Image Enhancement in Histopathological Images

Sandip Mehta

(Department of Electrical and Electronics Engineering, JIET Group of Institutions, Jodhpur, India)

Abstract: *Detection, segmentation and classification of nuclei in routinely stained histopathological images pose a difficult computer vision problem due to variations in dyes concentration, artefacts, noise and damaged nuclei boundaries during the slide preparation process, as imperfections in the staining and scanning of the slide. Furthermore, nuclei are clustered and heterogeneous in terms of both intensity gradient and color, even within the same nuclei. This may be due to uneven activation intensity leading to variable color intensity, the superposition of different colors on tissue layers and the variation of the illumination over the tissue specimen. This paper presents a technique for image enhancement of Histopathological medical image for pre-processing. Initially image intensity is adjusted. Thereafter, it is passed through median filter and converted from RGB to HSV.*

Keywords: *Medical imaging,Histopathological, Median filter, HSV, RGB.*

I. INTRODUCTION

Medical imaging is certainly one of the fields of medicine that has had a true revolution during the last two decades. These recent discoveries not only provide a better diagnosis but also offer new hopes for treatment of many diseases including breast cancer.

Prognostic assessments and successful treatments for breast cancer vary highly depending on the cancer type, stage, treatment and geographical location of the patient. Histopathological examination [1]-][3] is based on the visual observation of chromatin texture, shapes and sizes of nuclei, size of nucleoli, thickness of nuclear membrane, and regularity of nuclear contour of the population of tumor nuclei that can also be analyzed using quantitative image analysis techniques. Mostly, these image analysis techniques provide more objective prognostic clues, which may be insufficiently observed and quantified by human visual examination. Thus, a computer assisted quantitative image analysis in histopathology is likely to improve the diagnostic and prognostic capabilities and boost the efficiency of histopathologists by giving a reliable second opinion. These quantitative tools for tissue characterization are also important for understanding the biological mechanism of disease progression.

The most difficult challenge in quantitative image analysis is represented by the spatial analysis, more specifically by the automated nuclei detection, segmentation and classification. The objective of nuclei classification is to assign different labels to different types of nuclei as normal, cancer, mitotic, apoptosis, lymphocytes etc. Quantitative image analysis in cytology has been studied for years and numerous solutions have thus been proposed in the literature. The application of these solutions to histopathology is rather complicated due to the radical differences between the two imaging modalities and to the highly complex image characteristics. Indeed, in cytology imagery, the detection, segmentation and classification of nuclei are generally facilitated due to the well-separated nuclei and the absence of complicated tissue structures. In contrast, the detection, segmentation and classification of nuclei in histopathology imagery are relatively difficult since most of the nuclei are clustered, being parts of complex structures/architectures (tubules, blood vessels, nerves, muscles, DCIS) and zones/territories (neoplasm, fat, necrosis, connective tissue, hyperplasia, fibrosis) which provide a more comprehensive examination and understanding of the evolution of the disease. These complex structures formulate different challenges for quantitative image analysis. Nevertheless, recent works [4]-[9] show great potential for computer aided diagnostic of histopathological datasets for breast cancer grading.

Nuclei look different due to different tissue, nuclei type, cancer grade and nuclei life cycle. Having importance in cancer diagnosis and grading, these nuclei are broadly classified into two categories depending on nuclei type: lymphocyte and epithelial nuclei. Lymphocyte nuclei are inflammatory nuclei having regular shape and smaller size than epithelial nuclei [see in Fig. 1(a)].

