Detection of Mitotic Nuclei in Breast Histopathology Images using Localized ACM and Random Kitchen Sink based Classifier

Sabeena Beevi K,[1] Madhu S. Nair,[2] and Bindu G.R [1]

***Abstract*— The exact measure of mitotic nuclei is a crucial parameter in breast cancer grading and prognosis. This can be achieved by improving the mitotic detection accuracy by careful design of segmentation and classification techniques. In this paper, segmentation of nuclei from breast histopathology images are carried out by Localized Active Contour Model (LACM) utilizing bio–inspired optimization techniques in the detection stage, in order to handle diffused intensities present along object boundaries. Further, the application of a new optimal machine learning algorithm capable of classifying strong non-linear data such as Random Kitchen Sink (RKS), shows improved classification performance. The proposed method has been tested on Mitosis detection in breast cancer histological images (MITOS) dataset provided for MITOS-ATYPIA CONTEST 2014. The proposed framework achieved 95% recall, 98% precision and 96% F–score.**

I. INTRODUCTION

Mitotic rate is an important factor in histological grading of breast tumor cells by Nottingham international grading system [1]. Quantitative measurement of mitotic count requires expert assessment of different locales of various biopsy slides and frequently need to estimate number of malignant cells or nuclei amid a group of 5000 to 6000 cells. With Digital Whole Slide Imagining (WSI) and use of computer-aided diagnosis for digitized histopathology, the objectivity and reproducibility of analysis results are greatly improved. In breast histopathology images nuclei are seen with a background of fibro epithelial stroma (there are muscular cells, epithelial cells, adipose tissues and vacuoles) [2]. Due to difference in shape, size and density of nuclei undergoing mitosis, the detection process became one of the critical stages in automatic evaluation. In addition, presence of artefacts like irregular illumination level, cells with partial nuclei, feeble edges and closeness in gray level of different soft tissues makes it more challenging. In this frame work, the accuracy of mitotic detection is improved with careful design of algorithms in every stage of the processing pipeline. In order to segment the nuclei properly from the back ground, the mask image given to the segmentation phase should perfectly delineate the nuclei regions. Hence in this paper, Localized Active Contour Model (LACM) [3] combined with an automatic technique for optimal curve placement with Krill Herd Optimization algorithm (KHO) [4] is proposed so as to provide the inherent advantages of both techniques. Finally, the segmented nuclei are categorised using a relatively new optimal machine learning algorithm namely Random Kitchen Sink (RKS), [5-7], which is capable of classifying strong non-linear data. The frame work is simple and shows excellent classification performance compared to existing techniques.

The rest of the paper is structured in such a way that in section II different techniques recently reported in literature relevant to this research work are reviewed. The proposed method is described in section III. Section IV presents experimental results and discussion of the proposed technique. Conclusions and future work are given in section V.

II. REVIEW OF RELATED LITERATURE

Sertel et al. [8] proposed probability based likelihood functions along with binary thresholding to identify mitotic nuclei in neuroblastoma images. To detect cell nuclei, Hough transform is used by Cosatto et al. [9]. But it suits only for circular nuclei and require excessive computation. Lu and Mandal [10] detected the nuclei by linear discriminant analysis on spectral band images. Further, to perform classification of mitotic cell candidates, an imbalanced classification framework is applied. H.Irshad et al. [11] detected nuclei with blue ratio images and RF classifier. Mitotic nuclei detection is carried out by isolating tumor region from non-tumor areas by gamma-gaussian mixture model [12]. It requires a context aware post processing step in the classification stage. Hysteresis thresholding technique together with morphological processing is used for mitotic detection in breast histology [13] images. The algorithm detects almost all significant nuclei, but allows too many false positives [14]. Fuzzy C-means clustering algorithm along with ultra-erosion operation in L*a*b color space is proposed by Anari et al. [15] for detection of mitosis index in Immuno Histo Chemical (IHC) images. But the optimal solution is seldom reached and is time consuming for large histopathology images. Eventhough a number of works have been published, there are still progress to be made to achieve clinically acceptable results.

