorithm for opened degree respectively. $\alpha°$ represents the value of opened degree and calculated by applying trigonometry's method as shown in figure 5.

Equations 13 and 14 have been applied to calculate the value of $\alpha°$ [22].

$$\beta = \cos^{-1} \frac{(n_x \times n_y) + (m_x \times m_y)}{\sqrt{((n_x^2 + n_y^2) \times (m_x^2 + m_y^2)}} \quad (13)$$

$$\alpha° = \frac{\beta}{\pi} \times 180° \quad (14)$$

where;

m_x represents the substrate outcome between x_{center} and x_1, $(x_{center} - x_1)$.

m_y represents the substrate outcome between y_{center} and y_1, $(y_{center} - y_1)$.

n_x represents the substrate outcome between x_{center} and x_2, $(x_{center} - x_2)$.

n_y represents the substrate outcome between y_{center} and y_2, $(y_{center} - y_2)$.

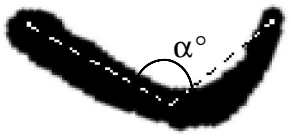

Fig. 4. Opened degree feature extraction using image Dawley01 as an example.

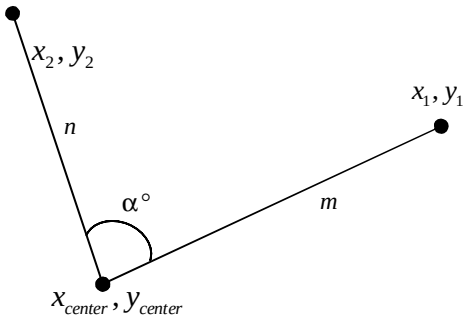

Fig. 5. Opened degree feature architecture.

Based on the observation of normal and abnormal sperms, the different in opened degree values is quite clear and distinguishable. However, comparison between hookless and banana shape abnormal sperms produce slightly different values of opened degree. Figures 6, 7 and 8 show the example analysis of opened degree for image samples Dawley01, Dawley02 and Dawley03 respectively. Image Dawley01, Dawley02 and Dawley03 are taken from sample normal, hookless abnormal and banana shape abnormal classes respectively.

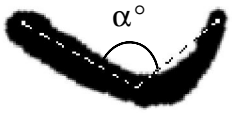

Fig. 6. Opened degree value for image Dawley01, $\alpha° = 93.18°$.

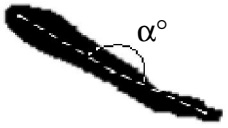

Fig. 7. Opened degree value for image Dawley02, $\alpha° = 174.94°$.

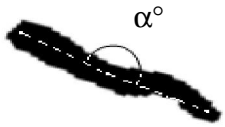

Fig. 8. Opened degree value for image Dawley03, $\alpha^\circ = 172.80^\circ$.

Based on figures 6 to 8, the normal and both abnormal sperm samples show an obvious difference for the opened degree values. However, the opened degree values for hookless and banana shape abnormals show a small different. Nevertheless, the HMLP network has been expected to produce high accuracy of classification using opened degree feature as one of the network inputs.

4.2 Width of the Curve Feature

The second feature selected to be the HMLP network input is the width of the curve. The width of the curve refers to the length between two center points as shown in figure 9. Both center points are measured from the head sperm's length and the direct line between two edges respectively. The measurement for this feature is shown in figure 10.

Fig. 9. Width of the curve feature extraction using image Dawley01 as an example.

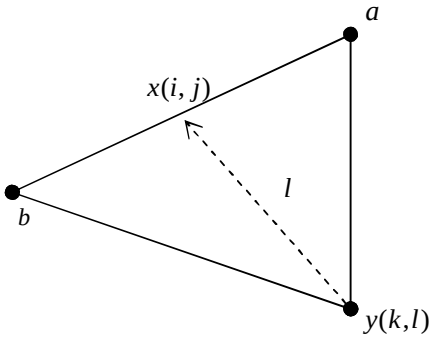

Fig. 10. Width of the curve feature architecture.

