Middle-East Journal of Scientific Research 20 (8): 986-991, 2014
ISSN 1990-9233

DOI: 10.5829/idosi.mejsr.2014.20.08.114043

Skin Histopathological Image Used LDED Based Algarithm with Automated Detection of Melanocytes Skin

P. Ramya

Vins Christian College of Engineering, Bharat University, Chennai, India

XYZ

Abstract: Skin cancer is the most frequent and malignant type of cancer. Melanoma is the most aggressive type among skin cancers. In melanoma diagnosis, the detection of the melano cytes in the epidermis area is an important step. This paper presents an effective computer-aided technique for segmentation and detection of the melanocytes in the skin histopatho logical image. The nuclei regions in the epidermis area are segmented using the K-means clustering algorithm with k value as 3. K-means clustering algorithm clusters the nuclei region based on space and color information. Then a local region recursive segmentation (LRRS) algorithm is applied to detect the candidate nuclei regions from the initially segmented image. LRRS uses two parameters: intensity and size of nuclei to filter out the candidate nuclei regions. Finally, a novel descriptor, named local double ellipse descriptor (LDED) is applied to differentiate melanocytes from keratinocytes. LDED is based on the double ellipsoidal model and utilizes the local features to distinguish melanocytes from the candidate nuclei regions.

Key words: Histopathological image analysis · Image segmentation · Local descriptor · K-means clustering · Pattern recognition

INTRODUCTION

The most dangerous form of skin cancer, these cancerous growths develop when unrepaired DNA damage to skin cells (most often caused by ultraviolet radiation from sunshine or tanning beds) triggers mutations (genetic defects) that lead the skin cells to multiply rapidly and form malignant tumors. These tumors originate in the pigment-producing melanocytes in the basal layer of the epidermis. Melanomas often resemble moles; some develop from moles. The majority of melanomas are black or brown, but they can also be skin-colored, pink, red, purple, blue or white. Melanoma is caused mainly by intense, occasional UV exposure (frequently leading to sunburn), especially in those who are genetically predisposed to the disease. Melanoma kills an estimated 8,790 people in the US annually.

If melanoma is recognized and treated early, it is almost always curable, but if it is not, the cancer can advance and spread to other parts of the body, where it becomes hard to treat and can be fatal. While it is not the most common of the skin cancers, it causes the most deaths. The American Cancer Society estimates that at present, about 120,000 new cases of melanoma in the US are diagnosed in a year.

In melanoma diagnosis, the segmentation and detection of the melanocytes in the epidermis area is an important step before the diagnosis is made. If the melanocytes can be found correctly, architectural and cellular features (e.g., size, distribution, location) can be used to grade or determine the malignancy of the skin tissue [1-3].

The digitized histopathological images used in this study are stained with haematoxylin and eosin (H&E). The cell nuclei are observed as dark blue, whereas the intracellular material and cytoplasm are observed as bright pink. It is also noted that there exist colour variations in interimages and intraimages due to non-uniform absorption of the stain and different handling procedure or other factors, e.g., stains fading. In addition, the high similarity between the melanocytes and other cytological components make it difficult to perform consistent quantitative analysis.

Corresponding Author: P. Ramya, Vins Christian College of Engineering, Bharat University, Chennai, India

The paper further organized as follows. Various existing nuclei segmentation schemes and their pitfalls are described in section II. The proposed automated detection scheme is described in Section III. Simulation results and conclusions are respectively described in sections IV and V.

Existing Algorithms: Several works have been conducted on the segmentation or detection of various biological components in a histopathological image using image-processing techniques such as thresholding and watershed. The technique first employs morphological operations to reduce the background signal. These threshold-based techniques [8] typically fail when considerable intensity variations are present in the images. Nuclei detection technique using trained artificial neural networks [5] incorporates the color, texture and shape information present in an image. Although this technique had been reported to provide good performance, the performance is sensitive to the training samples. This method also needs large databases to store the training samples.

The Mean shift segmentation algorithm [1] for segmentation of nuclei region makes use of color and spatial information to segment the nuclei regions of the histopathological image. But this method suffers from interimage and intraimage variations present in the images.

The histopathological images may contain interimage and intraimage intensity variations [4-7] due to the staining imperfection. Therefore, many of the aforementioned techniques seem unsuitable for our scenario.

