Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 167 (2020) 666–676

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Computational Intelligence and Data Science (ICCIDS 2019)

Performance Analysis of Segmentation Algorithms for the Detection of Breast Cancer

M.A. Aswathy[a], M. Jagannath[b,*]

[a]*Research Scholar, Vellore Institute of Technology (VIT), Chennai, Tamil Nadu 600127, India*
[b]*Associate Professor, Vellore Institute of Technology (VIT), Chennai, Tamil Nadu 600127, India*

Abstract

One of the main reasons for cancer death globally is breast cancer. The detection of breast cancer from Hematoxylin (H) and & Eosin (E) stained pathology image is significant and digital pathologists are struggling to get accurate final decision. The objective of this study was to carry out the performance analysis of segmentation algorithms for the detection of breast cancer. This study proposed a method for computer assisted diagnosis and classification of breast cancer from microscopic slide images by means of biologically explicable features. The suggested methodology consists of different stages involving image enhancement, nuclei segmentation, extraction of features, and to end with the classification. A comparative analysis was done on various segmentation algorithms and finally, a suitable and effective method was utilized in the proposed approach. The contrast-limited adaptive histogram equalization is performed on the input images for contrast enhancement. Similarly, *k*-means clustering segmentation algorithm is used for segmenting the nuclei in the proposed work because it outperforms the other commonly used algorithms during comparative analysis. Gray level texture features were extracted in the feature extraction step. Finally, support vector machine classifier was employed for classification of breast cancer histopathology images into benign and malignant classes because it is the most efficient classifier comparing to the other usual classifiers for this purpose. The performance analysis of the suggested design was estimated using familiar parameters such as accuracy, sensitivity, specificity, etc. and achieved an accuracy of 91.1 %.

Peer-review under responsibility of the scientific committee of the International Conference on Computational Intelligence and Data Science (ICCIDS 2019).

Keywords: Breast cancer; Histopathology; *k*-means clustering; Computer assisted detection; Digital pathology.

*Corresponding author. Tel.: +91-9884386262; Fax: 044-3993-2555.
E-mail address: jagan.faith@gmail.com

1. Introduction

Cancer is a chronic disease that became apparently increased affecting human health. The World Cancer Research Fund International Statistics states that breast cancer is the utmost prevalent cancer in women population

Peer-review under responsibility of the scientific committee of the International Conference on Computational Intelligence and Data Science (ICCIDS 2019).
10.1016/j.procs.2020.03.333

worldwide having 2 million new cases in 2018 [1]. In India, about 12% of new cancer cases and 25% of all types of cancers among women are represented by the breast cancer [2]. The number of women-death due to cancer has been increased and breast cancer has ranked five. Furthermore, as age decreases, there is an increase in the incidence rate.

Detection of breast cancer has constantly been a key problem for the clinicians and physicians for disease analysis and proposing the right therapy. The physical diagnosis of breast cancer from histopathology images is highly subjective and may differ from person to person subject to their expertise level [3]. Also, the manual identification depends on other factors like precise environment and lighting arrangement. The computer assisted diagnosis systems helps in lessening the aforementioned problems and offers improved outcomes in detecting the breast cancer from histopathology images [4]. Figure 1 shows the main differences in cell structure of a normal and abnormal cell. Abnormal cells are growing in an uncontrolled manner and keeps on dividing the cells where as normal cells have a precise shape and size. Abnormal cells will be morphologically different from normal cells by means of abnormal number of chromosomes and nucleus size. As in Fig. 1(e), the abnormal cells do not have any precise boundary like normal cells.

