ISSN (Print) : 0974-6846
Indian Journal of Science and Technology, Vol 8(34), DOI: 10.17485/ijst/2015/v8i34/79369, December 2015 ISSN (Online) : 0974-5645

Feature Extraction from Immunohistochemistry Images to Classify ER/PR Scores

P. Smitha[1*], Varghese Paul[2], L. Shajy[1] and K. Sujathan[3]

[1]Department of Computer Science and Engineering, College of Engineering, Karunagappally, Kollam - 690523, Kerala, India; smithakannur@gmail.com, shajikalluvila@gmail.com
[2]TocH Institute of Science and Technology, Arakkonam, Ernakulam - 682313, Kerala, India; vp.itcusat@gmail.com
[3]Division of Research, Regional Cancer Centre, Thiruvananthapuram - 695011, Kerala, India; ksujathan@gmail.com

Abstract

Objectives: Abnormalities of protein receptors in the cell induce cancer. Detection of protein receptors such as Estrogen Receptor (ER)/Progesterone Receptor (PR) helps in hormone treatment, which improves the prognosis factor. **Methods:** Immunohistochemistry stained breast cytology images are used for finding the protein receptors. Separate stains are used for finding each receptor status. The presence of receptors is identified based on the brown color present in the nucleus. Brown color extracted through the channel separation, thresholding and relevant features are obtained from Gray Level Co-occurrence Matrix (GLCM). Based on these feature values an Artificial Neural Network (ANN) will classify the scores. **Findings**: Manual procedure for ER/PR scoring is based on the value of HSCORE, which is calculated by counting the brown colored nuclei and its intensity levels by the pathologist. This is a subjective procedure and has the risk of human fatigue errors. The medical expert decides the treatment plan based on the scores. Here we developed a new technique by which the manual scoring process could be imitated using the optimal set of features through an Artificial Neural Network (ANN), and obtained a result of 95.52 percent. **Application**: This could be a step towards the automation of Immunohistochemistry images and help in the survival of the patients. Hormone treatments are costlier procedure because of, the large amount of data to be processed manually.

Keywords: Estrogen Receptors, Feature Extraction, Immunohistochemistry, Progesterone Receptors

1. Introduction

Magnetic Resonance Imaging (MRI) can be used for finding the high risk of developing breast cancer in women. Fine Needle Aspiration Cytology (FNAC) is a diagnostic method for investigating superficial lumps or masses. From the mass a thin, hollow needle is inserted and samples of the cells are taken. Figure 1 shows an example of a FNAC image.

Immunohistochemistry (IHC) is a special staining process to determine the presence of hormone receptors on the surface of cells. Both Estrogen receptors (ER) and Progesterone Receptors (PR) give more information for diagnosis and prognosis. Different kinds of stains are available to identify the presence of different receptors. Use of stain is based on the application. Staining is necessary because cells and tissues are essentially transparent when viewed with transmitted light as is commonly employed in routine microscopy.

Special stains are simply stain mixtures or combinations that provide selective or enhanced contrast for specific structures within the cell or tissue. This receptor presence can be measured using a scoring method. ER/PR scoring is based on the value of HSCORE. So by analyzing the

* Author for correspondence

hormone on the nuclei and protein receptor the cell surface it is possible to identify the severity of cancer and make sure the appropriate treatment.

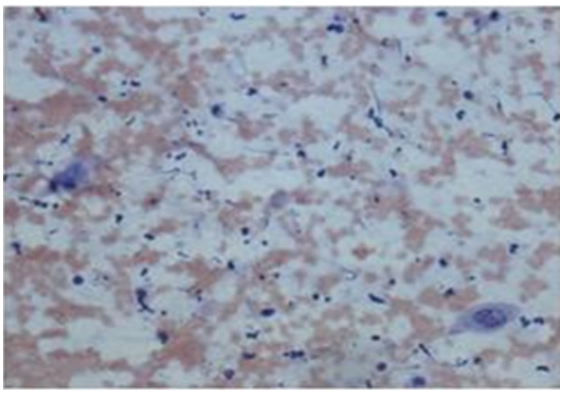


Figure 1. Cytological Images of Breast.