Based on figure 10, the width of the curve, l is calculated using equation 15 [23].

$$l = \left[(l-j)^2 + (k-i)^2 \right]^{\frac{1}{2}} \quad (15)$$

where;

l represents the distance between points x and y.

(i, j) represents the coordinate for point x.

(k, l) represents the coordinate for point y.

x represents the center point on the straight line between two edges of sperm. Both edges of sperm are symboled as a and b respectively.

y represents the center point of the sperm length where the head sperm length is between a and b.

This research encounters the width of the curve values for normal and abnormal sperms produce different range of values based on the different classes. As result, normal sperm produces larger value of the curved width compared to the abnormal sperm. Image samples Dawley01, Dawley02 and Dawley03 have been selected for the example analysis to obverse the feature's capability as an input for the HMLP network. The results of the analysis are shown as figures 11, 12 and 13 for Dawley01, Dawley02 and Dawley03 respectively.

Fig. 11. Width of the curve value for image Dawley01, $l = 33.24$.

Fig. 12. Width of the curve value for image Dawley02, $l = 4.12$.

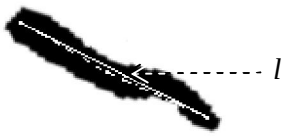

Fig. 13. Width of the curve value for image Dawley03 $l = 7.21$.

According to figures 11 to 13, the widths of the curve values produce big difference between normal and abnormal sperms. Eventhough the width of the curve values for hookless and banana shape abnormals are varied in a small amount, but this research has expected the HMLP network could still produce good performance for the sperm classification.

4.3 Template Matching Percentage Feature

As the last input for the HMLP network, this paper proposes the percentage value produced by the template matching process. The percentage values are obtained by the image processing technique called template matching method using cross correlation algorithm. The template matching method will be briefly discussed in this paper.

The cross correlation algorithm has been used in the template matching method. The cross correlation equation is shown as equation 16 [24], [25].

$$Out(i,j) = \sum_{k=0}^{K-1}\sum_{l=0}^{L-1} In1(k,l)\ In2(i+k, j+l) \qquad (16)$$

where;

In1 represents the input image,

In2 represents the template image,

Out(i,j) represents the output image.

K and *L* are defined as columns and rows respectively for the input image (using pixel as unit). *In1, In2* and *Out(i,j)* are defined as 2D images. The matching process is uninterrupted until all pixels in the input image have been matched with the template image.

For the matching process, an ideal normal rat sperm has been selected as the template image. This template image will be the reference image for rat sperm sample in the matching process. Therefore, normal sperms should have higher percentage compared to the abnormal sperms. This percentage is selected as one of the HMLP network's input for the rat sperm classification.

To show the capability of the template matching percentage as input data to the HMLP network to differentiate between normal and abnormal sperms, images Dawley01, Dawley02 and Dawley03 have been used as the tested images. The results obtained are tabulated in Table 1. Based on Table 1, the normal and abnormal sperms produce different matching percentages. Thus, this paper proposes this feature as one of the input for the HMLP network.

Table 1: The result of matching percentage for images Dawley01, Dawley02 and Dawley03.

Image sample	Matching percentage (%)
Dawley01	98.99
Dawley02	80.71
Dawley03	79.84

5 Results and Discussions

In order to observe the capability of the HMLP network on rat sperm image classification into normal and abnormal classes, 469 samples from the Sprague Dawley rat sperm images have been tested. 275 and 194 from the samples are normal and abnormal respectively. The 5-fold analysis method has been used based on [26]. For the HMLP performance analysis, this paper uses the accuracy, specificity and sensitivity percentages analysis. The equations of accuracy, specificity and sensitivity are shown in equations 17, 18 and 19 respectively.

$$accuracy = \frac{A+Y}{Total\ sperm\ sample} \times 100\% \qquad (17)$$

$$specificity = \frac{A}{A+B} \times 100\% \qquad (18)$$

$$sensitivity = \frac{Y}{X+Y} \times 100\% \qquad (19)$$

where;

A represents number of normal sperms correctly classified as normal

B represents number of normal sperms incorrectly classified as abnormal

X represents number of abnormal sperms incorrectly classified as normal

Y represents number of abnormal sperms correctly classified as abnormal

This paper divides the classification into two parts. The first part classifies the rat sperm images into two classes namely normal and abnormal sperms. The second part classifies the rat sperm images into three classes namely normal, hookless and banana shape abnormals. For the two classes classification, the data quantity divisions for training and testing phases are tabulated in Tables 2 and 3 respectively. For the three classess classification, the data quantity divisions for training and testing phases are tabulated in Tables 4 and 5 respectively.