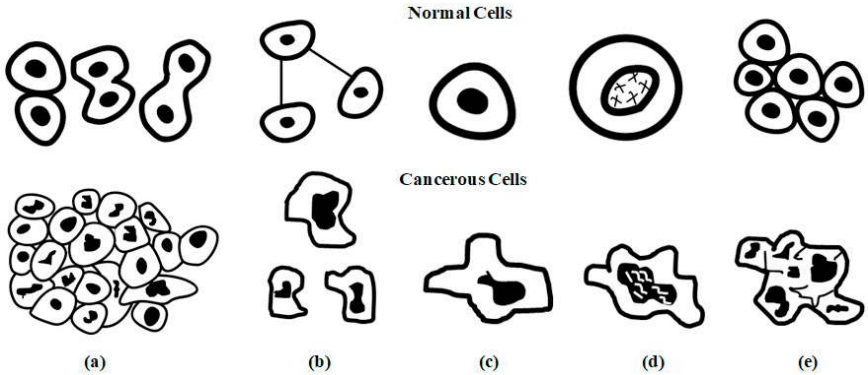

Fig. 1. (a) Cell division and growth; (b) Variation in cell size; (c) Morphology; (d) Arrangement of chromosomes; and (e) Cell boundary.

Women who are addicted to tobacco and liquor are highly prone to breast cancer. Lack of physical activities, breast feeding and reduction in the dietary conducts are some other reasons for getting breast cancer [5]. The combined recent advancements in the field of medicine and engineering give a way for curing this cancer. Early detection is the only answer for curing the breast cancer. The stage of malignancy determines the choice of treatment of breast cancer. Initially, medical experts will advise screening programs like Ultrasound, Computer Tomography (CT), X-ray, Positron Emission Tomography (PET), Mammogram and Magnetic Resonance Imaging (MRI). If some suspicions have found in the patient records, a biopsy will be suggested by the doctors [6]. A biopsy is a clinical procedure in which a tissue from the breast is removed and examined under a microscope for detecting suspicious cells. Thus, analysis of this biopsy image is a vibrant method for detecting breast cancer [7]. Histopathology analysis is the review of signs and indications of the breast cancer from the histopathology images. A procedure called staining is performed on these images to clearly visualize the structures and nuclei in the tissue. The tissue undergoes staining process using various dyes. The pathologists were using Hematoxylin & Eosin (H&E) staining from past few years. Hematoxylin is responsible for blue colour in cell (nuclei) whereas Eosin is responsible for pink colour in cytoplasm and other connective structures [8]. Abnormal and normal tissues can be differentiating from histopathology images with high resolution.

This study compared three machine learning (ML) algorithms for automated segmentation of breast cancer from histopathology images. The paper is structured as follows. Section 2 briefly overviewed existing researches in the histopathological segmentation and classification field. Section 3 presented the methodology used in the study and Section 4 described the quantitative performance analysis and Section 5 discussed the results and scope of the study. Finally, conclusion of the study is given in Section 6.

2. Related works

The associated works on histopathology image analysis are always employed in a limited classification environment. A variety of studies regarding histopathology analysis have intensive trust in efficient application of existing methods for other tasks by means of multi-resolution [9,10] and multi-spectrum [11] gliding windows. In recent studies in the literatures, in-depth window-based handling of histopathology images is employed as an alternative to manually drawing region of interests. But there will be the problem of tiling. That means tiling may cause cracks in the image which will leads to the image distortion. Balazsi et al. [12] put forward a method to reduce the problem of tiling to some point by employing hybrid feature extraction from super-pixels along with radiofrequency classifiers. Finally, the authors concluded that detection of invasive ductal carcinoma is possible with generic features but the differentiation of ductal carcinoma and invasive carcinoma is difficult. Mercan et al. [13] just presented a multi-occurrence multilabel classification framework to evaluate the vagueness concerning the communication among the slide-level remarks of pathologists and the patient's Region of Interest (ROI) extracted from their observations for poorly supervised learning on histopathology images.