Table 1. Allred Scoring Method

Proportion of Score	Observation values	Intensity Score given	Final Observation
0	NONE	0	None
1	1%	1	Weak
2	1-10%	2	Intermediate
3	10-33%	3	Strong
4	33-66%		
5	66-100%		
Total Score		**Interpretation**	
0-2		Negative	
3-8		Positive	

Breast cancer scoring based on the hormone receptors like ER and PR protein gives more information about the cancer condition and the possibility of treatment. In[1] discusses about the different features used in medical images and the classification techniques used for this purpose. How to score HER2 images based on features are discussed in paper[2]. Importance of ER, PR and HER2 receptors and its comparisons are discussed in[3]. Paper[4] presents about how cancer types related on the histopathological features. In[5] reported the correlation of cancer subtypes and auxiliary lymph node status. The work of[5,3] describes the types of breast cancer based on the receptors and its relations. In[6] proposed a method called Allred scoring, which differentiate ER+ and ER- based on proportion of positive cells and staining intensity and is the base for all ER/PR scoring. Allred method is the standard method used by pathologists for the scoring of ER/PR protein. It is a semi quantitative method in which it uses proportion of positive cells and the staining intensity. Positive cells are scored from 0 to 5 and staining intensity from 0 to 3. The two values are summed to get a value from 0 or 2 to 8. A score from 0-2 is regarded as negative and 3-8 is considered as positive. Table 1 explains the Allred scoring method and its interpretation.

In[7] mentions adaptive thresholding for the segmentation step. According to[8,10] and color deconvolution method is better for ER/PR scoring. In[9] discusses about the usage of color spaces for segmenting lung granular cells. In[11,16] authors proposes RGB to HSI color space as better method for cytological image preprocessing. In[12] uses Multi layer perceptron and Radial bias function networks for breast image classification. In[8] HSCORE calculation equation is used for ER/PR classification. In paper [17,21] also explain color decomposition method for histopathological images.

Segmentation method is the critical step in the scoring process because of the identification of cell nuclei is necessary[21] suggest system architecture with Deconvolution for stain separation and Otsu's method for segmentation to find HSCORE for finding ER/PR scoring using software. This was a good attempt and succeeded with average ER accuracy of 82.89% and average PR accuracy of 85.64% results. This result should be improved for better prognosis. Clustering based segmentation method is preferred by[13]. Mathemetical morphology and its importance are discussed in paper[15]. Some papers discusses about the different preprocessing techniques applicable for medical images[16]. Deconvolution plays an important role in cytology images which separates the staining process[17,21]. It is a time consuming task and complexity plays importance in automated systems. Different texture analysis methods are discussed clearly in paper[18,19,20].

In enhancement of immunohistochemistry images with our previous work it is clear except some enhancement techniques all are reducing the information available in the resulting image. If the images are taken with proper staining then there is no need of any enhancement techniques.

Count of positively stained cells are referred as, features in the work of[6, 7]. Mean, hue variance, mean hue and mean saturation are the features used by[11]. In[13] signal energy is used as a feature to determine HER2 status. So GLCM is used as feature extraction method. In[14] authors used wavelets for extraction of nuclei texture features. Different feature extraction methods used for grading is discussed in the thesis work[22]. It also discusses about

all the stages of image processing used for grading breast cancer images.

In paper[23] suggested an adaptive neural network for tumor classification. Classification is the final step and[11,24] uses back propagation neural network for the classification of HER2 status. There are other receptors like P53 which can also cause aggressive cancers and it should also included in further study[25].

2. Materials and Methods

2.1 Materials

Fine-Needle Aspiration Biopsy (FNAB), is a diagnostic procedure for investigating superficial lumps or masses. In this technique, hollow needle is inserted into the mass for sampling of cells. That is then stained and examined under a microscope. Cytology images could be of histological (biopsy) images. Biopsy images are very safe and offer minor surgical procedure. If an abnormality is detected on an imaging test such as x-ray, ultra-sound or mammography then FNAC or FNAB is suggested. In this work a FNAB is used.

