New Features of Cervical Cells for Cervical Cancer Diagnostic System Using Neural Network

N. Mustafa[1] **N. A. Mat Isa[2]** **M. Y. Mashor[3]** **N. H. Othman[4]**

[1,3]Electronic & Biomedical Intelligent Systems (EBItS) Research Group,
School of Mechatronics Engineering, Universiti Malaysia Perlis,
02600 Jejawi, Arau, Perlis, Malaysia
E-mail: nazahah@unimap.edu.my[1], yusoff@unimap.edu.my[2]

[2]School of Electrical & Electronic Engineering,
Universiti Sains Malaysia, Engineering Campus,
14300 Nibong Tebal, Penang, Malaysia.
E-mail: ashidi@eng.usm.my

[4]Pathology Department, School of Medical Science,
Universiti Sains Malaysia, Medical Campus,
16150 Kubang Kerian, Kelantan, Malaysia
E-mail: hayati@kb.usm.my

Abstract

Currently, Pap test is the most popular and effective test for cervical cancer. However, Pap test does not always produce good diagnostic performance. This problem has encouraged several studies to develop diagnosis system based on neural networks to increase the diagnostic performance. In order for neural networks to be used as cervical cancer diagnostic system, the features of cervical cell are used as inputs for neural networks and the classification of cervical cell type are used as output target. This study proposes new features of cervical cell that are suitable and can be used as inputs for neural networks for cervical cell classification system. The new cervical cell features are extracted from ThinPrep® images and their suitability are tested by using neural network. The results shows that all the nine extracted features which are size, grey level, perimeter, red, green, blue, intensity1, intensity2 and saturation are suitable for cervical cell classification in cervical cancer diagnostic system.

Keywords

Cervical cancer, ThinPrep®, feature extraction, neural network

1. Introduction

Cervical cancer is the second most common type of cancer that affects women, ranked after breast cancer. Cervical cancer is largely preventable if precancerous lesions are detected by effective screening and then adequately treated [1]. The most popular and preferred screening test for cervical cancer is Pap test [2]. Pap test begins with a wooden scraper (spatula) and/or a small brush that is used to collect a sample of cells from cervix and upper vagina. The cells are placed on a glass slide and any abnormal changes of cervical cells are determined. Based on Bethesda System, cells are classified into normal, low grade squamous intraepithelial lesion (LSIL) and high grade squamous intraepithelial lesion (HSIL. However, Pap test does not always produce good diagnostic performance due to bad samples, technical and human errors, and small size of cancer development area [3][4][5].

Due to limitations of diagnosis performance by Pap test, automated or semi-automated screening systems based on computer aided visualization and intelligent diagnosis have been developed to increase the diagnostic performance of the Pap test. Example of systems based on computer aided visualization and intelligent diagnosis are PAPNET®, AutoPap® Primary Screening, AutoCyte SCREEN and NeuralPAP [6]. Besides, there is also a few system designed to improve the cervical cell sample obtained from patients. The prominent systems in this category are ThinPrep®, PrepStain™, CytoScreen® and Labonord Easy Prep® [7]. These systems use liquid based cytology (LBC) technique to improve the quality of the conventional smear through an improved slide preparation technique following the collection of the sample in the standard way.

Normally, cervical cancer diagnosis systems are developed based on algorithmic image analysis [8]. Most of these systems help the expert to perform better diagnosis by improving cell images quality so that the morphological features can be seen easily. Recently, a few study has work on cervical cancer diagnostic system

based on neural network in order to increase the diagnosis performance [9][10][11][12][13]. In order for neural networks to be used as cervical cancer diagnostic system, features of cervical cell which are extracted by human expert could be used as neural networks inputs and the classification of cervical cells type could be the output [14].

Generally, the nucleus size increases from normal cell to HSIL cell while cytoplasm size decreases from normal cell to HSIL cell. As a result, the nucleus to cytoplasm ratio (NC ratio) increases from normal cell to HSIL cell [15]. At the same time, nucleus and cytoplasm regions become darker from normal cell to HSIL cell [16]. Therefore, features of cervical cell such nucleus size, cytoplasm size, nucleus grey level and cytoplasm grey level are commonly used to classify cervical cell into normal and abnormal (LSIL and HSIL) cells. In NeuralPAP system, these four features have been selected to be used as input to the hierarchical hybrid multilayered perceptron (H^2MLP) neural network for classification of cervical cell into normal and abnormal [17]. Besides, classification of cervical cell can also be done by using the grey level co-occurrence matrix textural features [18].

This study proposes a few new features to be extracted from the cervical cells. The features are perimeter, red, green, blue, intensity1, intensity2 and saturation. The suitability of these features in classifying the cervical cell into normal, LSIL and HSIL cells will be tested by using neural network. For this purpose, the H^2MLP network [17] has been chosen.

2. Approach/Methodology

The image of cervical cells has been captured from ThinPrep® slides using a computerised microscope. A total of 508 data (each data of ThinPrep® image consists of a cervical cell) were collected from Hospital University Sains Malaysia. First, the captured images were revised by the pathologist to determine appropriateness and type of the cervical cell. The cervical cells is classified into three categories; normal, LSIL and HSIL. Then, features of cervical cell of each ThinPrep® images will be extracted manually using image analyzer. This process will be done by experienced cytotechnologists with assistance and supervision by experienced pathologists.