*This work is supported by Ministry of Human Resource Development, Govt of India.

*The experimental procedures involving human subjects described in this paper were approved by the Institutional Review Board.

[1]Sabeena Beevi K is with Electrical Engineering Department, College of Engineering Trivandrum, Thiruvananthapuram-695016, Kerala, India and also with Electrical & Electronics Department, T. K. M College of Engineering, Kollam, Kerala, India. sabeena3000@gmail.com

[2]Madhu S. Nair is with Department of Computer Science, University of Kerala, Kariavattom, Thiruvananthapuram-695581, Kerala, India. madhu_s_nair2001@yahoo.com

[1]Bindu G. R. is with Electrical Engineering Department, College of Engineering Trivandrum, Thiruvananthapuram-695016, Kerala, India. bgr100@gmail.com

978-1-4577-0220-4/16/$31.00 ©2016 IEEE

III. Proposed Methodology

The proposed technique comprises of mainly three phases such as stain normalization and color component selection, candidate detection and segmentation together with feature selection and classification.

1) Stain Normalization and Color Component Selection: Staining process vary widely due to different stain manufacturers, different staining practices, and different storage times. Hence images may have different dynamic range and required to normalise for further analysis to avoid complications in quantitative assessment of their processed results. Initially images with stain variations are normalized with image specific color deconvolution based technique proposed by Adnan et al. [16]. A reference image with ideal staining conditions is manually selected for the stain normalization process. The technique automatically regulates the RGB color distribution $\Re$ of the source image to $\hat{\Re}$, corresponding to actual colors of the stains in the reference image. A good contrast exist between cell nuclei and other cell structures in R component compared with corresponding G and B components of the stain normalized RGB image. Hence R component is selected for further processing. In order to detect feebly stained nuclei and edges, R component image is processed with adaptive Wiener filter. From the filtered image (I_W), nuclei are detected by applying entropy based optimal detection.

2) Candidate Detection and Segmentation: As the second step of the proposed algorithm nuclei detection is carried out by maximising Kapur's entropy criteria [17] with KHO algorithm. Optimal threshold values of different regions can be obtained by imitating the motion characteristics of the Krill individuals through the following steps:

Step I. Parameter Initialization.

a). Initialize the number of threshold values which is equal to the number of Krill individuals in the KHO algorithm.

b). Specify the lower and upper boundaries of the threshold values.

c). Initialize KHO motion parameters and maximum number of iterations.

Step II. Position Calculation.

Randomly create the initial population of thresholds in the search space between 0-255 which is equal to the initial position of the Krill individuals.

Step III. Objective Function Evaluation

Calculate the fitness of current threshold position using Kapur's objective function.

Step IV. For each threshold value calculate

a). Motion induced by the presence of other individuals (N_k)

b). Foraging activity (F_k) and

c). Random diffusion (D_k)

The locus of a Krill individual through the interval $[t, t+\Delta t]$ is decided as follows:

$$X_i(t+\Delta t)=X_i(t)+\Delta t\frac{dX_i}{dt} \tag{1}$$

where $\frac{dX}{dt}$ is given by Eqn.(2)

$$\frac{dX}{dt}=N_k+F_k+D_k \tag{2}$$

Step V. Update threshold values using the motion vector obtained from Step IV.

Step VI. If maximum number of iterations is reached, select the threshold values corresponding to the best thresholds. Otherwise, repeat the process from step III.

The actual detection of nuclei is performed by fixing the centroid of the detected objects. While segmenting pathologic cell nuclei with diffused boundary in medical images, region based ACM also results in inaccurate segmentation. Hence, in this paper a new class of active contour energies, Localized Active Contour Model (LACM) proposed by Lankton et al. [3], is considered which utilizes local information of boundary points for accurate segmentation of nuclei.