Immunohistochemistry is the process of detecting proteins from cells of the tissue section. Preparation of samples requires proper tissue collection, fixation and sectioning. A solution of paraformaldehyde is normally used to fix tissue. The tissue is then sliced dependent upon the purpose of the experiment. Sections can be sliced on a microtome, and sliced at a range of 4 micro meter. The slices are then mounted on slides, then stains are applied for the duration specified (Hematoxylin) with different stains for ER (ER Clone SP1 Rabbit Mono) and PR (PR Clone Rabbit Mono), then dehydrated using alcohol washes of increasing concentrations and water contents are removed using a xylene solution then used under the microscope for imaging.

2.2 Methods

The scoring method involves different stages like preprocessing, segmentation, feature extraction and classification for scoring. The following section gives an overview about the methods for these steps.

2.2.1 Preprocessing

Preprocessing stage is used to enhance the visual appearance of images and to improve the manipulation of datasets. Color space conversion is used here as the preprocessing method. A color space is a specific organization of colors. Color model is a specification of coordinate system and a sub-space within that system where each color is represented by a single point. For the preprocessing method, separation of red, green and blue channels are used. The infected nuclei are in brown color, so the red channel is extracted for further calculation.

If the segmented image contains no brown nuclei or less than certain number of connected components, it is classified as negative. Otherwise find all the connected components in the segmented image and calculate the gray levels in each connected component. From these identified gray levels find the gray level value for each group. From these gray level values, found a value which is suitable to divide the image into different categories.

2.2.2 Segmentation

Image segmentation partitions a digital image into multiple segments. Segmentation is performed to simplify or change the representation of an image so that image becomes more meaningful and easier for further analysis. Then thresholding is applied on that red channel for segmenting the nuclei. After the segmentation process, we get the brown nuclei. Thresholding is the simplest method among segmentation methods.

2.2.3 Feature Extraction

Feature extraction is the process of extracting features from the enormous set of potentially useful features which may be available in a given problem domain. Features are numerical values computed from each image properties which are present or absent. For ER/PR classification energy feature is used as the criteria. For getting energy measure GLCM (Gray Level Co-occurrence Matrix) is used. The GLCM is a tabulation of how often different combinations of pixel brightness values (gray levels) occur in an image.

$$p(i,j) = \frac{v_{i,j}}{\sum_{i,j}^{n} v_{i,j}} \quad (1)$$

Where 'i' is the row number and 'j' is the column number. The following features are extracted using this method.

Contrast: This captures the dynamic range of gray levels from an image, with the polarization of the

distribution of black and white pixels.

$$Contrast = \sum_{i,j} |i - j|^2 P(i,j) \quad (2)$$

Homogeneity: It is largely related to the local information extracted from an image and reacts how uniform a region is.

$$Homogeneity = \sum_{i,j} \frac{1}{1-(i-j)^2} p(i,j) \quad (3)$$

Correlation: This is a measure of how connected pixel is to its neighbor over the whole image. It calculates the correlation of each attribute with the label attribute and returns the absolute or squared value as its weight.

$$Correlation = \sum_{i=0}^{G-1} \sum_{i=0}^{G-1} \left((i * j) p(i,j) - \{\mu_x - \mu_y\} \right) / \left(\sigma_x * \sigma_y \right) \quad (4)$$

Energy: Energy is the Angular Second Moment.

$$Energy = \sum_{i,j} p(i,j)^2 \quad (5)$$

Entropy: Entropy is the measure of randomness. Inhomogeneous scenes have low first order entropy, while a homogeneous scene has high entropy.

$$Entropy = -\sum_{i,j} p(i,j) \log (p(i,j)) \quad (6)$$

2.2.4. Classification

Artificial Neural Networks (ANNs) are computational models of biological neural networks. This neural network is provided with large number of unknown inputs and used to estimate or approximate functions. Artificial neural networks are system of interconnected "neurons" which are capable of machine learning and pattern recognition and can compute values from the inputs. The main advantage of ANN is their adaptive nature. There are many types of artificial neural networks. In this work Pattern network one of the feed forward network is used. This network classifies based on the input vector and the target vector specified.

