Classification of Optical Images of Cervical Lymph Node Cells

Salim J. Attia*
* College of Dentistry, University of Baghdad
* Email: salimj11@avic.uobaghdad.edu.iq; j.salim08@yahoo.com

Abstract: the study considers the optical classification of cervical nodal lymph cells and is based on research into the development of a Computer Aid Diagnosis (CAD) to detect the malignancy cases of diseases. We consider 2 sets of features one of them is the statistical features; included Mode, Median, Mean, Standard Deviation and Maximum Probability Density and the second set are the features that consist of Euclidian geometrical features like the Object Perimeter, Area and Infill Coefficient. The segmentation method is based on following up the cell and its background regions as ranges in the minimum-maximum of pixel values. The decision making approach is based on applying of Minimum Distance which give accuracy of 97%.

Keywords: lymph node cells, segmentation, minimum distance, optical imaging.

Introduction:
Cervical lymph nodes are laid on the neck. In the neck of the human body there are 300 of the 800 lymph nodes [1]. Several cases affect the lymph nodes like infection, tumors, inflammation. The lymph nodes in adults can be palpable even in a healthy state. In children at 12 years down, the cervical nodes take size of 1 cm and may be palpable, but this may not mean any disease. In persons above 50 years, cancers lead to metastatic enlargement [2].

The involvement of lymph nodal reflects the clinical, pathological mark in primary lymphomas, because it indicates the transformation or disease development from localized to systemic state; so the following up of nodal lymph is important for staging them. Fine Needle Aspiration (FNA) is broadly applied in the detection of lymphadenopathy, staging and follow-up them [3]. Lymphomas may take a long-term clinical course [4] in which nodal lymph may develop and enlarge for different reasons. A surgical biopsy is difficulty accepted by patients and rarely performed, so the surgical biopsy for only detection goals might be looked at as an overuse choice and excessive intervention in sometimes.

Fine needle aspiration (FNA) has a crucial role in diagnosing lymphadenopathy [5,8]. It is noticed in clinical settings, the lymph nodes FNA of lymph nodes present problems related to sampling, cells amount, different cellular properties. In general, some institutes are requested FNA for all patients suffering from enlargement of nodal lymph and suspected for malignancy by any diagnostic method such as Ultrasound or clinical. Other researchers had used image processing and analysis for recognition and classification of lymph node images from different imaging techniques [9,12].

Hardware and Methodology
1. Samples and imaging
Samples of different types of benign and malignant of cervical nodal lymph tissue were prepared as slides using the Fine Needle Aspiration (FNA) and Pap stain. These slices were imaged by the standard optical microscope connected to a CCD camera.
2. Image segmentation
In segmentation process, the image is divided into separated regions. The best case is to get two regions: object of interest and its background. The algorithms of image segmentation are a wide and demanding area like medical imaging, satellite imagery and computer vision systems. Its aim is to divide an image into meaningful area which are in to generate a feature vector for a certain task. Segmentation is the first and necessary step prior to the detection and classification of images and extracting their features. The segmentation depends on similarities are revealed or boundary based approaches, which depend on determining the edges.

In this study the color of the images is exploited by calculating the average value of Red, Green and Blue colors for the inter-cellular points and getting minimum-maximum of the values of these colors [13].

The result provides the segmented images and the examples being given in Figure 1. Two hundred images were collected; 100 benign and 100 malignant. The images were divided into two groups (50 benign + 50 malignant) for each; the first one represents the control basis and the second as the test group.

| No. | 1 | 2 | 3 | 4 | 5 | …… | 100 |
|-----|---|---|---|---|---|-----|-----|
| Benign cells | ![Image](benign1.png) | ![Image](benign2.png) | ![Image](benign3.png) | ![Image](benign4.png) | ![Image](benign5.png) | …… | ![Image](benign100.png) |
| Malignant Cells | ![Image](malignant1.png) | ![Image](malignant2.png) | ![Image](malignant3.png) | ![Image](malignant4.png) | ![Image](malignant5.png) | …… | ![Image](malignant100.png) |

Figure 1. Benign and Malignant of cervical node cells as segmented images

**Features of images**

The features are grouped in a set of numeric values as a feature vector or an input to a decision making (Minimum Distance) method. The feature vector included the statistical parameter ‘mode, median, mean, standard deviation, Maximum Probability Density’ and Euclidean geometrical features ‘area, perimeter, infill coefficient’ as follows:

1. **Mode.** The pixel value which repeatedly occurs more than any other pixel [14].

2. **Median.** That pixel value in the middle of other pixels if they are arranged in descending or ascending way. If there is more than one mode, the average value of them represents the mode [15].