In Localized ACM approach, the foreground and background regions are assumed to have distinct mean intensities and optimal curve evolution is obtained when the mean intensity difference is maximized along the contour. A characteristic function, B(x, y) is used to identify local regions in terms of a radius parameter r.

$$B(x,y)=\begin{cases}1, & \|x\text{-}y\| <= r\\ 0, & otherwise\end{cases} \tag{3}$$

Local energies are computed by splitting the local neighborhoods $N_{(x,y)}$ into local interior and local exterior by the evolving contour. The energy functional $F(C)$ can be defined as

$$F(C)=N_{(x,y)}E_{ext}(C)+\lambda\Psi \tag{4}$$

Based on a region-based energy, E_{YR}, proposed by Yezzi et al. [18] as in Eqn.(5), F can be written as given in Eqn.(6)

$$E_{YR}=\int \text{-}[I_{int}\text{-}I_{ext}]^2 \tag{5}$$

$$F(C)=\text{-}1/2[(I_{int}\text{-}I_{ext})^2]+\lambda\Psi \tag{6}$$

I_{int} and I_{ext} are average intensity value inside and outside the object region, respectively. λ is a positive fixed parameter which help to smoothen the growing contour. Since these energies utilizes local information and also integrates the use of region-based techniques, they are capable of handling diffused intensity variations.

3) Feature Selection and Classification: In the beginning of the feature computation stage, four intensity based features and three shape based features are computed along with Haralick texture features [19]. The intensity based features include mean, variance, skewness and kurtosis. Area, perimeter and solidity are the shape based features considered. All fourteen texture features are calculated for every candidate at first. Out of the fourteen features, the visual texture characteristic features include Angular Second Moment (ASM), Contrast (CON), and Correlation (COR). Variance (VAR), Inverse Difference Moment (IDM), Sum Average (SA), Sum Variance (SV) and Difference Variance (DV)

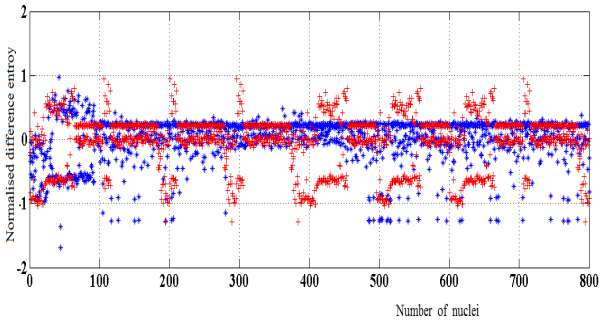

Fig. 1. Scatter plot of the normalised feature vector

are based on statistics. Information theory based features include Entropy (ENT), Sum Entropy (SENT) and Difference Entropy (DENT). Finally, features based on information measures of correlation such as Information Measures of Correlation (IMC1, IMC2) and Maximal Correlation Coefficient (MCC) are considered. All these features are calculated along the four directions such as 0^o, 45^o, 90^o and 135^o. After normalizing in the four directions, mean and range are computed. Throughout the investigation it was found that maximal correlation coefficient devours considerable time of computation. Besides, for many nuclei the range of the sum average (SA) of variance is observed as zero. Hence these two features are not considered in subsequent stages of analysis. For every candidate object twenty four texture features are computed along with intensity and shape based features, resulting in a total of 31 features. Before the classification stage feature normalization followed by sequential feature selection is carried out. The selected features include all the intensity and shape based features, but mean and range of ASM, COR,VAR,IDM and DENT are found to be redundant and discarded. The complexity of the data is displayed in Fig.1 which shows the scatter plot of the feature set for 800 mitotic nuclei and non mitotic nuclei in red and blue color, respectively.

A classifier evaluation is carried out with different classifiers such as Neural Network (NN), non-linear SVM (NLSVM), ensemble classifiers such as Ada Boost(AB), Robust Boost(RB) and Random Forest (RF) [20]. Despite the inherent advantages of less error rate for learning tasks and ability to handle large data set, the RF model is not easily interpretable and require fine tuning with the data to be used. Hence, in order to differentiate mitotic and non mitotic nuclei from the highly non linear feature vectors a new machine learning algorithm namely Random Kitchen Sink (RKS) [5] is utilised in this work.

