Detection of Malignant Tissues by Segmentation of Histology Images using Histograms of Color and Filter Responses

Ms. Akhila E.
Department of Computer Science and Engineering
NCERC, University of Calicut Kerala, India

Ms. Preethymol B.
Department of Computer Science and Engineering
NCERC, University of Calicut Kerala, India

Abstract

In medical science, digital image processing has relevance as several techniques such as MRI, CT-scan, laparoscopic and endoscopic surgeries and cancer diagnosis tools are being used currently. We present an approach for segmentation of images which are having vague boundaries between regions. The traditional segmentation techniques exhibit weak performance on edgeless images such as histology images. In contrast, the recently proposed segmentation framework exposed better results on histology dataset. It modeled images as the occlusions of realizations of textures. This concept directed us to suggest a variety of this framework. Our method achieves segmentation through convolution, factorization and deconvolution using the histograms of filter responses instead of color distributions. Based on the theoretical study, the system reveals the occlusion of textures in histology images and attains the segmentation in a faster manner. Also, we introduce a method to diagnose cancer by looking for abnormal mitosis and mutations. The histology image is classified as normal and cancer affected, for the diagnosis.

Keywords: Filter Response, Convolution, Deconvolution, Occlusion of Textures, Mitosis, Malignant Tissues

I. INTRODUCTION

Image processing is being used in many application areas. Some important areas are computer vision for tasks like navigation of robots, medical field such as disease diagnosis from CT scan and MRI images, remote sensing, video processing, and identification of number plates of moving vehicles etc. Purpose of digital image processing is automating different tasks and image segmentation is an unavoidable step in it. Image Segmentation refers to partitioning an image into several parts, based on textures, color, edges or regions in the image. It aims at dividing the image into visually homogeneous and distinct regions. Color is considered a relevant feature when dealing with the perception of static and moving images. Visual contrast is useful to filter information present in each color component and to distinguish among similar gray-scale intensities. The combination of color and texture have been proved to achieve better results and could be exploited more effectively when integrated than when treated in isolation.

Image segmentation is an important aspect of digital image processing. The image segmentation can give a result that simplifies the presentation of an image and makes the image analysis easier. In segmentation, a label is assigned to every pixel that is having similar characteristics, like color, texture or intensity, which will help to separate the regions and identify the objects and their boundaries. But the issue while processing is the chance of over-segmentation or under-segmentation.

II. RELATED WORKS

Hundreds of image segmentation algorithms are widely used in several areas. Such segmentation techniques may be broadly classified into discontinuity based segmentation and similarity based segmentation [1]. First category includes image segmentation algorithms like edge detection by detecting the edges or pixels between different regions that have rapid transition in intensity are extracted and latter includes segmentation algorithms such as region growing and region splitting and merging or thresholding method. Image segmentation has important role in medical imaging mainly in diagnosing abnormalities in images of human body parts. The way of segmenting image varies from image to image and also depends on the purpose of segmentation.

Fuzzy C-means and K-means clustering methods [2] have been used for medical image segmentation especially in MR Images of human brain. For CT scan images some active shape models [3] can be used and multiple model approaches [4] are there for segmenting ultrasound images of the heart.

Medical field uses different segmentation methods, [5] demonstrates the use of segmentation of punctate patterns in fluorescence microscopic images. Its statistical modeling helps the multiple masks converge from a random initial configuration to a relevant one. Another work in [6] presents a method to segment a large number of cells from 3-D images characterized by non-homogeneous intensity and gradient signal and capable to complete surface discontinuities without any compromise

between precision and ability to integrate the incomplete contours. The segmentation method is a generalized version of the Subjective Surfaces technique.

But most challenging segmentations in medical imaging include the processing of edgeless images such as histology images which consists of bright field microscopy images of hematoxylin and eosin (H&E)-stained slices of tissues. Some segmentation works on histology dataset are [7], which deals with slices of teratoma tumors and [8], where the intensity neighborhoods are used to segment bone, cartilage, and fat. Another method for identifying the histological grade of breast cancer tissue sections based on pattern classification and image analysis algorithms is given in [9]. A framework for unsupervised segmentation of tissue images is given in [10]. Digital pathology is becoming an increasingly important tool for automated biopsy analysis. It has been implemented for different purposes including classification, retrieval, and segmentation, which can further be categorized into two in terms of its objective. The automated analysis of histopathological images improves both throughput and reproducibility.