3. **The Mean.** The average value of all image pixels and reflects the brightness of the image and got from equation [16]:

\[
\mu = \sum_{k=0}^{K-1} k P(k) \tag{1}
\]

k is the intensity of pixel value from 0 to 255
P(k)is the pixel probability of k value

4. **Standard Deviation:** it is used as a measure for how far the average of all values from their mean. It is considered a sign of contrast in an image. It is given by [15]:

\[
\sigma = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} \left(I_{ij}^{2} - \mu^{2}\right) \tag{2}
\]
I: Intensity; µ: mean

5. **Maximum Probability Density**: The histogram represents the total number of each gray level [7]. Maximum Probability Density $P_{\text{max}}(k)$ is resulted by dividing the sum of the most frequent gray level by total number of image pixels [9]. We consider $P_{\text{max}}(k)$ [17]:

$$P_{\text{max}} = \frac{h(k)}{K} \quad (3)$$

Where $h(k)$ is the sum of most frequent pixel and $K$ is the number of all pixels.

6. **Object Area**: if the function $I_q(i, j)$ refers to a segmented region or cell $q$, the following function is considered [17]:

$$I_q(i, j) = \begin{cases} 1 & \text{if} \quad I(i, j) = n^{th} \ \text{object number} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

The area of the object is obtained from [8]:

$$A_q = \sum_{x \in \text{cell}} 1(x) \quad (5)$$

7. **Object Perimeter**: it is the total number of boundary pixels of the object or cell. The boundary pixel is determined by taking 1 value and at least 0 neighbor value. So it is given by [17]:

$$P = \sum_{x \in \text{edge}} 1(x) \quad (6)$$

8. **Infill Coefficient**: it is defined by the area of cell area divided by polygon area around it.

**Minimum Distance**

The minimum distance approach is applied as decision making to reflect the degree of similarity of features values and knows if any set of these values belonged to a specific class or not. If a certain class is given in space $S$. Minimum Distance is obtained by [19]:

$$d(x_g, m_{ig}) = \sqrt{(x_i - m_{i})^2 + \ldots + (x_K - m_{iK})^2} \quad (7)$$

Or in another form:

$$d(x_g, m_{ig}) = |x_i - m_{i}| + \ldots + |x_K - m_{iK}| \quad (8)$$

$x_g$ is any point in $K$ space, $m_{ig}$ are two vectors of class $i$ and $j$ respectfully in dimension $K$.

For condition ($i \neq j$) the Minimum Distance is:

$$d(x_g, m_{ig}) < d(x_g, m_{jg}) \quad (9)$$

Based on above specific features, we apply Minimum Distance as decision making to differentiate between benign and malignant classes. The class belonging depend on if the
most object features are nearer to specific class than the other. This provides a result of classification that the research aims. The images are divided into two groups; 100 training images and 100 testing images (50 Benign and 50 Malignant for each one).

### Results and discussion

The application was depended on setting 2 groups of features for any cell; its statistical and geometrical ones. It is found that the statistical features of malignant cases are larger than those of benign cases except the Maximum Probability Density values and this refers to the usual observation which are compatible with previous information in this field, as seen in Table 1.

**Table 1: Average values of the features**

| Features                  | Benign | Malignant | Overlapped Ratio |
|---------------------------|--------|-----------|------------------|
| Mean                      | 86     | 114       | 12%              |
| Mode                      | 85     | 156       | 18%              |
| Median                    | 86     | 151       | 16%              |
| Standard Deviation        | 7      | 13        | 19%              |
| Max. Prob. Den.           | 0.12   | 0.06      | 17%              |
| Perimeter                 | 81     | 112       | 5%               |
| Area                      | 502    | 1811      | 13%              |
| Infill Coef.              | 0.951  | 0.892     | 35%              |

Max. Prob. Den. = Maximum Probability Density

The malignant appearance of cell is light and pale, whereas, the benign is dim and dark, so the malignant take higher values and benign lower values. The Maximum Probability Density values are less in malignant because the variety of image pixels and there is not dominant pixel as in benign images. The perimeter values and area are larger in malignant than benign as predicted, whereas the Infill Coefficient is less according to the irregularity of malignant cells taken into consideration the definition of this feature. There is generally an overlapping between feature values of benign and malignant images cells and their ratios are explained in Table 1. The classification is not precise with one or few features, thus, it is needed to take suffice number of variety features as done in this study. The Minimum Distance that has been associated with these parameters has given a good solution to the overlap problem and reaches successfully for classification aim.

### Conclusions

The methodology in this study represents a multi-factors estimation of cervical lymph nodes cells images coupled with a Minimum Distance approach to classify benign and malignant of these cells. The method obviously depends on the reliability of biopsy process, staining, imaging setting and parameters computation as feature vectors. Taking all these issues into account the study gives good result and 97% as classification accuracy according to the samples used to date.