A. Random Kitchen Sink algorithm

In kernel methods, data points are mapped to higher dimensional space and has pair-wise similarity between mapped points in terms of inner products. The inner product between lifted datapoints of a positive definite function $p(x, y)$ with $x, y \in R^n$ and a lifting Θ can be computed as in Eqn.(7)

$$\langle \Theta(x), \Theta(y) \rangle = p(x, y). \tag{7}$$

Without any implicit lifting through kernel functions, data can be mapped to a low-dimensional inner product space using a randomized feature map z : $R^n \longrightarrow R^N$ so that

$$p(x, y) = \langle \Theta(x), \Theta(y) \rangle = z(x)', z(y). \tag{8}$$

Hence the input data can be transformed without kernel's lifting Θ, to a hyperplane w, and fast linear learning methods can be computed by evaluating, $f(x) = w' * z(x)$. The Random Features accelerates the supervised learning algorithms so that they can be applied on big data sets. The data is passed through random features, and a linear machine is trained on the featurized data. In this work RKS utilise dimensionality, *D =500* for creating random features and regularization parameter τ*=20.* Entire program of the proposed technique is consolidated in the form of a pseudocode and displayed in Algorithm 1.

Algorithm 1 Detection of Mitotic Nuclei

procedure MITOSIS DETECTION
  for each H & E stained image $i \in N$ **do**
  Stain Normalization & Color component selection
  $\hat{\Re}(p) = S^{-1}(-log_{10}(\Re(p)))$ /* S is the saturation value of stain matrix, $\hat{\Re}$ color distribution in the new colorspace*/
  $(I_R, I_G, I_B) \leftarrow I_{RGB}$ // color component selection
  $I_W \leftarrow I_R$ // Wiener filtering
  Optimal candidate detection & Segmentation by Localised ACM
  $I_{B(x,y)} \leftarrow$ KHO (I_W,n) // optimal mask by KHO
  $I_{LACM} \leftarrow LACM(I_{RGB}, I_{B(x,y)})$ /* Segmentation by Localised ACM*/
  Feature Selection and Classification
  $FEAT \leftarrow Featgen(I_{LACM})$ /* computing feature vectors for nuclei */
  $[Trainset, Testset] \leftarrow FEAT$ /* compute train and test set by sequential feature selection and selective sampling */
  $ClassifierTraining \leftarrow Train(Trainset)$ /* Training of classifier */
  $[Mitotic, NonMitotic] \leftarrow RF(Testset)$
  $[Mitotic, NonMitotic] \leftarrow Rnd\_KS(Testset)$
  end for
end procedure

IV. Experimental Results

The proposed frame work is evaluated on the scanned images of two sets of stained breast biopsy slides from MITOS dataset [21]. The data set is composed of 86-96 high power field (HPF) images of breast tissue scanned at 40X magnification using two different scanners, Aperio (AP) and Hamamatsu (HM), with a resolution of 0.23-0.24 μm. All images in the data set are of 1376 $\times$ 1539 $\times$ 3 size. Out of the 305 mitotic nuclei from the 182 HPFs considered,

the training set is formed with 200 mitosis and remaining 105 mitosis is used for the testing set. From the 20450 non mitotic nuclei present, 14315 are used in the training set and 6135 used for testing. In order to reduce the skewness of data, selective sampling of non mitotic data and random up sampling of mitotic data are carried out and a new training set is prepared with 600 mitosis and 3075 non mitosis. The proposed technique is developed with Matlab 14 (a) on Windows 8 and GURLS toolbox [22] with Intel(R) Core (TM) i7 CPU, 3.4 GHz, 4 GB RAM.