Our proposed system is capable of segmenting images in histology dataset in an efficient manner. Existing techniques for modeling complicated tissue images include [11], which proposed a supervised segmentation algorithm that modeled images as occlusions of textures and [12] proposed a novel framework for segmenting edgeless images using the concepts of local histograms and image occlusion models. It is an unsupervised segmentation technique based on local histograms of colors. Our proposed work uses the algorithmic framework that works based on histogram factorization and deconvolution mentioned in [12]. But in our work we used the histograms of filter responses instead of color distribution. Around four to five different filters are used to prepare the local histogram transform. The textures in the image are extracted through local histogram transform, non-negative matrix factorization and deconvolution. It works well on broad class of images such as histology images with poorly defined boundaries between regions.

III. BACKGROUND

We follow the framework proposed in [12], which uses a concept that the images can be modeled as the occlusions of textures. That means every image is constituted of a number of textures, which are random-valued images. The term occlusion refers to 'hide something' or 'being hidden'. Here the difficulty to realize the textures in an image is termed as occlusion of textures. In other words, there is a stack of textures with one blocking or occluding the others at each pixel in the image. The segmentation algorithm is responsible for untying those textures.

We summarize some basic concepts, which the framework in [12] used for texture modeling. Let the input image I: $X \rightarrow V$, where X is a discrete set of pixel locations and V is a discrete set of pixel values. A texture is a random- valued image and it becomes flat when the distribution of values at every pixel is uniform. Now the local histogram transforms and occlusion operators have to be defined. The local histogram transform gives the fraction of pixels in m×n neighborhood around x that have value, when convolution is performed using a window of size m×n. A labeling function is to be estimated in such a way that the perfect segmentation is achieved. The image will be modeled as occlusion of realization of set of textures with respect to this labeling function. Fig. 1. illustrates the idea.

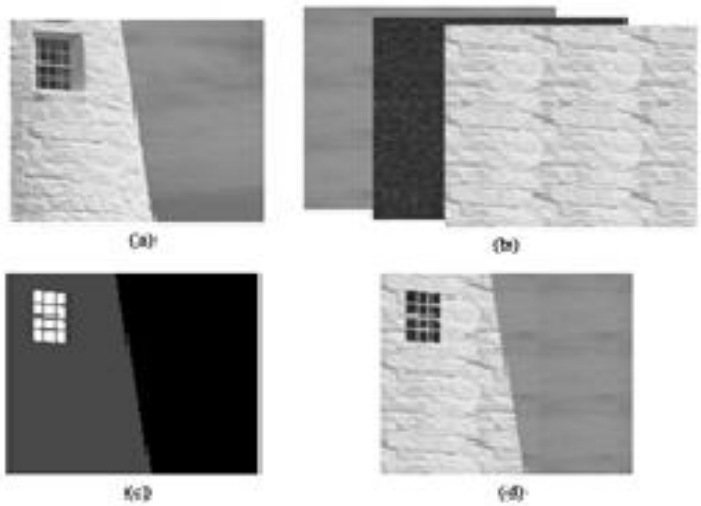

Fig. 1: (a) Input image (b) occlusion of realization of textures (c) labeling function (d) segmented image

IV. SEGMENTATION BASED ON OCCLUSION OF RANDOM TEXTURES

The images in histology dataset are edgeless images and their analysis is relatively difficult. ORTSEG [12] is an algorithm proposed for this purpose. It works effectively on histology images and brings light into the area of medical imaging. The ORTSEG framework works purely based on color distribution in images. The objective is to find out a labeling function such

that the images can be demonstrated as occlusions of realizations of random textures in it. The main tool used for segmentation is the local histogram transform of the textures. The edge detection is achieved by focusing on color distribution of regions rather than edges. Nonnegative matrix factorization and parametric deconvolution also carry important role in the algorithm. The framework performs well on synthetic texture mosaics and random textures. But exhibits pure performance on Prague dataset.