The HMLP network analysis for the sperm image classification applied three features as its inputs namely opened degree, width of the curve and template matching percentage. Generally, the application of neural network as a classifier employs three steps as reported in [27] and [28]. The steps are determination of optimum epoches, determination of optimum hidden node for the neural network and analyzation of the classification performance for each neural network.

Table 2: Amount of sperm images for training phase using 5-fold analysis

5-fold	Number of image samples	
	Normal sperm	Abnormal sperm
1st fold	225	151
2nd fold	228	148
3rd fold	227	149
4th fold	226	150
5th fold	225	151

Table 3: Amount of sperm images for testing phase using 5-fold analysis

5-fold	Number of image samples	
	Normal sperm	Abnormal sperm
1st fold	50	43
2nd fold	47	46
3rd fold	48	45
4th fold	49	44
5th fold	50	43

Table 4: Amount of sperm images for training phase using 5-fold analysis

5-fold	Number of image samples		
	Normal sperm	Hookless abnormal sperm	Banana shape abnormal sperm
1st fold	226	142	8
2nd fold	226	142	8
3rd fold	228	140	8
4th fold	226	142	8
5th fold	227	141	8

Table 5: Amount of sperm images for testing phase using 5-fold analysis.

5-fold	Number of image samples		
	Normal sperm	Hookless abnormal sperm	Banana shape abnormal sperm
1st fold	49	39	5
2nd fold	49	39	5
3rd fold	47	41	5
4th fold	49	39	5
5th fold	48	40	5

The results of accuracy, specificity and sensitivity for two classes classification for both training and testing phases are tabulated in Tables 6 and 7 respectively. Tables 8 and 9 represent the results of classification into three classes namely normal, hookless and banana shape abnormals for training and testing phases respectively. Table 10 shows the average percentages of accuracy, specificity and sensitivity for training and testing phases for two and three classes classifications.

Table 6: Results for training phase (i.e. classification into normal and abnormal classes)

5-fold	Training phase		
	Specificity	Sensitivity	Accuracy
1st fold	99.11%	100%	99.73%
2nd fold	98.25%	100%	99.47%
3rd fold	97.80%	100%	99.47%
4th fold	99.12%	100%	99.73%
5th fold	98.22%	100%	99.47%
Average	98.50%	100%	99.57%

Table 7: Results for testing phase (i.e. classification into normal and abnormal classes)

5-fold	Testing phase		
	Specificity	Sensitivity	Accuracy
1st fold	100%	100%	100%
2nd fold	100%	100%	100%
3rd fold	100%	100%	100%
4th fold	100%	100%	100%
5th fold	100%	100%	100%
Average	100%	100%	100%

Table 8: Results for training (i.e. classification into normal, hookless abnormal and banana shape abnormal classes)

5-fold	Training Phase		
	Specificity	Sensitivity	Accuracy
1st fold	98.67%	95.33%	97.61%
2nd fold	98.67%	93.33%	97.61%
3rd fold	99.12%	94.59%	97. 34%
4th fold	99.12%	94.67%	97.61%
5th fold	99.12%	93.96%	97.34%
Average	98.94%	94.38%	97.50%

Table 9: Results for testing phase (i.e. classification into normal, hookless abnormal and banana shape abnormal classes)

5-fold	Testing phase		
	Specificity	Sensitivity	Accuracy
1st fold	100%	88.64%	94.62%
2nd fold	100%	88.64%	94.62%
3rd fold	97.87%	89.13%	94.62%
4th fold	100%	88.64%	94.62%
5th fold	100%	88.89%	94.62%
Average	99.57%	88.79%	94.62%