Subsequently, some other frameworks focused on cell analysis for binary classification (benign-malignant classification). Kowal et al. [14] in his study employed some clustering approaches for segmenting nuclei using microscopic slide images. The study was conducted on 50 patients with around 500 images and the features such as morphological, texture and topological were extracted. Kowal et al. [14] managed to achieve accuracy between 84% and 93%. Filipczuk et al. [15] used nuclei features for classification by employing Hough transform for ROI selection and Otsu thresholding for False Positive (FP) reduction. George et al. [16] used watershed segmentation algorithm for cell detection. These two works extracted texture and structural features of nuclei for classification using various classifiers. George et al. [16] achieved an accuracy of about 97.15% using 92 images. On a contrast, Belsare et al. [17] took both nuclei linked information and tissue related info for twofold classification. This study assessed seventy complex images from a publicly available dataset with a 40× amplification level. Epithelial layer segmentation was done with the help of three-dimensional, color and texture features. The statistical texture features were used to train the classifiers and acquired an accuracy of about 85%.

Pawar and Patil [18] proposed a novel automatic system for breast cancer detection using feed forward neural network. This back propagation neural network was then compared with radial basis function (RBF) network and calculated the quantitative measures for performance evaluation. The results showed that feed forward network outperformed the RBF network in detecting breast cancer. Sameti et al. [19] presented a new feature extraction technique for detecting breast cancer from mammogram images. This method was used for early detection of breast cancer immediately after finding a suspicious region. For individual mammographic projections of the malignant breast the two specific regions were categorized. Finally, the authors introduced a discriminant analysis procedure for effectively extracting various features that are biologically important in detecting breast cancer.

3. Methodology

A histopathology image may contain very complex cell structures in a disorganized manner making the detection of breast cancer difficult. In this study, a study for automatic detection and classification is proposed using various machine learning algorithms. The different stages implicated in the proposed framework contain histopathology image enhancement, nuclei segmentation, feature extraction, and classification. The flowchart for the proposed work is depicted in Fig. 2.

The methodology section is organized as follows: Section 3.1 deals with the preprocessing of images. The contrast limited adaptive histogram equalization is used for image enhancement. Section 3.2 deals with comparison of different ML algorithms for the segmentation of histopathology images. Section 3.3 deals with the GLCM feature extraction and Section 3.4 deals with the support vector machine classifier for classification of breast cancer images.

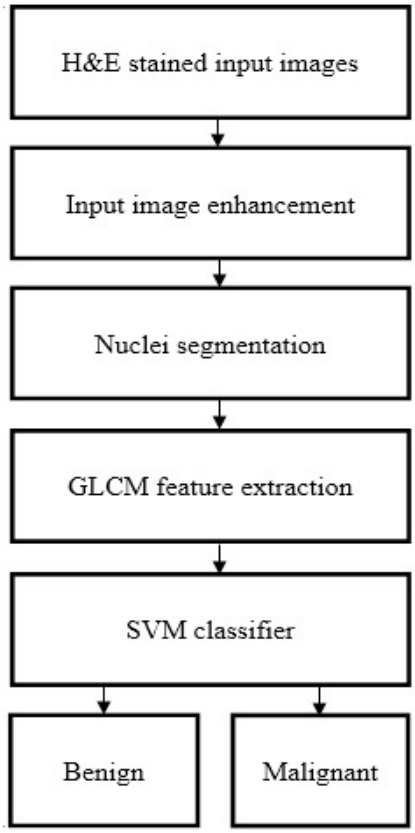

Fig. 2. Flowchart for automated breast cancer detection and classification from Haematoxylin & Eosin stained images.

3.1. Preprocessing

The input images are H&E stained histopathology images. To get better visualization, image enhancement has to be done on these images. For that, the contrast limited adaptive histogram equalization (CLAHE) [20] method is performed. The CLAHE method improves the contrast of the input image *I* by changing the hue, saturation and value. This method functions on minor patches in the input image relatively than the whole input image. So, each patch's contrast is improved separately. Finally, the histogram of the enhanced image will match with the histogram quantified by a distribution parameter. In case of induced boundaries, bilinear interpolation can be performed on neighbouring patches to combine together. Fig. 3 represents the enhanced image using CLAHE.