### Acknowledgements

The author would like to thank the Department of Basic Sciences at College of Dentistry/ University of Baghdad for their support.

### References

[1] Park, Chang Ho; Song, Chang Myeon; Ji, Yong Bae; Pyo, Ju Yeon; Yi, Ki Jong; Song, Young Soo; Park, Yong Wook; Tae, Kyung. "Significance of the Extracapsular Spread of Metastatic Lymph Nodes in Papillary Thyroid Carcinoma". Clinical and Experimental Otorhinolaryngology, 2015, 8 (3): 289–294. doi:10.3342/ceo.2015.8.3.289
[2] Kerawala C, Newlands C, “Oral and maxillofacial surgery”, Oxford University Press, Oxford, 2010.
[3] Elena Vigliar1, Immacolata Cozzolino1, Marco Picardi2, Anna Lucia Peluso2, Laura Virginia Sosa Fernandez4, Antonio Vetrami1, Gerardo Botti3, Fabrizio Pane2, Carmine Selleri4 and Pio Zeppa4, “Lymph node fine needle Cytology in the staging and follow-up of Cutaneous Lymphomas”, Vigliar et al. BMC Cancer, 2014, 14(8), 1-10. http://www.biomedcentral.com/1471-2407/14/8
[4] Kim YH, Liu HL, Mraz-Gernhard S, “Long-term outcome of 525 patients with mycosis fungoides and Sezary syndrome: clinical prognostic factors and risk for disease progression”, Arch Dermatol 2003, 139:857–866.
[5] Zeppa P, Vigliar E, Cozzolino I, “Fine needle aspiration cytology and flow cytometry immunophenotyping of non-Hodgkin lymphoma: can we do better?”, Cytopathology 2010, 21:300–310.
[6] Zeppa P, Marino G, Troncone G, “Fine-needle cytology and flow cytometry immunophenotyping and subclassification of non-Hodgkin lymphoma: a critical review of 307 cases with technical suggestions”, Cancer 2004, 102:55–65.
[7] Galindo LM, Garcia FU, Hanau CA, “Fine-needle aspiration biopsy in the evaluation of lymphadenopathy associated with cutaneous T-cell lymphoma (mycosis fungoides/Sézary syndrome)”, Am J Clin Pathol 2000, 113:865–871.
[8] Pai RK, Mullins FM, Kim YH, Kong CS, “Cytologic evaluation of lymphadenopathy associated with mycosis fungoides and Sezary syndrome: role of immunophenotypic and molecular ancillary studies”, Cancer 2008, 114:323–332.
[9] Kempf W, Sander CA, “Classification of cutaneous lymphomas”, Histopathology 2010, 56:57–70.
[10] David V., Schacht Karen, Drukker Karen and Maryellen L. Giger, “Using Quantitative Image Analysis to Classify Axillary Lymph Nodes on Breast MRI”, Eur. J. Radiol. 2015, 84(3):392-397. doi: 10.1016/j.ejrad.2014.12.003.
[11] Laszlo Papp, Norbert Zsoter, Charlotte LohShow, “Automated lymph node detection and classification on breast and prostate cancer SPECT-CT images”, Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference:3431-4, 2011. doi: 10.1109/EMBS.2011.6090928
[12] Jamaluddin M. F., Fauzi M. F. A. and Abas F. S., ”Tumor detection and whole slide classification of H&E lymph node images using convolutional neural network”, IEEE International Conference on Signal and Image Processing Applications (ICSIPA), Kuching, 2017, pp. 90-95, doi: 10.1109/ICSIPA.2017.8120585.
[13] Blackledge J. M. and Dubovitskiy D. A., “Object Detection and Classification with Ap-plications to Skin Cancer Screening”, International Transactions on Intelligent Systems, 2008, 1(1), 34-45.
[14] Attia S., “Optical Diagnosis of Prostate Cancer”, IOP Conf. Series: Materials Science and Engineering 571 (2019) 012123 IOP Publishing. doi:10.1088/1757-899X/571/1/012123
[15] Smoller S. W., “Biostatistics and Epidemiology”, Third Edition, Springer, 2004.
[16] Gomes J. and Velho L., “Image Processing for Computer Graphics”, Silvio Levy, New York, 1997.
[17] Merchant Wu. F. A. and Castle-man K. R., “Microscope Image Processing”, Elsevier, 2008.
[18] Jelen L., Fevens T. and Krzyzak A., “Classification of Breast Cancer Malignancy using Cytological Images of Fine Needle Aspiration Biopsies”, Int. J. Appl. Math. and Comp. Sci., Vol. 18, No. 1, 75-83, 2008.
[19] Umbaugh S. E., “Computer Vision and Image Processing”, Prentice-Hall, 1998.