TABLE I
CLASSIFIER PERFORMANCE EVALUATION (Mitosis-105, Non-Mitosis-249)

Classifier	TP	FP	Sensitivity	Precision	F-Score
AdaboostM1	90	50	85.71	64.28	73.4
RobustBoost	97	59	92.38	62.17	74.33
SVM	94	46	89.5	67.14	76.73
RandomForest	98	32	93.33	75.38	83.4
RandomKS	98	1	93.33	98.98	96.07

Two sets of test data are prepared with and without skewness from the initial test data and classifier predictions are observed. The performance of different classifiers are quantitatively evaluated with respect to sensitivity, specificity, accuracy, precision and F-score measures. The first test set contains 105 mitosis and 249 non mitosis nuclei. The second test is formed with 100 mitosis and 1000 non mitosis nuclei. Table I & II presents the performance of different classifiers obtained by the proposed method over the first and second test set, respectively. For creating RF model repeated 10 fold cross validation is adopted with *repetitions=10*. The tuning process selected 4 random features at each split. NLSVM is used with rbf kernal of degree 3 and boosting algorithms with 500 trees. RKS has got highest F-score of more than 96% in both test sets. The Receiver Operating Characteristic (ROC) Curve of the proposed framework on the second data set is presented in Fig.3. The area under curve (AUC) is about 96.47%. Fig. 2 displays visual results of the proposed technique in a H & E stained image. Fig.2 (a) is the original image. Fig.2 (b) and (c) displays the mask image derived by the KHO based detection stage and segmented image obtained after Localised ACM based segmentation process, respectively. Fig.2 (d) shows 4 rightly detected mitotic nuclei (TP) in green circles and one missed mitotic nuclei (FN) in red circle. The missed mitosis is not segmented out along with other nuclei in the segmentation stage, which is shown in blue circle in Fig.2 (c).

TABLE II
CLASSIFIER PERFORMANCE EVALUATION (Mitosis-100, Non-Mitosis-1000)

Classifier	TP	FP	Sensitivity	Precision	F-Score
AdaboostM1	85	85	85	50	62.96
SVM	91	75	91	54.81	68.42
RobustBoost	91	62	91	59.47	71.93
RandomForest	93	33	93	73.8	81.05
RandomKS	95	2	95	97.93	96.44

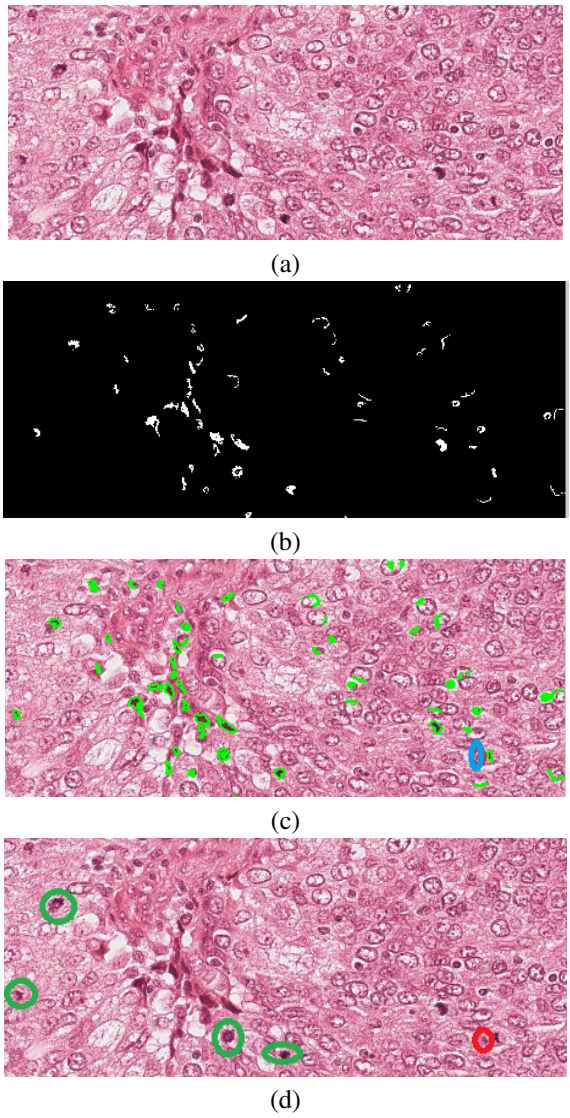

Fig. 2. Visual result of the proposed framework. (a) Original image (b) mask image derived by KHO optimized candidate detection stage. (c) segmented image obtained by the LACM. (d) Predictions of RKS classifier, green circular markings represent TP and red circle represents FN