The Proposed Technique: In order to address the earlier mentioned problems, a novel technique to segment and detect the melanocytes in the skin epidermis area is proposed. Unlike the existing techniques which usually assume relatively uniform background, the proposed technique considers the interimage and intraimage variations due to the staining imperfection. Also, the proposed technique can provide good detection performance on histopathological images, where the background is complex and has similar appearance with the foreground (i.e., the melanocytes). Furthermore, the proposed technique models the natural biological features, i.e., the shape and the distribution of intensity, as the parameters which make the technique robust.

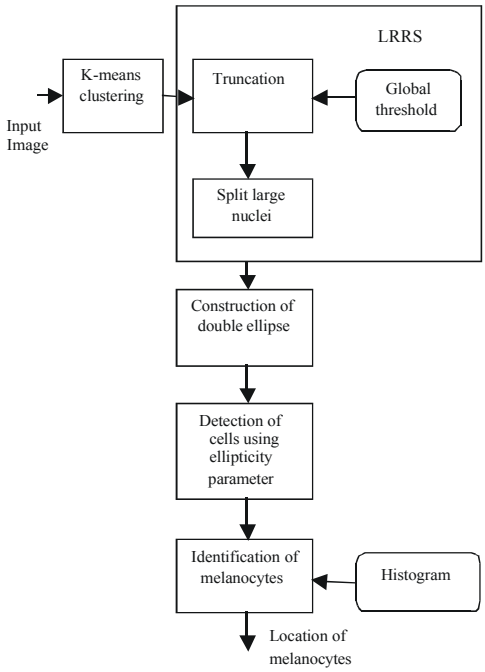

Fig. 1: Block diagram of the proposed technique

The block diagram of the LDED based automated detection of melanocytes is shown in Fig. (1). In the first step, the candidate nuclei regions are segmented in the epidermis area using the k-means clustering algorithm with k value as 3. The k-means clustering algorithm clusters the pixels into local regions. Next split-and-merge-based local region recursive segmentation (LRRS) algorithm is used to segment the candidate nuclei regions. LRRS is carried out on two steps; truncation and splitting large nuclei [8-10].

In the last step, a novel descriptor, named local double ellipse descriptor (LDED), is proposed to perform the quantitative analysis. This descriptor builds two ellipsoid models based on the segmented candidate regions. The LDED then measures the local feature information using two well-defined parameters which incorporate the biological pattern of the melanocytes. Based on the biological features present in the image the location of the melanocytes can then be identified by the LDED algorithm.

K-Means Clustering: Due to the staining imperfection and variations, the appearance of individual cytological components is not homogeneous and have complex texture surface. In order to reduce such variations, initial segmentation is required to decompose the original image

into homogeneous biological components. K-means clustering algorithm is used in this step to segment the candidate nuclei regions in the epidermis area.

The K-means clustering algorithm is a simple method for estimating the mean (vectors) of a set of K-groups. The algorithm iterates over two steps:

- Compute the mean of each cluster.
- Compute the distance of each point from each cluster by computing its distance from the corresponding cluster mean. Assign each point to the cluster it is nearest to.
- Iterate over the above two steps till the sum of squared within group errors cannot be lowered any more.

The initial assignment of points to clusters can be done randomly. In the course of the iterations, the algorithm tries to minimize the sum, over all groups, of the squared within group errors, which are the distances of the points to the respective group means. *k*-means clustering aims to partition *n* observations into *k* clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster. The k-means clustering algorithm is as follows:

- Initialize cluster centroids $\mu_1, \mu_2, \ldots, \mu_k$ randomly.
- Repeat until convergence:

  {

  For every i, set

  $c(i) := \arg\min \|x^{(i)} - \mu^j\|^2$

  For each j, set

$$\mu_j = \frac{\sum_{i=1}^{m} \{c^{(i)} = j\} x^i}{\sum_{i=1}^{m} c^{(j)} = j}$$

  }

In the algorithm above, k (a parameter of the algorithm) is the numbers of clusters want to find; and the cluster centroids μ^j represent our current guesses for the positions of the centers of the clusters.