Table 10: Average percentage of accuracy, specificity and specificity for two and three classes classification

	Two classes classification		Three classes classification	
	Training phase	Testing phase	Training phase	Testing phase
Accuracy	99.57%	100%	97.50%	94.62%
Specificity	98.50%	100%	98.94%	99.57%
Sensitivity	100%	100%	94.38%	88.79%

Table 10 shows the HMLP network produces high percentages of accuracy, specificity and sensitivity for both training and testing phases for two classes classification. In the training phase, the HMLP network produces 99.57% of accuracy, 98.50% of specificity and 100% of sensitivity. The HMLP network has successfully produces high percentages of accuracy, specificity and sensitivity with 100% in the testing phase. Overall, HMLP network produces high performance for normal and abnormal sperms classification.

Furthermore, based on Table 10, the HMLP network has also been proven to produce good performance for three classes classification of rat sperm namely normal, hookless and banana shape abnormals. The HMLP network produces 97.50% of accuracy, 98.94% of specificity and 94.38% of sensitivity in the training phase. For the testing phase, the HMLP network produces 94.62% of accuracy, 99.57% of specificity and 88.79% of sensitivity. Based on these percentages, it can be concluded that the HMLP network is still capables to produce good performance for the three classes classification.

6 Conclusion

In this paper, three features from Sprague Dawley rat sperm have been proposed as inputs to the neural network application namely opened degree, width of the curve and template matching percentage. Using the HMLP network as a classifier, the proposed input features are successfully operate as expected. Using the proposed features as the neural network inputs, the HMLP network has been proven to produce high performance for two classes rat sperm classification namely normal and abnormal. Furthermore, the HMLP network has also produced high performance for three classes of rat sperm classification namely normal, hookless abnormal and banana shape abnormal.

As conclusion, the HMLP network trained with the MRPE algorithm has been proven to successfully produce high accuracy for Sprague Dawley rat sperm classification using the proposed features namely opened degree, width of the curve and template matching percentage as its inputs.

References:

[1] Raab, S. S., & Grzybicki, D. M. Anatomic Pathology Workload and Error. *American Journal Clinical Pathology 125*. 2006. pp. 809-812.

[2] Roberts, M. S. The Use of Decision Analysis for Understanding the Impact of Diagnostic Testing Errors in Pathology. *American Journal Clinical Pathology 126* (Suppl 1). 2006. pp. S36-S43.

[3] Chorneyko, K., & Butany, J. Canada's pathology. *CMAJ 178*(12). 2008. p. 1523-1524.

[4] Fletcher, C. D. M. Symposium on Errors, Error Reduction, and Critical Values in Anatomic Pathology. *Archives of Pathology and Laboratory Medicine*: Vol. 130, No. 5, 2006. p. 602–603.

[5] Duggan, B. Pathologist to Face New Inquiry Over Cancer Error [Online]. [Accessed 31 January 2009]. Available from World Wide Web: http://www.independent.ie/national-news/pathologist-to-face-new-inquiry-over-cancer-error-1248229.html. 2007.

[6] Joshi, S. A., Shaikh, S., Ranpura, S., & Khole, V. V. Postnatal Development and Testosterone Dependence of a Rat Epididymal Protein Identified by Neonatal Tolerization. *Reproduction 125*. 2003. p.495-507.

[7] Kawai, Y., Hata, T., Suzuki, O., & Matsuda, J. The relationship between Sperm Morphology and In Vitro Fertilization Ability in Mice. *Journal of Reproduction and Development 52*(4). 2006. p. 561-568.

[8] Hu, J. H., & Yan, Y. C. Identification of γ1 Subunit of GABAA Receptor in Rat Testis. *Cell Research 12*(1). 2002. p. 33-37.

[9] Computer Assisted Sperm Analysis (CASA). Rat Sperm Morphological Assessment. Available from World Wide Web: http://www.irdg.co.uk/Sperm_morphology.pdf. 2000.

[10] Mashor, M. Y. Nonlinear System Identification Using HMLP Network. *J. Institution of Engineers, Malaysia*. 1999.