3.2. Segmentation

This study employed three segmentation algorithms fuzzy *c*-means algorithm, *k*-means clustering algorithm and active contour model. Then their performances were compared by using various metrics to decide the efficient one for the breast cancer detection application. Finally, the *k*-means clustering algorithm is employed in the framework because it worked well in segmenting a H&E nucleus rather than using other methods. The original H&E image and *k*-means segmented H&E image were shown in results and discussion. Fifty-seven H&E stained histopathology images are taken for the study from UCSB dataset [21] for testing and evaluation. In the same way, the corresponding ground truth images which are available in UC Santa Barbara library are also used in this study for calculating dice coefficient. The segmented results of *k*-means clustering, fuzzy *c*-means segmentation and active contour segmentation algorithms are also presented in this study.

3.2.1. k-means clustering algorithm

The *k*-means clustering is a segmenting technique that partitions a collection of data into *k* groups. The algorithm moves iteratively through the following steps. Evaluate the mean of each and every group or cluster. Then estimate the distance of each data point from each group to the cluster center. Finally, allocate each data point to the nearest cluster on the basis of distance calculated. Once the allocation is over, the cluster center again recalculated and based on that center, new distance vector is measured [22]. Imagine an $x \times y$ image that has to be partition into k clusters and c_k be the center for clusters. The algorithm of *k*-means clustering method is as follows:

Step 1: Set the number of clusters, *k* and its center.

Step 2: Calculate the distance between each data point in the image and the cluster center by using the equation (1):

$$d = \left\| p(x,y) - c_k \right\| \tag{1}$$

Step 3: Allocate all the data points into clusters based on the nearest distance, *d*.

Step 4: Once allocation is done, recalculate the new center for every cluster using the equation (2):

$$c_k = \frac{1}{k} \sum_{x \in c_k} \sum_{y \in c_k} p(x,y) \tag{2}$$

Step 5: Repeat the procedure until it converges.

Step 6: Assign data points again accordingly and reshape the image.

3.2.2. Active Contour Model (ACM)

Active contour approach uses energy forces of pixels in the region of interest (ROI) and their constraints for the separation of foreground pixels from background pixels for the image processing and evaluation. The contours in the active contour model are the boundary around the region of interest. This contour consists of pixels that endure interpolation procedure. The curve in the image can be described by various interpolation procedures like linear, spline and polynomial functions [23]. Active contour model is primarily using for segmenting ROI by defining a smooth curve around the image. The different models coming under ACM are snake model, gradient vector model, geodesic contours etc.

External and some internal energy forces are used by these contour algorithms to obtain the curvature of the image. The external force is used to position the contour correctly over the ROI and the internal forces helped to control the deformable variations. This deformable variation means defining a smooth contour around the image and this contour is defined by an energy function which has to be minimize is given by equation (3).

$$E_{acm}(S) = \int_0^1 \left\{ -\mu \left| \nabla I(C(S)) \right|^2 + w_1(S) \left| C'(S) \right|^2 + w_2(S) \left| C'(S) \right| \right\} \tag{3}$$

where E_{acm} is the energy function, S is the curvature, $C(S)$ is the contour, w_i are weights, I represents the image and μ is the mean.