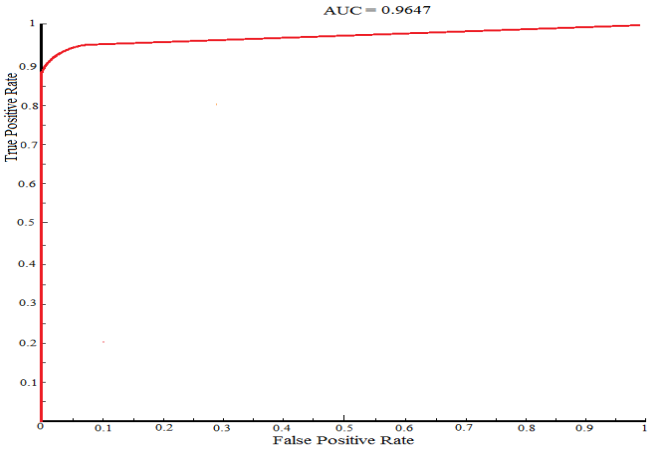

Fig. 3. ROC Curve of the proposed framework

For the first test data a ratio of 30:70 is selected for mitosis and non mitosis. RF and RKS classifier shows equal sensitivity, but RF has got a low precision of 75.38% compared with RKS which has got good value of 98.98%. From the available literature for mitosis detection in breast histopathology images, best F-score of 0.782 and 0.718 are reported with deep max-pooling convolutional neural networks and random forest classifier [11], respectively. With proposed method, RF classifier predictions gives an F-score of 83.4% for the test data 1 and 81.05% for data set 2, a fairly good value, which accounts for the improved performance of the segmentation phase. In the case of RKS, F-score of 96.07% is obtained for the first data set and 96.44% for the second data set. Increased number of false positives is a major issue in traditional classifiers. By taking selective samples of clusters from majority class, false positives can be considerably reduced. Table II displays the performance of classifiers on the second test data with 100 mitotic and 1000 non mitotic nuclei. RKS shows excellent reduction of false positives compared to all other classifiers with good precision and recall even with a skewed test data.

V. Conclusion

An integrated frame work is proposed to carry out segmentation and classification of mitotic nuclei in breast histopathology images subject to Localized ACM along with RKS classifier. Since mitosis are normally rare and seen well-separated, they are very hard to distinguish from non-mitotic nuclei. To handle this difficulty an optimal way of segmenting exact nuclei boundary is adopted by the use of KHO and LACM. More over reduced number of features is proved to be sufficient for discriminating mitotic nuclei from non mitotic nuclei, which makes this method less complex. RKS based classifier shows excellent mitosis detection rate together with perfect reduction of false positives. Results prove that the proposed methodology succeeds in improving sensititvty, precision and F–score. As a future work, the deep learning techniques can be incorporated to explore the possibility of further performance enhancement of the proposed technique.

ACKNOWLEDGMENT

The authors acknowledge the timely guidance and support rendered for this work by pathologist Dr.Abitha and team.

References

[1] Fattaneh A. Tavassoli and Peter Devilee, *Tumours of the Breast and Female Genital Organs. World Health Organization Classification of Tumours.* Pathology & Genetics. International Agency for Research on Cancer, September 2003.

[2] W. W. Ma and A. A. Adjei, “Novel agents on the horizon for cancer therapy”, A Cancer Journal for Clinicians, vol.14, No. 59, pp. 111-137, 2009.

[3] Lankton S, Tannenbaum A, “Localizing region-based active contours”. *IEEE Transactions on Image Processing*, vol. 17, no.11, pp: 2029–39, 2008.

[4] A.H.Gandomi, A.H.Alavi, “Krill herd: A new bio-inspired optimization algorithm”. *Communications in Nonlinear Science and Numerical Simulation*,vol. 17, pp. 4831–4845, May.2012.