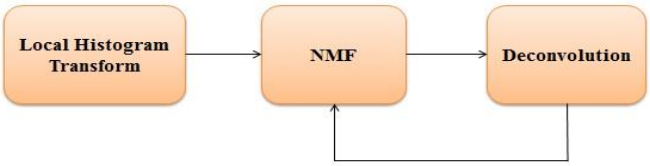

Fig. 2: Block diagram for ORTSEG algorithm

Initially the histology image is quantized to avoid the problem with use of memory. K-Means clustering is used for reducing the color space. Fig. 2. shows the block diagram for ORTSEG algorithm. The local histogram transform is the main tool used in the algorithm and is calculated by single convolution. Larger window size is preferred as the resulting image gives the blurred version and it helps to alleviate the influence of noise. The local histograms are factorized into weights and histograms using any of the nonnegative matrix factorization methods. The approach recommends ALS algorithm for factorization.

The weights are then reshaped into image of weights which corresponds to blurred version of one level of labeling function. It is then deconvolved for recovering the labeling function. The regions to be labeled are the Voronoi cells of three seed pixels and deconvolve using gradient descent those pixels.

Even though the algorithm gives better results in the case of edgeless images, it performs worse on histology images as compared to random textures, because histology images are much more complex. On Prague dataset ORTSEG performs very poor compared to JSEG [13]. The reason is that the design of ORTSEG does not detect sharp edges between regions while others do and the Prague images are having increased scale of textures.

V. SEGMENTATION BASED ON FILTER RESPONSES

The proposed technique works exactly based on the histograms of filter responses. Fig. 3. Exhibits the working environment of the proposed system.

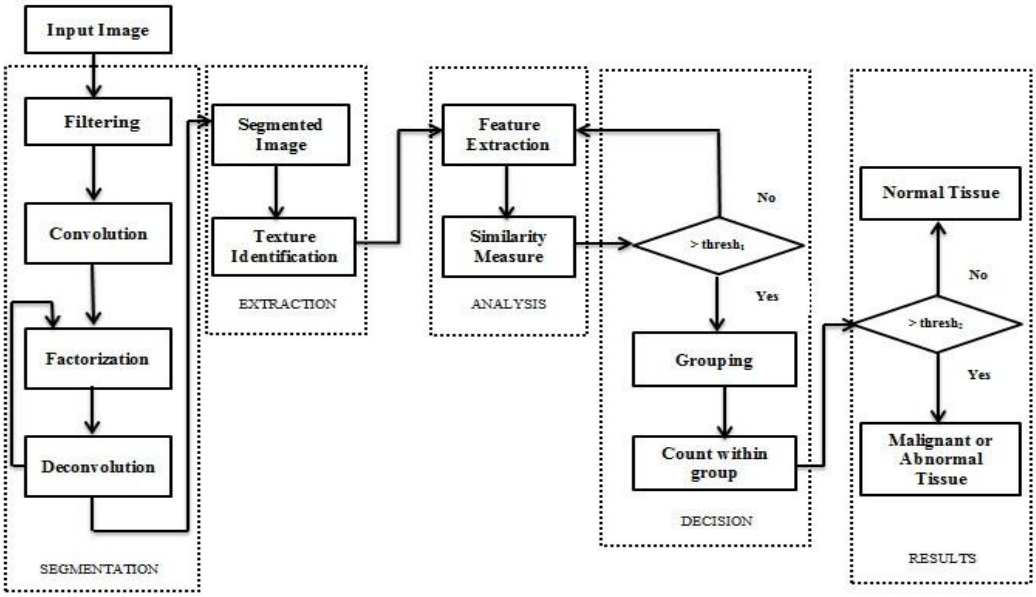

Fig. 3: Block diagram for proposed system

Our method concentrates on histology dataset. The blurring and deblurring is done accordingly on the input histology image. Both are attained by filtering operations. Filter responses are capable of characterizing the image features like intensity or textures in an effective way compared to the color distributions in the image. Filtering can be done in both frequency and spatial domain.