Epithelial nuclei have nearly uniform chromatin distribution with smooth boundary [see Fig. 1(b)]. High grade cancer tissue, epithelial nuclei, often called cancer nuclei, are larger in size, having heterogeneous chromatin distribution, irregular boundaries and clearly visible nucleoli as compared to normal epithelial nuclei [see Fig. 1(c)]. The variation in nuclei shape, size and texture during nucleilife cycle, mitotic nuclei, is another
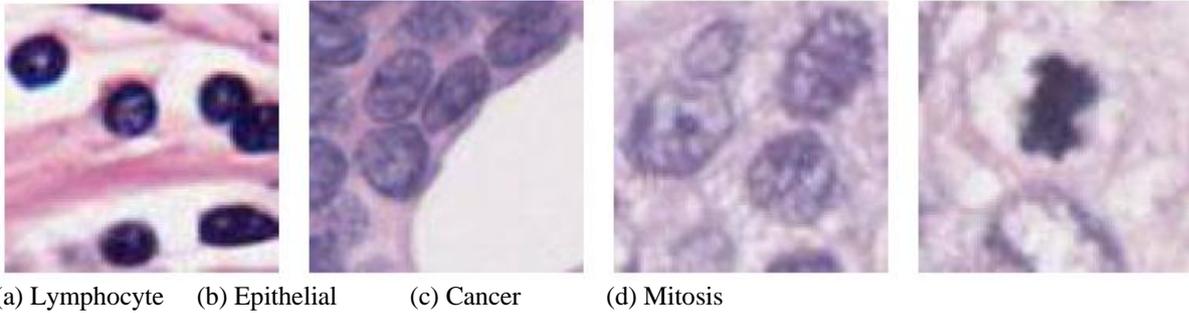
(a) Lymphocyte   (b) Epithelial   (c) Cancer   (d) Mitosis
Fig. 1. Different types of nuclei

factor of complexity [see Fig. 1(d)]. Nuclei detection, segmentation and classification are important steps in cancer diagnosis and grading. The presence of nuclei and their aspect are critical signs for evaluating the existence of disease and its severity. For example, infiltration of lymphocyte in breast cancer histopathology images are related to patient survival and outcome [10]. Similarly, nuclei pleomorphism has diagnostic value for cancer grading. Furthermore, mitosis count is also an important prognostic parameter in breast cancer grading. Therefore, nuclei detection, segmentation and classification are prerequisites to cancer diagnosis and prognosis. Automated nuclei detection, segmentation and classification is now a well-studied topic for which a large number of methods have been described in the literature and new methodologies continue to be investigated.. All these problems, as shown in Fig. 2, make the nuclei detection, segmentation and classification a challenging problem.A successful quantitative image analysis approach will have to overcome these issues in a robust way, in order to maintain a high level in the quality and accuracy of nuclei detection, segmentation and classification.
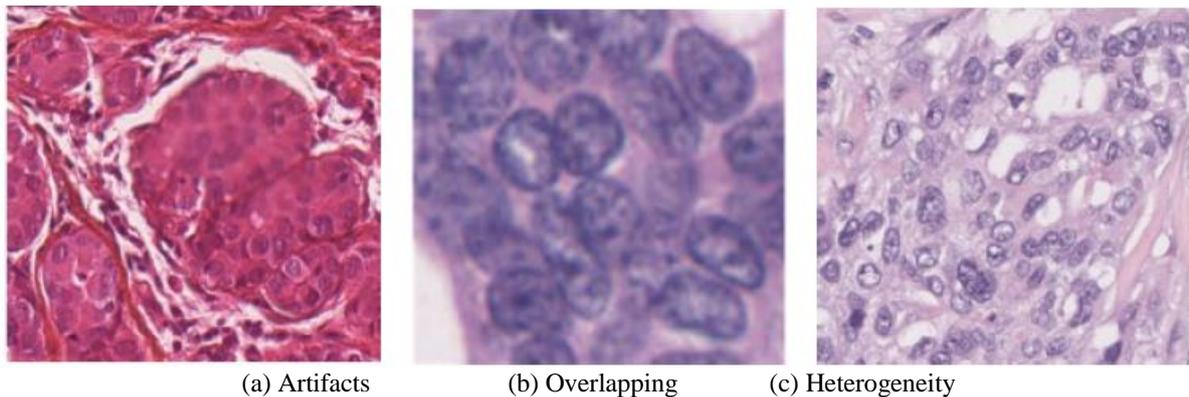

(a) Artifacts   (b) Overlapping   (c) Heterogeneity
Fig. 2. Examples of challenges in nuclei detection and classification

The paper is arranged as follows. The proposed methodology is introduced in section II. The simulation and results are presented in section III while conclusion is given in section IV.