Manual way for classification is to find the HSCORE. This can be done by dividing the gray values into 4 categories. First category contains very dark gray values. Second group contain next dark gray values. Third group contains next level of gray values and the last group contains remaining values up to the maximum. After computing the categories, count of each category is calculated and total count is computed to find the percentage of each category. Percentages of each category is denoted as p1, p2, p3 and p4, based on these values HSCORE is calculated with the category weight value as follows

$$HSCORE = (p1*3) + (p2*2) + (p3*1) + (p4*0) \quad (7)$$

Classification is done based on the following Table 2. Score 1 is taken as negative and the remaining scores are determined as positive.

Table 2. ER/PR score calculation from HSCORE.

Resultant Score	Condition
Score 1	if HSCORE<50
Score 2	If 51<HSCORE<=100
Score 3	if 101<HSCORE<=200
Score 4	if HSCORE>200

3. Result and Discussion

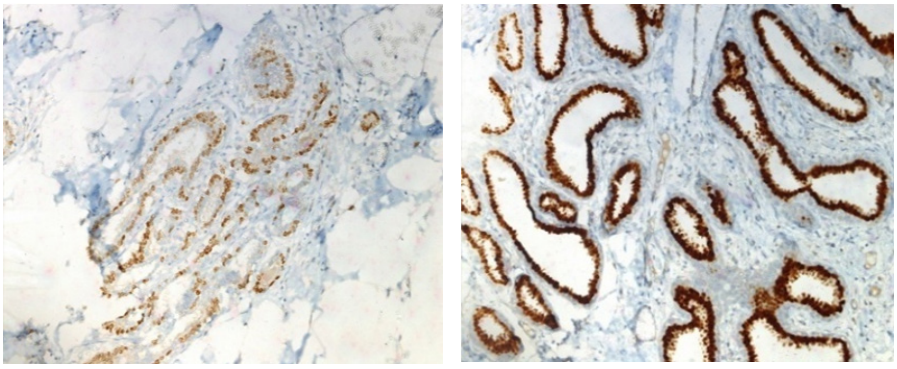
MATLAB is used as the software tool for implementing this work. Here we used images from a private scanning centre. The results are verified and authenticated by the experts from RCC, Trivandrum. In this work we used ER images of 15 different patients, output are generated with 66 images. Here we used some GLCM texture features with entropy value for training with the Pattern Network. Total five features are extracted and used for classification.

Input and segmented ER images are shown in figure below. Figure 3. shows the sample ER and PR image and its segmented results are shown in Figure 4. In this segmented image the brown nuclei is separated for further processing. After the feature extraction ANN is trained based on this feature set and the accuracy, precision, sensitivity and specificity values are noted. Then feature set is reduced to repeat the same procedure but got increasing accuracy, sensitivity and other parameters. While using energy and entropy as features accuracy increased to 95.52%.

Table 3 shows different precision, sensitivity and specificity values noted against different feature sets. For the classification purpose a multiclass classification is needed with more than 2 classes. Multiclass pattern network is used for this purpose. In single feature cases energy feature misclassifies some classes and this is avoided by including the entropy feature. While using entropy feature alone also provides the same situation in some other classes. So for getting perfect results a combination of energy and entropy features are used and