A. Filtering

Filtering is a technique for modifying or enhancing the characteristics of an image. Several filters can be applied based on the requirement. Low-pass filters are responsible for smoothing or blurring and high-pass filters are responsible for sharpening. The

use of filters can improve the performance of entire system as it reduces the delay and provide fast processing. Filters are much helpful in edge detection operations.

In this system five different filters are used. First is the Gaussian filter, which is a low-pass filter, second is the Wiener filter which is for noise removal, third is Gabor filter that helps for edge detection, fourth one is the Sobel filter that is for edge detection and finally the Log filter that helps to make the image more clear.

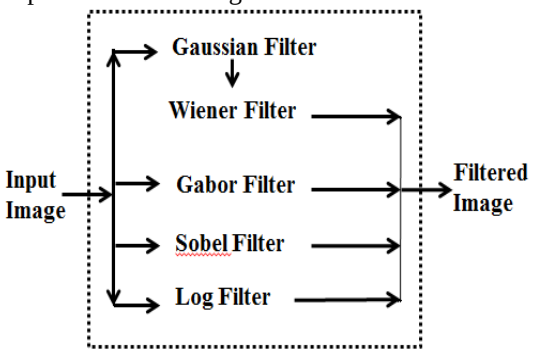

Fig. 4: Filtering

Fig. 4. shows the step of filtering in detail. Five filters have been used here. Three different possible combinations of these five filters are selected. Purpose of applying these filters on input image is given below.

1) Gaussian Filter:

Gaussian filters are low-pass filters which are normally used for noise removal. Here we selected this filter to get the blurred version of the image. Initially the input histology image is filtered using the Gaussian filter. The filtered image is the convolved image which is then input to the next filter.

2) Wiener Filter:

Wiener filter removes blur in the image. The effect is deconvolution and noise reduction. The resulting histology image is clear compared to the input image. Sequential application of Gaussian and Wiener filters gives enhanced histology image, which is then passed through next filter.

3) Gabor Filter:

It is a linear filter, which is commonly used for edge detection and texture analysis. It gives a result, of which the orientation and frequency are much similar to those of human visual system and so it is very useful for texture representation and texture discrimination.

4) Sobel Filter:

It is for horizontal edge detection by applying a horizontal mask. It is a discrete differentiation operator that works based on convolving the image. The resulting image returns the edges in the image clearly.

5) Log Filter:

Log filter is the Laplacian of the Gaussian filter. Laplacian is applied to an image which has been smoothened using Gaussian filter in order to reduce noise. It enhances the image characteristics. The resulting image gives the improved features in the histopathology image.

The Gabor, Sobel and Log filters are applied separately on the input histology image. Four different filtered image components are formed like result of sequential application of Gaussian and Wiener Filters, result of Gabor filtering, result of Sobel filtering and finally the result of Log filtering. Then the final filtered image is formed by concatenating three among these results appropriately. Three concatenations are possible here, and segmentation using each is implemented separately as three different cases.

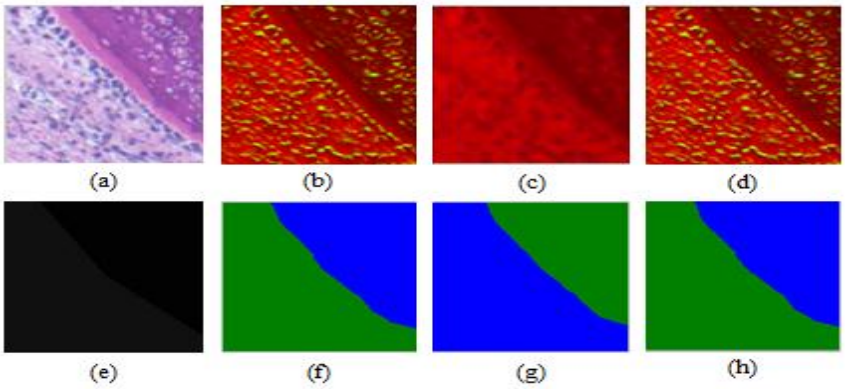

Fig. 5: (a) Input image, (b)case 1 result, (c) case 2 result, (d)case 3 result, (e)ground truth, (f) segmented image using (b), (g), segmented image using (c), (h) segmented image using (d).