Local Region Recursive Segmentation: After applying the k-means clustering algorithm, the pixels which have intensity similarity and geometric closeness are clustered together. The initially segmented image consists of many segmented regions which are denoted as $\{R_p\}p = 1...Z$, where *Z* is the total number of regions. In this paper, the object of interest is the nuclei regions in the epidermis area. The next step is to segment the candidate nuclei regions based on the clustered regions $\{R_p\}p = 1...Z$. In this section, a split-and-merge-based algorithm, named local region recursive segmentation (LRRS) to segment the nuclei regions is presented. In the LRRS algorithm, two domain specific conditions are incorporated:

- The intensity of the nuclei is lower than that in the cytoplasm.
- The size of a candidate nuclei region is within a predefined range.

The LRRS algorithm has two steps that are detailed in the following.

Truncation: In this step the mean intensity $\{q_p\}$ p = 1...*Z* for each region in $\{R_p\}p = 1...Zc$ is calculated. Then a global threshold *Tg* is calculated using Otsu's method for the mean intensity set $\{q_p\}$ $p = 1...Z$, followed by truncation of the region whose mean intensity is greater than T_g . After the truncation, most of the regions representing the cytoplasms are removed. Figure 2 shows an image obtained by applying step 1.

Splitting Large Nuclei: The remaining adjacent regions are merged to form the new regions set $\{R_p\}p = 1...Z$. Note that in these merged regions, there are under segmented regions, due to the intensity variation in the epidermis. Based on the domain knowledge that the nuclei region should be within an area range, a size prior criterion T_{area} is defined. T_{area} is the upper bound of the candidate nuclei region. For each merged region *Rp,* the number of intensity values *v* and the area $A(R_p)$ are estimated. The local region *Rp* which satisfies the following conditions will be further split into sub regions using the mean value of current region *Rp*.

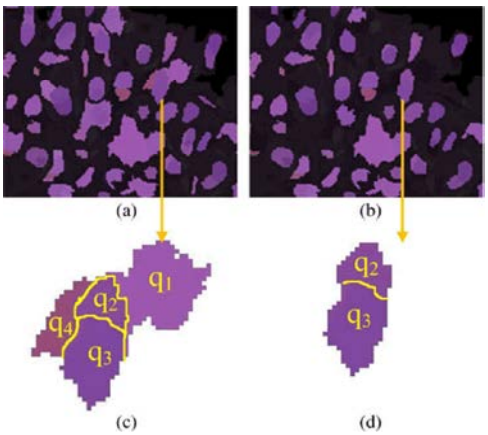

Fig. 2: Output of the local region recursive segmentation

$A(R_p) > T_{area}$ and $V > 2$ (1)

For the regions which do not satisfy the earlier mentioned conditions, assi`gn these regions to the candidate nuclei regions set $\{N_p\}p=1...K$. This split and-merge strategy is repeated until there is no region satisfying the condition.

Local Double Ellipse Descriptor Analysis: The LDED utilizes the candidate regions $\{N_p\}_{p=1...K}$ and its surrounding local features to discriminate the melanocytes and other cytological components. The details of the LDED analysis are presented in the following subsections.

Construction of the Elliptical Model: At first, an ellipse is fitted based on the boundary points of a candidate region N_p using the direct least-squares fitting algorithm [2]. It is observed that most of the nuclei are fitted well with ellipses. However, there are a few regions that have irregular shapes and cannot be fitted well by an ellipse. These regions need to be eliminated for efficient nuclei detection.

Construction of the Double Ellipse Descriptor: Using the aforementioned elliptical model the shape of the nuclei region can be measured. In order to capture the local information of the nuclei region, another elliptical model is build that has larger capturing range. Let E_{IN} denote the inner elliptical model. Now build another elliptical model E_{OT}, named outer/enlarged elliptical model, such that it has the same centroid position with that of E_{IN}, but has larger minor and major axes. The outer elliptical model is proposed to capture the surrounding local information of the current candidate nuclei region. Typically, the enlarged major and minor axes have a factor of 1.4.