[5] Rahimi, A. and Recht, B. “Weighted sums of random kitchen sinks: Replacing minimization with randomization in learning”, *In Advances in Neural Information Processing Systems*,volume 21, pp. 1313–1320, 2008.

[6] Kumar SS, Mohan N, Poornachandran P, Soman KP, “Condition Monitoring in Roller Bearings using Cyclostationary Features”, In Proceedings of the Third International Symposium on Women in Computing and Informatics 2015 Aug 10, pp. 690-697, ACM.

[7] Karthieswari.R, Sreethivya.M, DevaRajendraKumar.D and Soman.K.P.“Thamil Character Classification Using Rndom Kitchen Sink Algorithm”, ICONIAAC ’14, October 10 - 11, 2014, Amritapuri, India

[8] O. Sertel, U. V. Catalyurek, H. Shimada, and M. N. Gurcan, “Computeraided prognosis of neuroblastoma: Detection of mitosis and karyorrhexis cells in digitized histological images,”*in Proc. IEEE 31st Annu. Int. Conf.Eng. Med. Biol. Soc.*, Sep. 2009, pp. 1433–1436.

[9] E.Cosatto, M. Miller, “Grading nuclear pleomorphism on histological micrographs”*in Proc.19th international conference on Pattern recognition*, pp. 1-4, December 2008.

[10] Cheng Lu, Mrinal Mandal, “Toward automatic mitotic cell detection and segmentation in Multy spectral Histopathological images”, *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no.2, pp. 850–870, March 2014.

[11] H. Irshad,“Automated mitosis detection in histopathology usingmorphological and multi-channel statistics features,”*J. Pathol. Inform.*, vol. 4, pp. 10–15, May 2013.

[12] M. Khan, N. M. Rajpooth, “A Gamma- Gaussian mixture model for detection on mitotic cells in breast histology images”, Proceedings of 21st International Conference on Pattern Recognition, October 2012.

[13] Beevi.K Sabeena, Bindu G.R., Nair.S Madhu.,“Automatic Segmentation and Classification of Mitotic Cell Nuclei in Histopathology Images Based on Active Contour Model”*in Proceedings of IEEE International Conference on Contemporary Computing (IC3I)*, pp. 740–744, 2014.

[14] K. Sabeena Beevi, G.R. Bindu, “Analysis of Nuclei Detection with Stain Normalization in Histopathology Images ”*Indian Journal of Science and Technology* , Vol. 8, pp. 234–238, September 2015.

[15] V. Anari, P. Mahzouni, and R. Amirfattahi, “Computer-aided detection of proliferative cells and mitosis index in immunohistichemically images of meningioma,”in MVIP, 2010.

[16] A. Khan, Nasir Rajpoot, “A Nonlinear Mapping Approach to Stain Normalization in Digital Histopathology Images Using Image-Specific Color Deconvolution” *IEEE Transactions on Biomedical Engineering*, VOL. 61, NO. 6, pp.1729–1738, JUNE 2014.

[17] Kapur, J. N., Sahoo, P. K., & Wong, A. K. C. “A new method for gray-level picture thresholding using the entropy of the histogram”. *Computer Vision, Graphics and Image Processing*, pp. 273–285, 1985.

[18] J. A. Yezzi, A. Tsai, and A.Willsky, “A fully global approach to image segmentation via coupled curve evolution equations,”J. Vis. Comm. Image Rep., vol. 13, no. 1, pp. 195–216, Mar. 2002.

[19] Robert.M.Haralick, K. Shanmugam and Dinsten,“Textural Features for image classification”, IEEE Transactions on Systems, Man and Cybernetics, no. 6, pp. 610-621, November 1973.

[20] Machine Learning with R, Packt Publishing Ltd. 2nd edition

[21] (2014). MITOS, ICPR 2014 Contest, IPAL UMI CNRS Lab Std., [Online]. Available: http://ipal.cnrs.fr/ICPR2014.

[22] “GURLS: a Least Squares Library for Supervised Learning”, Andrea Tacchetti,Pavan K.Mallapragada,Matteo Santoro,Lorenzo Rosasco.