Case 1 uses the filtered image formed by combining the outputs of Wiener, Sobel and Gabor filters. Case 2 uses the combination of Wiener, Gabor and Log filters and finally case 3 uses the outputs of Wiener, Sobel and Log filters. As these filters are applied appropriately, the image quality gets improved and the output image is processed instead of considering simply the color distribution. Fig. 5. Illustrates the results.

B. Convolution

Purpose of convolution is smoothing the image so that we get the blurred version of the histology image. The smoothing filters can be used to do this transformation. We recommend Gaussian filter [14], but any smoothing filter can be used. It removes the noise in the image and makes it blurred. Using the Gaussian filter the pixel [x, y] is weighted the distance of the neighborhood pixel [xk, yk] from the center pixel [xc, yc] of the output image where the filter is being applied. In frequency domain low pass filters can be used for smoothing and high pass filters do sharpening. Here we can select low pass filters to work in frequency domain. It reduces delay because there is no need for quantization. Filtering in frequency domain is the Fourier transform of filtering in the spatial domain. The simple multiplication in Fourier space is equivalent to the convolution in spatial domain as a result it is possible get the blurred version of image. The histogram transforms of this convolved image is then calculated.

C. Factorization

Factorization is the process of dividing the transformed image into weights and actual histogram values. Any method available for factorization can be used. Our aim is to separate the effect of the weights on the blurred version of image. So this will be used by the deconvolution step.

D. Deconvolution

It is the inverse process of convolution. The weights obtained as a result of factorization are used along with the sharpening filter. Sharpening filters are high pass filters that result in deblurring. We suggest Wiener filters [15], [16] for the reconstruction of the characteristics of the image. The aim is to estimate the labeling function such that the effect of weight on the blurred version of the histology image is minimal. Such a labeling gives the best result.

VI. IDENTIFYING NORMAL AND MALIGNANT TISSUES

The textures separated by the process of segmentation are analyzed for detecting the presence of abnormality in mitosis [16], [17] and malignancy [18]. In tissues there may be abnormal cell division which is referred to as the abnormal mitosis. It can lead to malignancy which is a symptom of cancer. The abnormal mitosis can be a reason for mutations [19] also. Mutations have role in normal as well as abnormal biological processes such as evolution and cancer.

The histology image is classified into normal and cancer affected, for the judgment. Fig. 6. shows an example of normal and abnormal tissues, which exhibits a change in texture as the tissue progresses towards malignancy. It is clear from the figure that the texture becomes finer as the nuclei spread across the tissue. The relationship between the characteristics of textures in the tissue and the presence of tumor malignancy has been confirmed in several studies [20]. There are two categories of statistical classifiers that are parametric and nonparametric classifiers. Parametric methods depend on the hypothesis that the functional form of class-conditional distributions is known and nonparametric methods make conventions about the form of the distributions. Both parameters together can also be used.

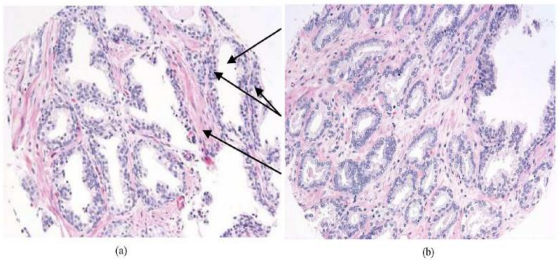

Fig. 6: Example of normal and abnormal tissues

The different images are analyzed and measure the similarity among them. The problem with this model is that both the similarity measurement and the decision making are purely dependent on the fixed threshold, which may cause the degradation of performance.

A. Analysis

From the segmented image we have to extract the necessary features. The features, which are helpful in differentiating the normal and abnormal tissues, must be identified properly. For the identification of tumors and cancerous tissues we seek for the similarity between patterns in the tissues by analyzing these extracted features.

1) Feature Extraction:

For each pixel, features like hue, saturation and intensity of the pixels, mean of hue, saturation and intensity, standard deviation of hue, saturation and intensity, intensity of pixel in Gaussian blurred planes etc are extracted and estimated using Fuzzy inferencing technique. For feature selection we can use sequential forward floating selection which starts with an empty feature set and adds or removes features when this improves performance.