Detection of the Nuclei Using Ellipticity Parameter: A few false positives, i.e., the regions which are not true nuclei, are expected to be present in the candidate regions $\{N_p\}_{p=1...K}$. Based on the assumption that a nuclei typically has an elliptical shape, the false positives can be filtered out by using the ellipticity of a region with the inner elliptical model. Denote S as the set of pixels in a candidate region N_p. A parameter which measures the ellipticity is defined as follows:

$$e_E \equiv \frac{|S \oplus Q_{IN}|}{|Q_{IN}|} \quad (2)$$

where Q_{IN} denotes the points of inner ellipse.

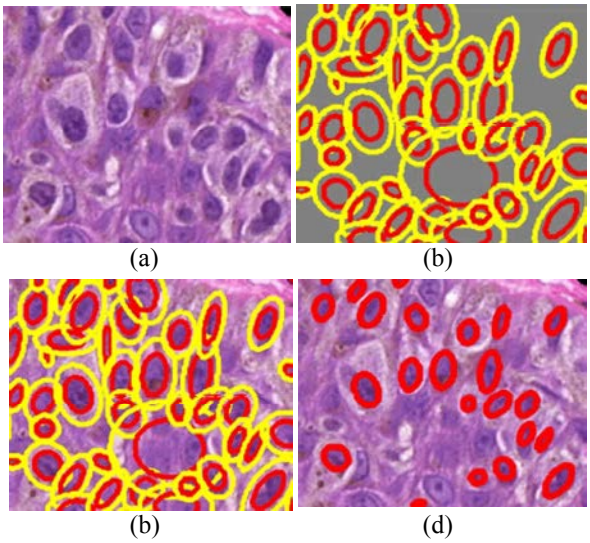

Fig. 3: Illustration of the LDED analysis. (a) Original image. (b) All the LDEDs. (c) LDEDs overlap onto the original image to perform the analysis. (d) Detection of the melanocytes

Detection of the Melanocytes: After the nuclei detection, the task is now to distinguish the melanocytes from other keratinocytes. In digitized image the nuclei become larger and has irregular contour, but the pattern is the same. This pattern can be easily observed. It is observed that the nuclei of the melanocytes prefer to have low intensity value, while its spatial surrounding space presents higher intensity value. It is shown in the histogram that there are two distinct modes present. The centers of these two modes lie around intensity value 110 and 175 [8-15]. On the other hand, in the case of other keratinocytes, it is difficult to find the two distinct modes and the histogram usually is unimodel. The distance of the two Gaussian modes is then used to distinguish the melanocytes and other keratinocytes.

Experimental Results: The simulation result of the LDED based automated detection of melanocytes is shown in Fig. 3(d).

The regions highlighted by yellow color show the outer ellipse built by LDED. The regions highlighted by the red color in Fig. 3(d) shows the melanocyte regions.

Performance Evaluation: The proposed technique is evaluated on different skin histopathology images of epidermis. These images are captured from different skin tissue samples corresponding to normal skin, nevus and melanoma.

Table 1: Erformance Evaluation of Proposed Method

Index	μ	σ	PPR	SEN
1	135.54	31.95	73.59	84.78
2	170.04	24.24	63.78	85.47
3	116.16	41.92	74.78	81.97
4	118.43	30.34	77.32	88.81
5	130.27	22.72	78.81	85.32
6	142.57	25.78	76.00	79.59
7	146.57	35.35	79.71	71.59
8	141.65	32.31	66.79	89.29
9	141.65	32.31	66.79	90.29
10	141.06	34.08	73.55	82.22

For the performance evaluation, the melanocytes manually identified are treated as the ground truths. We define N_{GT} as the total number of ground truths, N_{DO} as the total number of detected objects, N_{TP} as the number of true positives (i.e., correctly detected objects), N_{FP} as the number of false positives (i.e., falsely detected objects). The positive prediction rate (PPR) and the sensitivity (SEN) are defined as follows:

$$PPR = \frac{N_{TP}}{N_{DO}} \times 100 \quad (3)$$

$$SEN = \frac{N_{TP}}{N_{GT}} \times 100 \quad (4)$$

The evaluation result is shown in Table 1. In Table 1, the image index is shown in the first column. In the second and third columns, the image properties (i.e., the mean intensity μ and the standard variance σ of red channel image are presented. The remaining columns show the performance of the proposed technique, in terms of PPR and SEN.

It is observed that the proposed technique provides a robust performance on test images with different intensity levels.