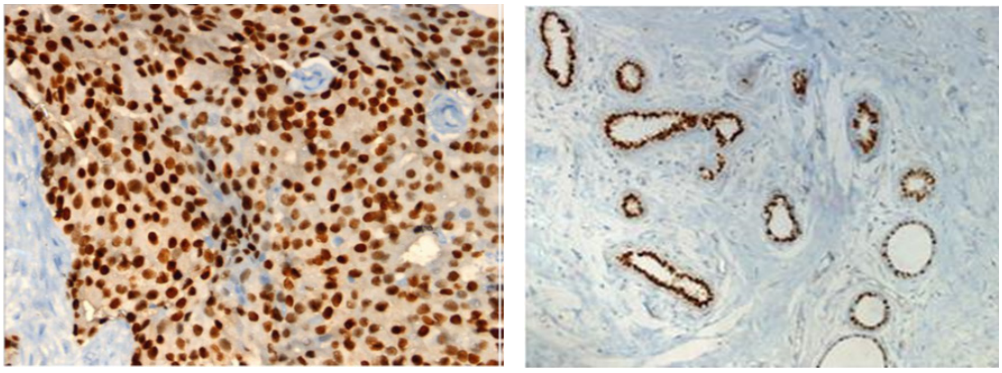
Table 3. Precision, Sensitivity and Specificity Values with different feature set for ER Images.

Class	5 features			4 features			2 and 3 features		
	Precision	Sensitivity	Specificity	Precision	Sensitivity	Specificity	Precision	Sensitivity	Specificity
1	33.33	50	96.88	66.7	66.7	98.41	100	100	100
2	87.88	82.86	87.1	96.97	84.21	96.43	100	91.67	100
3	74.07	80	82.93	81.48	95.65	88.37	89.29	100	92.85
4	66.7	50	98.39	66.7	100	98.44	100	100	100



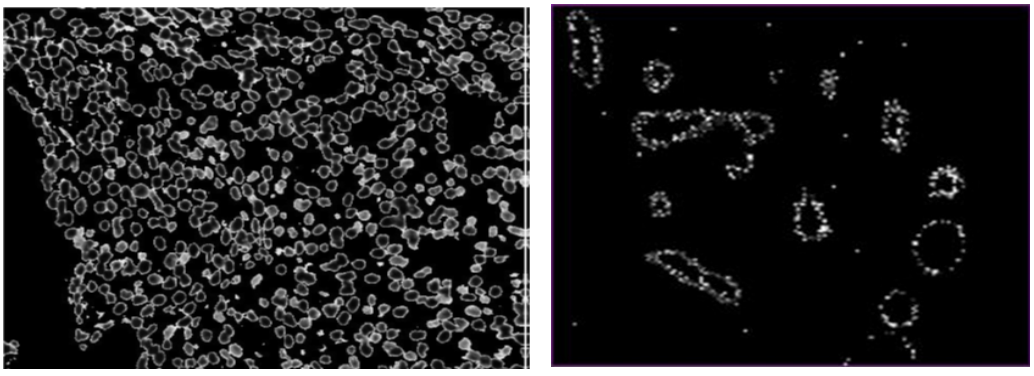
(a) **(b)**

Figure 2. **(a)** ER Image. **(b)** PR Image.



(a) **(b)**

Figure 3. **(a)** Original ER Image. **(b)** Original PR Image.



(a) **(b)**

Figure 4. **(a)** Segmented ER Image. **(b)** Segmented PR Image.

achieved an accuracy of 95.52 percentages. Total 66 test images and 20 train images are used for this purpose.

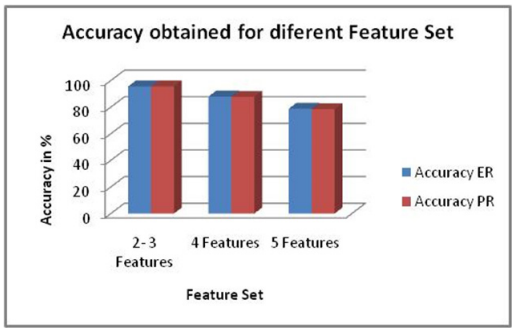


Figure 5. Accuracy Comparison for ER/PR Scoring with Different Feature Sets.

Figure 5 shows the accuracy values obtained for the tested images with different feature sets. Similar tests are done with PR images also. And similar results are obtained with PR images also. Total 65 images are used in PR cases and in Class 1, 8 images are used and in Class 2, 25 images are used. Similarly for Class 3 and Class 4, 23 and 9 images are used. With this set of test images we got an accuracy of 95.58%.