2) Similarity Measure:

Based on the selected and extracted features we can perform comparison between two tissue textures. For that comparison we can use the above mentioned features along with any similarity measure. Here we can use SSIM index. For measuring the similarity between two images the structural similarity (SSIM) index is more suitable. The SSIM index is a full reference metric. Based on an initial uncompressed or distortion-free image the image quality can be measured. SSIM is designed to improve on traditional methods like peak signal-to-noise ratio (PSNR) and mean squared error (MSE), which have proven to be inconsistent with human eye perception.

The result of similarity measure will be checked against a threshold for comparison. If the value of similarity measure is less, we will go for another pair of textures. Otherwise the textures will be tagged as similar and will be inserted to a labeled group.

B. Decision Making

To decide whether there is an abnormal mitosis, the number of regions can be compared with a threshold value. It helps to predict the chance for malignancy.

The similarity index obtained is checked against the thres1 which is a fixed integer. And if the index value is lesser another pair of textures is chosen and the steps are repeated. Else a group is formed and labeled to indicate the similarity of textures in it. This is termed as grouping. After grouping is completed we have to count the number of textures in each labeled group. If the count of textures in the groups falls beyond a limit, i.e. thres2, then malignancy is detected and the tissue is classified as ABNORMAL or MALIGNANT. Otherwise it is grouped under NORMAL. These steps are repeated till no textures are left out in the segmented image. The system directs the pathologist to make a decision on the diagnosis process.

VII. RESULTS AND DISCUSSIONS

As this segmentation framework works based on the filter responses on histology images, the results show better performance compared to the existing algorithm ORTSEG. We implemented this concept on histology dataset using Matlab. Among the three different cases of filtering, the quality of filtered image depends on the input image selected for segmentation and the performance of entire system varies. The results are compared with the ground truth image to finalize filtered image. We identified the best case of filtering for each image in the histology dataset and the corresponding filtered image is used for finding the local histogram transform. Then the segmented output is showing very good performance compared to that of ORTSEG. As soon as the segmentation produces a good output the classification is also capable of giving improved result.

VIII. CONCLUSION

From the recently existing framework for segmentation, we began with the idea of characterizing textures in histology image in a different way. The system is different from the framework that it works based on histograms of filter responses instead of color distributions. The convolution, nonnegative factorization and deconvolution are sequentially performed for separating the textures in the image. Filtering can be done both in spatial and frequency domain. The method concentrates on the edgeless images especially the histology images. The result of this segmentation technique is then analyzed for detecting the influence of tumors and cancer. In future we plan to continue the work with variety of filters for convolution and deconvolution.

ACKNOWLEDGMENT

Authors would like to thank Mr. Michael T. McCann, Mr. Mr. Dustin G. Mixon, Mr. Matthew C. Fickus, Mr. Carlos A. Castro, Mr. John A. Ozolek, andMs. Jelena Kovaˇcevi´c whose work inspired us to design this system.

REFERENCES

[1] Rafael C. Gonzalez, Richard E. Woods, “Digital Image Processing”, 2nd ed., Beijing: Publishing House of Electronics Industry, 2007.
[2] Ajala Funmilola A , Oke O.A, Adedeji T.O, Alade O.M, Adewusi E.A. “Fuzzy k c-means Clustering Algorithm for Medical Image Segmentation” Journal of Information Engineering and Applications ISSN 2224-5782 Vol 2, No.6, 2012
[3] T. F. Cootes, A. Hill, C. J. Taylor, and J. Haslam, “Use of active shape models for locating structures in medical images,” Image Vis. Comput., vol. 12, no. 6, pp. 355–365, Jul. 1994.
[4] J. Nascimento and J. Marques, “Robust shape tracking with multiple models in ultrasound images,” IEEE Trans. Image Process., vol. 17, no. 3, pp. 392–406, Mar. 2008.
[5] G. Srinivasa, M. C. Fickus, Y. Guo, A. D. Linstedt, and J. Kovaˇcevi´c, “Active mask segmentation of fluorescence microscope images,” IEEE Trans. Image Process., vol. 18, no. 8, pp. 1817–1829, Aug. 2009