The proposed features are perimeter, red, green, blue, intensity1, intensity2 and saturation. Intensity1, intensity2 and saturation were computed using Equation (1), (2) and (3) respectively [19][20]. Besides, this study also has extracted conventional features such as size and grey level. Therefore, the total numbers of features that have been extracted from nucleus and cytoplasm components of cervical cell in ThinPrep® images were nine.

$$\text{Intensity1} = \frac{1}{3}(\text{Red} + \text{Green} + \text{Blue}) \quad (1)$$

$$\text{Intensity2} = (0.299*\text{Red}) + (0.587*\text{Green}) + (0.114*\text{Blue}) \quad (2)$$

$$\text{Saturation} = \sqrt{c_1^2 + c_2^2} \quad (3)$$

where $c_1 = \text{Red} - 0.5*\text{Green} - 0.5*\text{Blue}$ (4)

$$c_2 = -\frac{\sqrt{3}}{2}*\text{Green} + \frac{\sqrt{3}}{2}*\text{Blue} \quad (5)$$

After the feature extraction process was completed, the suitability of the nine extracted features in classifying the cervical cell into three different categories was tested by using neural network individually. For this purpose, (H^2MLP) network [17] was used to test the suitability of each extracted features. The 508 data were divided into 255 training data and 253 testing data sets, where the distributions of the data are as shown in Table 1.

Table 1 : Distribution of training and testing data sets

Category of cervical cell	Training data	Testing data
Normal	192	192
LSIL	40	39
HSIL	23	22
Total	255	253

3. Results and Discussion

Using the 255 training data and 253 testing data sets, the diagnostic performance of the proposed features are summarised in Table 2.

The suitability of the proposed extracted feature in detecting of cervical cancer is determined based on their accuracy percentage. All nine features including the proposed features have provided high accuracy with percentage over than eighty percent in classifying the cervical cell into normal, LSIL and HSIL. These results show that all the proposed features are suitable to be used as input to neural network system for cervical cancer diagnostic system.

Besides, this study also provides result for all nine features as inputs to the H^2MLP network. Table 3 shows the diagnosis performance for all features. It is discovered that, the system achieved high accuracy with 94.29%. The obtain results shows that the combination of all nine features which are size, grey level, perimeter, red, green, blue, intensity1, intensity2 and saturation are suitable for cervical cell classification in cervical cancer diagnostic system.

Table 2 . Diagnosis performance of individual feature

Feature	Phase	Acc (%)	Sens (%)	Spec (%)	FN (%)	FP (%)
Size	Train	91.76	76.19	96.88	4.76	3.13
	Test	86.17	65.57	92.71	1.63	7.29
	Overall	88.98	70.97	94.80	3.22	5.21
Grey level	Train	85.88	42.86	100.00	28.57	0.00
	Test	77.87	31.15	92.71	52.46	7.29
	Overall	81.89	37.10	96.36	40.32	3.64
Perimeter	Train	92.55	80.95	96.35	3.17	3.65
	Test	85.38	62.30	92.71	3.28	7.29
	Overall	88.98	71.78	94.53	3.22	5.47
Red	Train	85.88	42.86	100.00	28.57	0.00
	Test	77.87	31.15	92.71	52.46	7.29
	Overall	81.89	37.10	96.36	40.32	3.64
Green	Train	87.45	57.14	97.40	6.35	2.60
	Test	75.49	37.70	87.50	39.34	12.50
	Overall	81.49	47.58	92.45	22.58	7.54
Blue	Train	83.14	33.33	99.50	34.92	0.52
	Test	79.84	24.59	97.40	67.20	2.60
	Overall	81.50	29.03	98.45	50.80	1.56
Intensity1	Train	88.63	58.73	98.44	9.52	1.56
	Test	74.70	44.26	84.38	37.70	15.63
	Overall	81.69	51.61	91.41	23.38	8.58
Intensity2	Train	88.63	53.97	100.00	20.63	0.00
	Test	79.84	34.43	94.27	49.18	5.73
	Overall	84.25	44.36	97.14	34.67	2.86
Saturation	Train	87.45	58.73	96.88	14.29	3.13
	Test	73.12	47.54	81.25	32.79	18.75
	Overall	80.31	53.23	89.07	23.39	10.92

Note : Acc=accuracy, Sens=sensitivity, Spec=specificity, FN=false negative, FP=false positive

Table 3 . Diagnosis performance for all features

Analysis	Train (%)	Test (%)	Overall (%)
Accuracy	97.65	90.91	94.29
Sensitivity	90.48	65.57	78.23
Specificity	100.00	98.96	99.48
False Negative	0.00	4.92	2.42
False Positive	0.00	1.04	0.52

4. Conclusion

This study has proposed new features of cervical cell to be used for classification of cervical cell in cervical cancer diagnostic system. The proposed features perhaps can be used as inputs to neural network in a diagnostic system in order to increase its diagnostic performance.