CONCLUSION

In summary, it has been concluded that the proposed technique is giving much better results than the existing ones. This project presents a simple but effective computer-aided technique for segmentation of the melanocytes in the skin histopathological image. The candidate nuclei regions are first extracted through the k-means clustering method and the proposed local region recursive segmentation algorithm. The local double ellipse descriptor then incorporates the biological feature of melanocytes and provides robust parameters to identify the melanocytes. Also, the proposed technique can provide good detection performance on histopathological images, where the background is complex and has similar appearance with the foreground.

REFERENCES

1. Cheng Lu, Muhammad Mahmood, Naresh Jha and Mrinal Mandal, 2013. Senior Member, Ieee, Automated Segmentation Of The Melanocytes In Skin Histopathological Images, Ieee Journal Of Biomedical And Health Informatics, 17(2).
2. Andrew W. Fitzgibbon, Maurizio Pilu Robert B. Fisher, 1999. Direct Least Squares Fitting of Ellipses, Pattern Analysis And Machine Intelligence, 21(5).
3. Ilias Maglogiannis and Charalampos N. Douka, 2009. Overview Of Advanced Computer Vision Systems For Skin Lesions Characterization, Ieee Transactions On Information Technology In Biomedicine, 13(5).
4. Xiaowei Chen, Xiaobo Zhou and Stephen T.C. Wong, 2006. Automated Segmentation, Classification and Tracking Of Cancer Cell Nuclei In Time-Lapse Microscopy", Ieee Transactions On Biomedical Engineering, 53(4).
5. Tim, W. Nattkemper, Helge J. Ritter and Walter Schubert, 2001. A Neural Classifier Enabling High-Throughput Topological Analysis Of Lymphocytes In Tissue Sections, Ieee Transactions On Information Technology In Biomedicine, 5(2).
6. Ajay Nagesh Basavanhally, Shridar Ganesan and Shannon Agner, 2010. Computerized Image-Based Detection And Grading Of Lymphocytic Infiltration In Her2+ Breast Cancer Histopathology Ieee Transactions On Biomedical Engineering, 57(3).
7. Rigel, D., J. Russak and R. Friedman, 2010. The Evolution Of Melanoma Diagnosis: 25 Years Beyond The Abcds, Ca: Canc. J. Clinic., 60(5): 301-316.
8. Coelho, L., A. Shariff and R. Murphy, 2009. Nuclear Segmentation In Microscope Cell Images: A Hand-Segmented Dataset And Comparison Of Algorithms, In Proc. Ieee Int. Symp. Biomed. Imag, Nano Macro, pp: 518-521.
9. Sezgin M. Ismail, Angela B. Colclough, John S. Dinnen, Douglas Eakins, David M.D. Evans, Ernest Gradwell, Jerry P. O'sullivan, Joan M Summerell, Robert G. Newcombe, 1989. Observer Variation In Histopathological Diagnosis And Grading Of Cervical Intraepithelial Neoplasia", Brit. Med. J., 298: 6675-707.

10. Shivang Naik, Scott Doyle Shannon Agner, Anant Madabhushi, Michael Feldman and John Tomaszewski, 2008. Automated Gland And Nuclei Segmentation For Grading Of Prostate And Breast Cancer Histopathology, 978-1-4244-2003-2/08/$25.00 ©Ieee.
11. Kumaravel, A., 2013. Web Portal Visits Patterns Predicted By Intuitionistic Fuzzy Approach, Indian Journal of Science And Technology, 6(5s): 4550-4553.
12. Kumaravel, A., 2013. Layered Approach For Predicting Protein Subcellular Localization In Yeast Microarray Data, Indian Journal Of Science And Technology, 6(5s): 4568-4571.
3. Kumaravel, A., 2013. Efficient Molecule Reduction For Drug Design By Intelligent Search Methods, International Journal Of Pharma And Bio Sciences, 4(2): 1023-1029.
14. Kumaravel, A., 2013. Consruction Of Meta Classifiers For Apple Scab Infections, International Journal Of Pharma And Bio Sciences, 4(4): 1207-1213.
15. Kerana Hanirex, D., 2013. Multi-Classification Approach For Detecting Thyroid Attacks, International Journal Of Pharma And Bio Sciences, 4(3): 1246-1251.