[6] C. Zanella, M. Campana, B. Rizzi, C. Melani, G. Sanguinetti, P. Bourgine, et al., "Cells segmentation from 3-D confocal images of early zebrafish embryogenesis," IEEE Trans. Image Process., vol. 19, no. 3, pp. 770–781, Mar. 2010.
[7] J. A. Ozolek and C. A. Castro, "Teratomas Derived from Embryonic Stem Cells as Models for Embryonic Development, Disease, and Tumorigenesis", Vellore, India: InTech., 2011, ch. 13.
[8] C. Chen, J. A. Ozolek, W. Wang, and G. K. Rohde, "A general system for automatic biomedical image segmentation using intensity neighborhoods," Int. J. Biomed. Imag., vol. 2011, pp. 606857-1–606857-12, Jan. 2011, doi:10.1155/2011/606857
[9] C. Loukas, S. Kostopoulos, A. Tanoglidi, D. Glotsos, C. Sfikas, and D. Cavouras, "Breast Cancer Characterization Based on Image Classification of Tissue Sections Visualized under Low Magnification," Hindawi Publishing Corporation Computational and Mathematical Methods in Medicine, Volume 2013, Article ID 829461, August 2013.
[10] A. Cagri, A. Burak, C. Aykanat, C. Sokmensuer, and C. Gunduz-Demir, "Multilevel segmentation of histopathological images using cooccurrence of tissue objects," IEEE Trans. Biomed. Eng., vol. 59, no. 6, pp. 1681–1690, Jun. 2012.
[11] M. L. Massar, R. Bhagavatula, M. Fickus, and J. Kova˘cevi´c, "Local histograms and image occlusion models," Appl. Comput. Harmon. Anal., vol. 34, no. 3, pp. 469–487, May 2013.
[12] Michael T. McCann, Dustin G. Mixon, Matthew C. Fickus, Carlos A. Castro, John A. Ozolek, and Jelena Kova˘cevi´c, "Images as Occlusions of Textures: A Framework for Segmentation, " IEEE Trans. Image Processing, vol. 23, no. 5, May 2014.
[13] Y. Deng and B. S. Manjunath, "Unsupervised segmentation of colortexture regions in images and video," IEEE Trans. Pattern Anal. Mach. Intell., vol. 23, no. 8, pp. 800–810, Aug. 2001.
[14] G. Deng and L. W. Cahill, "An Adaptive Gaussian Filter For Noise Reduction and Edge Detection," Department of Electronic Engineering, La Trobe University Bundoora Victoria 3083 Australia.
[15] Prodip Biswa,s Abu Sufian Sarka,r Mohammed Mynuddin, "Deblurring Images using a Wiener Filter," International Journal of Computer Application, Volume 109 - Number 7, 2015.
[16] Chikako Abe, Tetsuya Shimamura, "Iterative Edge-Preserving Adaptive Wiener Filter for Image Denoising," International Journal of Computer and Electrical Engineering, Vol. 4, No. 4, August 2012.
[17] Dan C. Cires¸an, Alessandro Giusti, Luca M. Gambardella, Jurgen Schmidhuber, "Mitosis Detection in Breast Cancer Histology Images with Deep Neural Networks," IDSIA, Dalle Molle Institute for Artificial Intelligence, USI-SUPSI, Lugano, Switzerland.
[18] E.J. Kaman, A.W.M. Smeulders, P.W. Verbeek, I.T. Young, and J.P.A. Baak , "Image Processing for Mitoses in Sections of Breast Cancer: A Feasibility Study," Pattern Recognition Group, Department of Applied Physics, Delft University of Technolog, 1983.
[19] Claire F. Taylor and Graham R. Taylor, "Current and Emerging Techniques for Diagnostic Mutation Detection: An Overview of Methods for Mutation Detection," Elles, R. (Ed.), Springer, 2004.
[20] Y. Smith, G. Zajieck, M. Werman, G. Pizov, and Y. Sherman, "Similarity measurement method for the classification of architecturally differentiated images," Comp. Biomed. Res., vol. 32, pp. 1–12, 1999.