II. PROPOSED METHOD FOR PRE-PROCESSING OF CANCER CELL ENHANCEMENT
The proposed work can be obtained by integer wavelet transform followed by JPEG algorithm. Fig. 3 shows the general architecture of the proposed system.
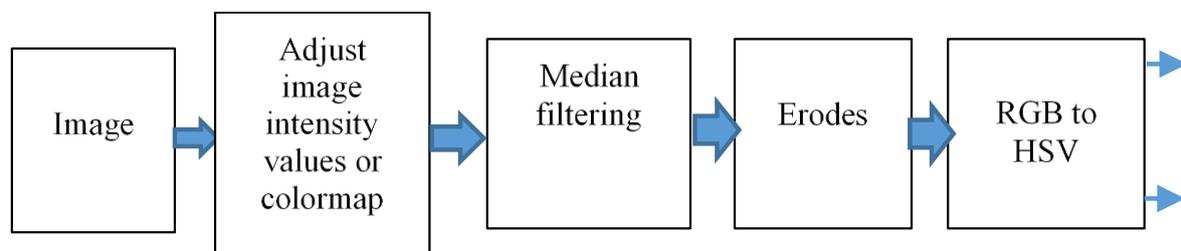

Fig. 3: Block diagram for proposed approach

A. Image Acquisition
The first stage of any vision system is the image acquisition stage. After the image has been obtained, various methods of processing can be applied to the image to perform the many different vision tasks required today. However, if the image has not been acquired satisfactorily then the intended tasks may not be achievable, even with the aid of some form of image enhancement.

B. Median Filtering

Median filtering is a noise reduction tool and is efficient for impulsive and pepper & salt noise cases[11]. It also preserves edges in an image while reducing random noise. Impulsive or salt-and pepper noise can occur due to a random bit error in a communication channel. In a median filter, a window slides along the image, and the median intensity value of the pixels within the window becomes the output intensity of the pixel being processed.

Let $\{X_i\}$ is the m-directional sequence with index $i \in Z^m$. A sliding window $W \in Z^m$ has $W_i = \{X_{i+r} : r \in W\}$ at position i.

The Standard median (SM) filter for output Y_i at i^{th} location is given in the terms of W_i as:

$$Y_i = med\{W_i\} = med\{X_{i+r} : r \in W\} \tag{1}$$

where, $med\{.\}$ is the median operator.

The rth order statistic of the samples inside the window W_i for odd number of r is:

$$[W_i]_{(1)} \leq [W_i]_{(2)} \leq \cdots \leq [W_i]_{(2N+1)} \tag{2}$$

C. Erosion Method

The erosion [12] of the colour image A by the structuring element B is defined by:

$$A \ominus B = \{z \epsilon E | B_z \subseteq A\} \tag{3}$$

where, B_z is the translation of B by the vector z, i.e.

$$B_z = \{b + z | b \in B\}, \forall_z \in E \tag{4}$$

When the structuring element B has a centre (e.g., B is a disk or a square), and this centre is located on the origin of E, then the erosion of A by B can be understood as the locus of points reached by the centre of B when B moves inside A. For example, the erosion of a square of side 10, centred at the origin, by a disc of radius 2, also centred at the origin, is a square of side 6 centred at the origin.