[11] Mashor, M. Y. Performance Comparison between HMLP, MLP, and RBF Networks with Application to On-line System Identification. *Proceedings of the 2004 IEEE Conference on Cybernetics and Intelligent Systems*, 2004, Singapore. 2004. p. 643-648

[12] Mashor, M. Y. Performance Comparison Between Back Propagation, RPE and MRPE Algorithms for training MLP Networks. 2000.

[13] Mashor, M. Y., Esugasini, S., Mat-Isa, N. A., & Othman, N. H. Classification of Breast Lesions using Artificial Neural Network. Biomed 06, *IFMBE Proceedings 15. Springer-Verlag Berlin Heidelberg*. 2007. p. 45-49.

[14] Mat-Isa, N. A., Subramaniam, E., Mashor, M. Y., & Othman, N. H. Fine Needle Aspiration Cytology Evaluation for Classifying Breast Cancer using Artificial Neural Network. *American Journal of Applied Sciences 4*(12). 2007. p. 999-1008.

[15] Mat-Isa, N. A., Mashor, M. Y., & Othman, N. H. Classification of Cervical Cancer Cells using HMLP Network with Confidence Percentage and Confidence Level Analysis. *International Journal of the Computer, The Internet and Management 11*(1). 2003. p. 17-29.

[16] Ramli, D. A., Kadmin, A. F., Mashor, M. Y. & Mat-Isa, N. A. Diagnosis of Cervical Cancer using Hybrid Multilayered Perceptron (HMLP) Network. *M.G h. Negoita et al. (Eds.): KES 2004, LNAI 3213*, 2002. p. 591–598.

[17] Salleh, N. M., Mat-Sakim, H. A., & Othman, N. H. Neural Networks to Evaluate Morphological Features for Breast Cells Classification. *IJCSNS International Journal of Computer Science and Network Security, 8*(9). 2008. p. 51-58.

[18] Mashor, M. Y., Osman, M. K., & Arshad, M. R. 3D Object Recognition using 2D Moments and HMLP Network. *Proceedings of the International Conference on Computer Graphics, Imaging and Visualization (CGIV'04), IEEE Computer Science.* 26-29 July 2004. p. 126- 130.

[19] Mashor, M. Y. Performance Comparison Between Back Propagation, RPE and MRPE Algorithms for training MLP Networks. 2000.

[20] Chen, S., Cowan, C. F. N., Billings, S. A., & Grant, P. M. A Parallel Recursive Prediction Error Algorithm for Training Layered Neural Networks. *Int. J. of Control*. *51*(6). 1990. p. 1215-1228.

[21] Hamidreza, Z. Shape Recognition by Clustering and Matching of Skeletons. *Journal of Computers*. Vol 3. 2008. p. 24-33.

[22] Ulaby, F. T. *Electromagnetics for Engineers*. Pearson Education Inc. Upper Saddle River, NJ 07458. 2005(a). p. 30-31.

[23] Ulaby, F. T. *Electromagnetics for Engineers*. Pearson Education Inc. Upper Saddle River, NJ 07458. New Jersey. 2005(b). p. 28-29.

[24] Myler, H. R., & Weeks A. R. *Computer Imaging Recipes in C*. Prentice Hall. A Simon & Schuster Company, Englewood Cliffs, New Jersey. 1993

[25] Kagawa, K., Ogura, Y., Tanida, J., & Ichioka, Y. Prototype Demonstration of Discrete Correlation Processor-2 Based on High-Speed Optical Image Steering for Large-Fan-Out Reconfigurable Optical Interconnections. *Optical Review 8*(1). 2001.np. 18-25.

[26] Hoang, A. *Supervised Classifier Performance on The UCI Database*. M. S. C. thesis. University of Adelaide. 1997.

[27] Mat-Isa, N. A. *Early Diagnosis System for Cervical Cancer Based on neural Network*. Ph.D thesis. Universiti Sains Malaysia. 2003.

[28] Mat-Sakim, H. A. *Application of Neural Networks on Fine Needle Aspirates of Breast Lesions*. Ph.D thesis. Universiti Sains Malaysia. 2003.