3.2.3. Fuzzy c-means clustering

The k-means clustering algorithm is the base for fuzzy c-means algorithm. Due to the imprecision of the data, a fuzzy concept is introduced into *k*-means algorithm [24]. Fuzzy c-means algorithm is developed by Bezdek and Dunn in 1981. In this algorithm, the classes are characterized by centroids [25]. Each observation can be classified into a degree of membership between 0 and 1. Here fuzzy *c*-means is built on minimizing a cost function, J_{fcm} iteratively which is given by the equation (4).

$$J_{fcm} = \sum_{k=1}^{n} \sum_{i=1}^{c} (u_{ik})^q \quad d^2 \left(x_k, v_i \right) \tag{4}$$

where $X = \{x_1, x_2, ...x_n\} \subseteq R^p$ is the dataset vector population with *p*-dimension, n is the over-all sum of data in the dataset, *c* is the cluster number whose range in between 2 and *n*, For i^{th} cluster, u_{ik} is the degree of membership, *q* is the weighting function, i^{th} cluster center is given by υ_i, and Euclidean distance is given by $d^2(x_k,\upsilon_i)$.

The algorithm is as follows:

Step 1: Initialize the values of *c*, *q* and ε.

Step 2: Set the values of $\{u_{ik}\} = U$.

Step 3: Initialize the count for loop as $b = 0$.

Step 4: Calculate the cluster centers $\{\upsilon_i^{(b)}\}$ using $U^{(b)}$.

Step 5: Estimate the membership function $U^{(b+1)}$.

For $k = 1$ to n, estimate $I_k = \left\{i \middle| 1 \le i \le c, d_{ik} = \left\|x_k - \upsilon_i\right\| = 0\right\}, or\ I$. Compute new membership values for k^{th} column for the below specified conditions in equation (5):

If $I_k = \xi$ then,

$$u_{ik}^{(b+1)} = \frac{1}{\sum_{j=1}^{c}\left(\frac{d_{ik}}{d_{jk}}\right)^{\frac{2}{(q-1)}}} \tag{5}$$

else $u_{ik}^{(b+1)} = 0$ *for all* $i \notin I$ and $\sum_{i \in I_k} u_{ik}^{(b+1)} = 1$;

Step 6: If $u_{ik}^{(b+1)} = 0$ *for all* $i \notin I$, then stop; else fix $b=b+1$ and go to Step 4.

3.3. Feature extraction

In this phase, various gray level texture features are extracted using gray level co-occurrence matrix (GLCM) [26]. GLCM resembles a second order statistics in which the pixel information is collected from pixel pairs. GLCM reveals how the pixel pair intensity in an image varies. GLCM matrix is constructed at a distance of *d* separated by *1* and at various angles of degrees. GLCM features are various measures i.e. entropy, energy, contrast and correlation. These measures can be calculated by using the group of equations (6)–(9).

$$Contrast = \sum_{i,j=0}^{n-1} I_{i,j}(i-j)^2 \tag{6}$$

$$Correlation = \sum_{i,j=0}^{n-1} I_{i,j}\left[\frac{(i-m)(j-m)}{s^2}\right] \ where\ m = \sum_{i,j=0}^{n-1} i \times I_{i,j} \ and\ s^2 = \sum_{i,j=0}^{n-1} I_{i,j}(i-m)^2 \tag{7}$$

$$Energy = \sum_{i,j=0}^{n-1} {I_{i,j}}^2 \tag{8}$$

$$Homogeneity = \sum_{i,j=0}^{n-1} \frac{I_{i,j}}{1+(i-j)^2} \tag{9}$$

3.4. Support Vector Machine (SVM)Classification

Finally, support vector machine (SVM) classifier is completed for classifying the benign and malignant H&E images. Initially SVMs were established for binary classification and later expanded for multi class classification

and regression. The multi class classification can be achieved through a coupling method by using binary classifiers as pairwise models. Consider the data in training set as *X.* Then *X* can be expressed as in (10),

$$X = \{(i_1, j_1), (i_2, j_2) \ldots\ldots (i_n, j_n)\} \tag{10}$$

where i_x is a real vector, i_y denotes the class labels and comes in the range of 1 or -1. The function of the SVM classifier, *S(i)* takes the form of equation (11):

$$S(i) = w \bullet i - \lambda \tag{11}$$

w is the weight updates and λ is the bias. The classification decision is based on the two conditions shown in equation (12).

$$\begin{aligned} &w \bullet i_x - \lambda > 0 \;\; if \;\; y_x = 1 \; and \\ &w \bullet i_x - \lambda < 0 \;\; if \;\; y_x = -1 \end{aligned} \tag{12}$$

Out of fifty-seven, ten images are tested for classification and the remaining images were used for training purposes. The performance of the classifier is assessed using well known performance metrics.