In the PR cases similar experiments are done with the feature sets and obtained the optimal set as energy and entropy feature set. In manual case also ER and PR images are treated as same type of images even though different stains are used for identifying these receptors. Table 3 shows the Precision, Sensitivity and Specificity Values obtained with different feature set for ER Images. Here while using two features the precision, sensitivity and specificity values are high compared to the other two feature sets used.

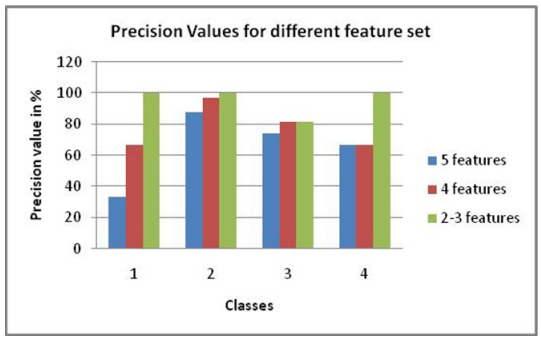


Figure 6. Precision Value Comparison for Different Feature Sets With ER.

The precision value chart for the ER image is shown in Figure 6. This chart exactly shows the reproducibility of the classes. With 2 features class 1 and class 2 and class 4 classifies with 100% classification and only few misclassifications in class 3 images reduces its reproducibility.

Table 4. shows the actual classification matrix of PR images with 2 features. From the total 25 class 2 images, one image is misclassified as class1 image and similarly in class 3 images, from 23 total images, two are misclassified as class 2 images.

Table 4. Obtained classification result for the PR images with two Features

	Class 1	Actual class			
		2	3	4	
Output Class	1	8	1	0	0
	2	0	24	2	0
	3	0	0	21	0
	4	0	0	0	9

4. Conclusion

In this work we are computing the scores for ER/PR images. One important aspect in the role of pathology is the evaluation of the breast cancer subtypes, with different biomarkers and the staining method. Slight changes in the staining will lead difference in scoring. This leads to over, or under dosage to the patients and decrease the prognosis factor. Specifically the accurate assessment of the Estrogen Receptor (ER), Progesterone Receptor (PR) receptors can increase the prognosis factor of nearly 60 percentage of patient's. This work is a good step towards automating the subjective HSCORE procedure, an objective one through a software implementation.

5. References

1. Smitha P, Shaji L, Mini MG. A Review of medical image classification techniques. Proceedings on International Conference on VLSI, Communications and Instrumentation (ICVCI)2011; UK: International Journal of Computer Applications. 2011; (11):34–8.
2. Smitha P, Paul V, Marichamy P, Sujathan K, Jalal BN. Scoring of HER2 expression from immunohistochemistry images of breast cytology. International Journal of Applied Engineering Research. 2015; 10(70):146–53.
3. Onitilo AA, Engel JM, Greenlee RT, Mukesh BN. Breast cancer subtypes based on ER/PR and Her2 expression:

Comparison of clinicopathologic features and survival. Clinical Medicine Research. 2009; 7(1):4–13.
4. Bhagat VM, Jha BM, Patel PR. Correlation of hormonal receptor and Her-2/neu expression in breast cancer study at tertiary care hospital in South Gujarat. National Journal of Medical Research. 2012 Jul; 2(3):295–8.
5. Ali EM, Ahmed RH, Ali AMA. Correlation of breast cancer subtypes based on ER, PR and HER2 expression with axillary lymph node status. Cancer and Oncology Research. 2014; 2(4):51–7.
6. Asim Qureshi SP. Allred scoring for ER reporting and it's impact in clearly distinguishing ER negative from ER positive breast cancers. Pakistan Medical Association. 2010; 60(5):350–3.
7. Singh S. A biopsy analysis system for cancer diagnosis and prognosis. International Journal of Computer Applications. 2012 Jun; 47(1):34–41.
8. Khan AM, Mohammed AF, Al-Hajri SA, Al Shamari HM, Qidwai U, Mujeeb I, Rajpoot NM. A novel system for scoring of hormone receptors in breast cancer histopathology slides. 2014 Middle East Conference on Biomedical Engineering (MECBME); Doha; UK: IEEE. p. 155–8.
9. Kecheril SS, Venkataraman D, Suganthi J, Sujathan K. Segmentation of lung glandular cells using multiple color spaces. International Journal of Computer Science, Engineering and Applications (IJCSEA). 2012 Jun; 2(3):147–58.
10. Tuominen VJ, Ruotoistenmki S, Viitanen A, Jumppanen M, Isola J. Immunoratio: A publicly available web application for quantitative image analysis of estrogen receptor (ER), progesterone receptor (PR), and ki-67. Breast Cancer Research. 2010; 12(4):2–12.
11. Prasad K, Zimmermann B, Prabhu G, Pai M. Datamining approach for automation of diagnosis of breast cancer in immunohistochemically stained tissue microarray images. The Open Medical Informatics Journal. 2010; 4:86–93.
12. Raad A, Kalakech A, Ayache M. Breast cancer classification using neural network approach: MLP and RBF. The 13th International Arab Conference on Information Technology (ACIT). 2012; 7(8):9.
13. Avoni M. Image analysis methods for determining her2/neu status of breast cancers. Nordic MATLAB User Conference Proceedings; Denmark. 2008.
14. Niwas SI, Palanisamy P, Sujathan K, Bengtsson E. Analysis of nuclei textures of fine needle aspirated cytology images for breast cancer diagnosis using complex Daubechies wavelets. Journal on Signal Processing. 2013 Oct; 93(10): 2828–37.
15. Lezoray O, Elmoataz A, Cardot H, Gougeon G, Lecluse M, et al. Segmentation of cytological image using color and mathematical morphology. European conference on Stereology; Amsterdam, Netherlands. 1998. p. 1–10.
16. Ponraj DN, Jenifer ME, Poongodi P, Manoharan SJ. A survey on the preprocessing techniques of mammogram for the detection of breast cancer. Journal of Emerging Trends in Computing and Information Sciences. 2011 Dec; 2(12): 656–64.
17. Gavrilovic M, Azar JC, Lindblad J, Wahlby C, Bengtsson E, Busch C, Carlbom IB. Blind color decomposition of histological images. IEEE Transactions on Medical Imaging. 2013 Jun; 32(6):983–94.
18. Kassner A, Thornhill RE. Texture analysis: A review of neurologic MR imaging applications. A Journal of Neuro Radiology. 2010 May; 31:809–16.
19. Materka, Strzelecki M. Texture analysis methods – A review. Brussels: Institute of Electronics, Technical University of Lodz, COST B11 report; 1998.
20. Jiji WG, Ganesan L, Ganesh SS. Unsupervised texture classification. Journal of Theoretical and Applied Information Technology. 2009 Apr; 5(4):373–81.
21. Khan AM, Rajpoot N, Treanor D, Magee D. A nonlinear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution. IEEE Transactions on Biomedical Engineering. 2014 Jun; 61(6):7291–8.
22. Basavanhally A. Automated image-based detection and grading of lymphocytic infiltration in breast cancer histopathology [Thesis report]. New Jersey: Graduate School—New Brunswick Rutgers, The State University of New Jersey; 2010.
23. Singh S, Saini S, Singh M. Cancer detection using adaptive neural network. International Journal of Advancements in Research and Technology. 2012 Sep; 1(4):1–5.
24. George YM, Elbagoury BM, Zayed HH, Roushdy MI. Breast fine needle tumor classification using neural networks. IJCSI International Journal of Computer Science. 2012 Sep; 9(5):247–56.
25. Sheikhpour R, Ghassemi N, Yaghmaei P, Ardekani JM, Shiryazd M. Immunohistochemical assessment of P53 protein and its correlation with clinicopathological characteristics in breast cancer patients. Indian Journal of Science and Technology. 2014 Apr; 7(4):472–9.