5. References

[1] Noorani, H. Z., Brown, A., Skidmore, B. & Stuart, G. C. E. (2003). Liquid-Based Cytology and Human Papillomavirus Testing In Cervical Cancer Screening**.** *Canadian Coordinating Office for Health Technology Assessment*: Technology Report. www.ccohta.ca/publications/pdf/197_cervical_cancer_tr_e.pdf. 2006.

[2] Cronjé, H. S. (2004). Review Article: Screening for Cervical Cancer In Developing Countries. *Int. J of Gynecol & Obstet*. 84(2). 101-108.

[3] Othman, N. H., Ayub, M. C., Aziz, W. A. A., Muda, M., Wahid, R. & Selvarajan, S. (1997). Pap Smears – Is It An Effective Screening Methods for Cervical Cancer Neoplasia? – An Experience With 2289 Cases. *The Malaysian Journal of Medical Sciences*. 4(1). 45-50.

[4] Tay, S. K. (1996). *Cervical Cancer: Its Causes and Prevention*. Singapore: Times Book Int..

[5] Hislop, T. G., Band, P. R., Deschamps, M., Clarke, H. F., Smith, J. M. & Ng, V. T. Y. (1994). Cervical Cancer Screening In Canadian Native Women: Adequacy of The Papanicolaou. *The Journal of Clinical Cytology and Cytopathology*. 38(1). 29-32.

[6] Medical Services Advisory Committee. (2003). *Computer-Assisted Image Analysis for Cervical Screening: Assessment Report*. Canberra: Commonwealth Minister for Health and Ageing. www.msac.gov.au/pdfs/reports/msacref12c.pdf. 2006.

[7] Karnon, J., Peters, J., Platt, J., Chilcott, J., McGoogan, E. & Brewer, N. (2004). Liquid-Based Cytology In Cervical Screening: An Updated Rapid and Systematic Review and Economic Analysis. *Health Technology Assessment.* 8(20). http://www.ncchta.org/fullmono/mon820.pdf. 2006.

[8] Mashor, M. Y., Mat-Isa, N. A. & Othman, N. H. (2002). Automatic Pap Smear Screening Using HMLP Network. *Proc. Of the Int. Conf. on Artificial Intelligence in Engineering & Technology*. 453-457.

[9] Tan, K. L., Mashor, M. Y., Mat-Isa, N. A., Ali, A. N. & Othman, N. H. (2003). Design of A Neural Network Based Cervical Cancer Diagnosis System : A Microcontroller Approach. *Proc. Of the 3rd Int. Conference On Advance In Strategic Technologies*. 2. 725-729.

[10] Mat-Isa, N. A., Mashor, M. Y. & Othman, N. H. (2002). Diagnosis of Cervical Cancer Using Hierarchical Radial Basis Function (HiRBF) Network. *Proc. Of the Int. Conf. on Artificial Intelligence In Engineering & Technology*. 458-463.

[11] Li, Z. & Najarian, K. (2001). Automated Classification of Pap Smear Tests Using Neural Networks. *Proc. of Int. Joint Conf. on Neural Networks*. 4. 2899-2901.

[12] Mitra, P., Mitra, S. & Pal, S. K. (2000). Staging of cervical Cancer with Soft Computing. *IEEE Trans. on Biomedical Engineering*. 47(7). 934-940.

[13] Balasubramaniam, R., Rajan, S., Doraiswani, R. & Stevenson, M. (1998). A Reliable Composite Classification Strategy. *Proc. Of IEEE Canadian Conf. on Electrical and Computer Engineering*. 2. 914-917.

[14] Mat-Isa, N. A., Mat-Sakim, H. A., Mashor, M. Y. & Othman, N. H. (2003). A Review on Application of Neural Networks In Cancer Disease : Case of Cervical and Breast Cancer. *Proc. of The Int. Conf. on Robotics, Vision, Information and Signal Processing*. 538-543.

[15] Thiran, J. P., Macq, B. & Mairesse, J. (1994). Morphological Classification of Cancerous Cells. *IEEE Proc. of Int. Conf. on Image Processing*. 3. 706-710.

[16] WebMD. (2002). How Can Cervical Cancer Be Prevented?. http://www.webmd.com/content/dmk/dmk_article_3961643. 2004.

[17] Mat-Isa, N. A. (2003). *Early Diagnosis System for Cervical Cancer Based On Neural Networks*. PhD thesis, University Science of Malaysia.

[18] Walker, R. F., Jackway, P., Lovell, B. & Longstaff, I. D. (1994). Classification of Cervical Cell Nuclei Using Morphological Segmentation and Textural Feature Extraction. *Proc. of 2nd Australian and New Zealand Conf. on Intelligent Information Systems*. 297-301.

[19] Zhang, C. & Wang, P. (2000). A New Method of Color Image Segmentation Based on Intensity and Hue Clustering. *Proc. of 15th Int. Conf. on Pattern Recognition*. 3. 613-616.

[20] Weeks, A. R. (1996). *Fundamentals of Electronics Image Processing*. Washington: Spie Optical Engineering Press.