4. Quantitative performance evaluation

The short description of the performance metrics used for the quantitative evaluation of segmentation algorithms is as follows [27]. The performance parameters accuracy, sensitivity, specificity, *F*-score, and balanced accuracy were calculated using equations (13)–(18).

Accuracy (*A*) can be estimated by comparing the segmented result with the ground truth image. It will give a measure of how accurately the proposed algorithm segmented an image.

$$A = \frac{t_P + t_N}{N} \times 100 \tag{13}$$

Sensitivity (α) and specificity (β) are true positive fraction and true negative fraction. That means how much percentage of positive fraction correctly segmented and how much of percentage of negative fraction correctly removed.

$$\alpha = \frac{t_P}{t_P + f_N} \tag{14}$$

$$\beta = \frac{t_N}{t_N + f_P} \tag{15}$$

The dice index (*D*) calculates the Sorensen-dice similarity coefficient between input image and the segmented image.

$$\begin{aligned} &I_{and} = Ground\ truth\ \&\ Segmented\ output \\ &I_{or} = Ground\ truth \mid Segmented\ output \\ &D = \frac{2\ Sum(I_{and})}{Sum(I_{or})} \end{aligned} \tag{16}$$

F-score or *F*-measure is determined by using precision, ρ and recall, Υ. It is the harmonic average of ρ and Υ.

$$\begin{aligned} &\rho = \frac{t_P}{t_P + f_P} \;\; \& \;\; \gamma = \frac{t_P}{t_P + f_N} \\ &F - measure = 2 \times \frac{\rho \times \gamma}{\rho + \gamma} \end{aligned} \tag{17}$$

Balanced accuracy (B_a) is a measure of balance between correctly predicting the two classes. It can be computed as:

$$B_a = \frac{\alpha + \beta}{2} \tag{18}$$

All these performance metrics are calculated for the three segmentation approaches and validated the results with previous studies.

5. Results and Discussion

This framework is implemented on MATLAB 2014b, on UCSB dataset on PC with 3.4 GHz AMD processor, 16 GB RAM, and windows 10 platform.

For the experimentation purpose, the dataset used in this work is the publicly available UCSB dataset. The dataset consists of 58 images having a resolution of 896×768. The ground truth images are also available in the UC Santa Barbara library. Figure 3 represents the input image and the CLAHE enhanced image. Compared to histogram equalization approach, CLAHE method will reduces the effect of intensity saturation. Moreover, it is useful in reducing noise amplification. It is clear from Fig. 3 that all the local areas got enhanced precisely.

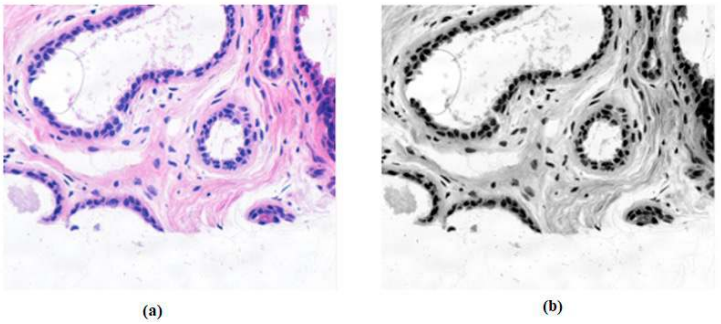

Fig. 3. Image enhancement using contrast-limited adaptive histogram equalization. (a) Original image; (b) Enhanced image.