The erosion of A by B is also given by the expression:

$$A \ominus B = \bigcap_{b \in B} A_{-b} \tag{5}$$

D. Convert from RGB to HSI

The pitch (H) and the saturation (S) are given in terms of the equations called chromatic coordinates, which are defined as:

$$r = \frac{R}{R+G+B} \tag{6}$$

$$g = \frac{G}{R+G+B} \tag{7}$$

$$b = \frac{B}{R+G+B} \tag{8}$$

These are nothing other than the normalized values of R, G and B. Adding eqs. (6) to (8) is the equation of the plane in the Euclidean space RGB that passes through the points (1,0,0), (0,1,0) and (0,0,1):

$$r + g + b = 1 \tag{9}$$

The geometric figure on the plane delimited by these points is an equilateral triangle (Fig. 4). It is known that the resulting color range by the combination of three colors is modeled precisely by this type of triangles [13]. The triangle of Fig. 4 corresponds to a triangular double pyramid slice.

The intensity I, is defined as the average of R, G and B:

$$I = \frac{1}{3}(R + G + B) \tag{10}$$

whose value is also in the interval [0, 1].

In order to find the tone, we can see Fig. 4, where the center of the equilateral triangle W has coordinates $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$; w is the vector from the origin to W; the point of color O has coordinates (r_0, g_0, b_0); the vector o goes from the origin to the point O. As mentioned, 'the tone is the angle' formed by the red axis and the vector that ends in a point of color O. With the auxiliary figure we notice that this angle is formed by the vectors $e_r - w$ where e_r is the vector ending at the point (1, 0, 0).
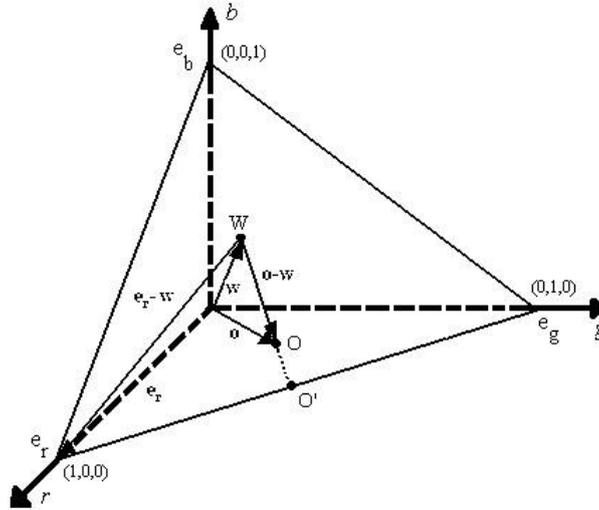
Fig. 4. Plane that cuts Euclidean space at points

Using the equation of the product point between two vectors a•b=$\|a\|\|b\|\cos\theta$ for $0 \leq \theta \leq 180$, we get that the tone H is expressed as follows:

$$H = \cos^{-1}\left[\frac{(o-w).(e_r-w)}{\|o-w\|\|e_r-w\|}\right] \tag{11}$$

By carrying out the operations concerned, we finally obtain:

$$H = \cos^{-1}\left[\frac{\frac{1}{2}[(R-G)+(R-B)]}{\sqrt{(R-G)^2+(R-B)(G-B)}}\right] \tag{12}$$

for $0 \leq H \leq 180$.

From Fig. 4, if b> g, or equivalently if B> G, then H> 180. For this case we first occupy Eq. (12) and then do H = 360 - H.

For the calculation of saturation we extend the line defined by the points W and O until we cut some side of the triangle. The point where we cut it is denoted by O' (Fig. 4). Since saturation is in the closed interval [0,1] and is proportional to the distance from W to O (|WO|), it is defined as:

$$S = \frac{|WO|}{|WO'|} \tag{13}$$

Let us focus our attention on the area of the triangle bounded by the points e_r, W, and e_g. Let us call it A_1 and suppose that O is in this area. To obtain S, we need to know the coordinates of O'. These coordinates can be obtained by solving the system of equations, which is the result of the intersection of the straight line passing through W and O' and which mathematically [3] is expressed as:

$$r = k\left(r_o - \frac{1}{3}\right) + \frac{1}{3}$$

$$g = k\left(g_o - \frac{1}{3}\right) + \frac{1}{3} \tag{14}$$

$$b = k\left(b_o - \frac{1}{3}\right) + \frac{1}{3}$$

(1,0,0), (0,1,0) and (0,0,1) with coordinate axes r, g and b, and equation r + g + b = 1.
With k within the interval [0,1] and the line passing through e_r, e_g whose parametric equation is

$$r = -l + 1 \tag{15}$$

$$g = l \tag{16}$$

$$b = 0 \tag{17}$$

where value of l is [0,1].
Once the system of equations has been solved we discover that O' has coordinates:

$$O' = \left[\frac{1-g_0-2b_0}{1-3b_0}, \frac{g_0-b_0}{1-3b_0}, 0\right] \tag{18}$$

Then, by substituting the coordinates of W, O and O' in Eq. (13)

$$S = 1 - 3b_0 \tag{19}$$

where b_0 is the minimum of r_0, g_0 and b_0 in the area A_1, whereby this equation is transformed into:

$$S = 1 - \frac{3 \min(R,G,B)}{R+G+B} \tag{20}$$

In analogy to the triangular areas A_2 and A_3 defined by the points e_g, W, e_b and e_r, W, and b respectively, we get Eq. (20).

III. SIMULATION AND RESULTS

The simulation results are shown in fig. 6 and fig. 7. The results are in visual or qualitative form.
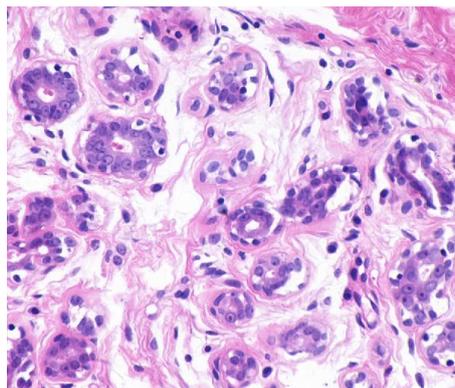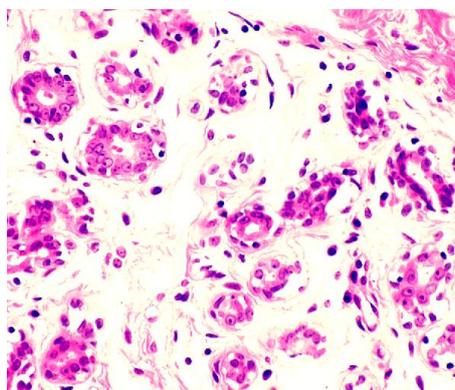

Fig. 5. Original image

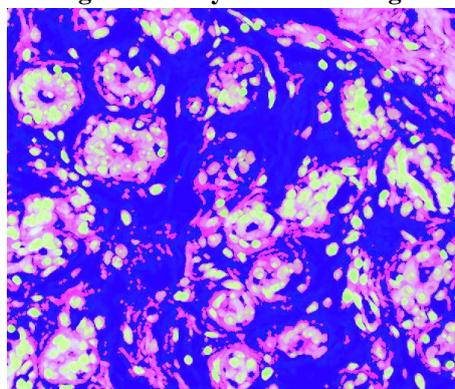**Fig 6. Intensity corrected image**

Fig. 7. HSV converted image

The image enhance process is very essential step for histopathological analysis, once proper image preprocessing is done, then theses image can be further segmented to extract the ROI and based on the extracted feature, it can classified.

IV. CONCLUSION

To avoid uneven activation intensity leading to variable color intensity, the superposition of different colors on tissue layers and the variation of the illumination over the tissue specimen. Initially image has gone through intensity correction then passed through median filter is further after eroding the image it is converted for the color change is made in the HSI color system because of its English meaning (H Hue, S Saturation, I Intensity). To maintain the appearance of the original objects, the matrices S and I are maintained, and the matrix H is transformed. In this way, the resulting image has the same shading effects, changing only the tonality. The HSI system associates the tonality of colors with angles.

References

[1]   Das, K., Karri, S.P.K., Roy, A.G., Chatterjee, J. and Sheet, D., 2017, April. Classifying histopathology whole-slides using fusion of decisions from deep convolutional network on a collection of random multi-views at multi-magnification. In Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on (pp. 1024-1027). IEEE.
[2]   Gupta, V. and Bhavsar, A., 2017, August. An Integrated Multi-scale Model for Breast Cancer Histopathological Image Classification with Joint Colour-Texture Features. In International Conference on Computer Analysis of Images and Patterns (pp. 354-366). Springer, Cham.
[3]   Spanhol, F.A., Cavalin, P.R., Oliveira, L.S., Petitjean, C. and Heutte, L., Deep Features for Breast Cancer Histopathological Image Classification.
[4]   Ali, S. and Madabhushi, A., 2012. An integrated region-, boundary- shape-based active contour for multiple object overlap resolution in histological imagery. IEEE transactions on medical imaging, 31(7), pp.1448-1460.
[5]   Basavanhally, A.N., Ganesan, S., Agner, S., Monaco, J.P., Feldman, M.D., Tomaszewski, J.E., Bhanot, G. and Madabhushi, A., 2010. Computerized image-based detection and grading of lymphocytic infiltration in HER2+ breast cancer histopathology. IEEE Transactions on biomedical engineering, 57(3), pp.642-653.
[6]   Fatakdawala, H., Xu, J., Basavanhally, A., Bhanot, G., Ganesan, S., Feldman, M., Tomaszewski, J.E. and Madabhushi, A., 2010. Expectation–maximization-driven geodesic active contour with overlap resolution (emagacor): Application to lymphocyte segmentation on breast cancer histopathology. IEEE Transactions on Biomedical Engineering, 57(7), pp.1676-1689.
[7]   Malon, C.D. and Cosatto, E., 2013. Classification of mitotic figures with convolutional neural networks and seeded blob features. Journal of pathology informatics, 4.
[8]   Qi, X., Xing, F., Foran, D.J. and Yang, L., 2012. Robust segmentation of overlapping cells in histopathology specimens using parallel seed detection and repulsive level set. IEEE Transactions on Biomedical Engineering, 59(3), pp.754-765.
[9]   Vink, J.P., Van Leeuwen, M.B., Van Deurzen, C.H.M. and De Haan, G., 2013. Efficient nucleus detector in histopathology images. Journal of microscopy, 249(2), pp.124-135.
[10]   Basavanhally, A.N., Ganesan, S., Agner, S., Monaco, J.P., Feldman, M.D., Tomaszewski, J.E., Bhanot, G. and Madabhushi, A., 2010. Computerized image-based detection and grading of lymphocytic infiltration in HER2+ breast cancer histopathology. IEEE Transactions on biomedical engineering, 57(3), pp.642-653.
[11]   Jadhav, A.S., Banerjee, S., Dutta, P.K., Paul, R.R., Pal, M., Banerjee, P., Chaudhuri, K. and Chatterjee, J., 2006, June. Quantitative analysis of histopathological features of precancerous lesion and condition using image processing technique. In Computer-Based Medical Systems, 2006. CBMS 2006. 19th IEEE International Symposium on (pp. 231-236). IEEE.
[12]   Bug, D., Feuerhake, F. and Merhof, D., 2015, March. Foreground Extraction for Histopathological Whole Slide Imaging. In Bildverarbeitungfür die Medizin (pp. 419-424).
[13]   Dikmen, M., Akbas, E., Huang, T.S. and Ahuja, N., 2010, November. Pedestrian recognition with a learned metric. In Asian conference on Computer vision (pp. 501-512). Springer, Berlin, Heidelberg.