Figure 4 represents the segmentation results using various algorithms. The study utilized fuzzy *c*-means, *k*-means clustering and active contour model for the segmentation purpose. The UCSB dataset consists of 58 images having a resolution of 896×768. The ground truth images are also available in the UC Santa Barbara library. Fig. 4 depicts the segmentation results of various algorithms and it is clearly visible how all these methods worked on the input image. After a visual examination on Fig. 4, it is confirmed that *k*-means clustering algorithm performed well. Therefore, various clustering segmentation methods were studied here to preserve the desired statistics during the segmentation procedure. The number of clusters was decided to be three and squared Euclidean distances are used as similarity measures.

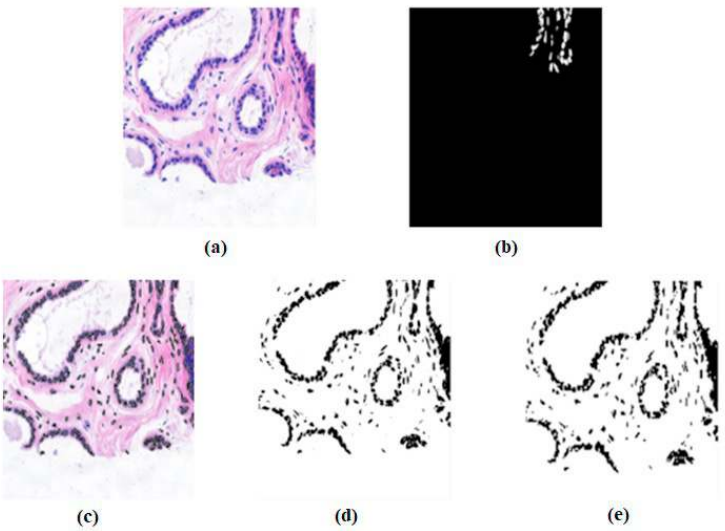

Fig. 4. Segmentation results for various algorithms. (a) Original image; (b) Ground truth image; (c) *k*-means clustering; (d) Fuzzy *c*-means clustering; (e) Active contour model.

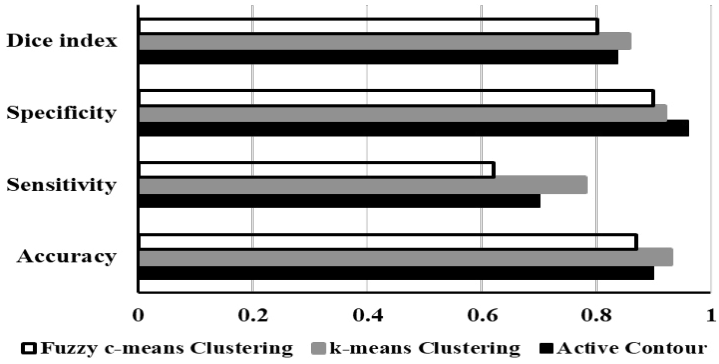

Fig. 5. Comparisons of various segmentation algorithms on the basis of performance measures.

From Fig. 5, it is noted that the *k*-means based clustering technique is connected with higher value of accuracy compared to fuzzy *c*-means and active contour model. Considering specificity, there is no much difference between *k*-means clustering and fuzzy *c*-means algorithms whereas sensitivity is very high for *k*-means clustering algorithm. Dice index value is high for *k*-means clustering rather than other two algorithms. Hence, it is shown that *k*-means clustering approach is preferred as the segmentation approach in this study for automatic cancer detection from histopathology images.

Randomly selected 48 images from UCSB dataset were used for classification. Thus 48 images were used for training and 10 images were used for testing. Then a 10-cross validation was performed to validate the results. The SVM classifier is a simple classifier in which a hyperplane is going to separate the different classes. For SVM's linear kernel function, a quadratic programming (QP) optimization constraint was applied to find the hyperplane. The SVM classifier performance was computed by means of a confusion matrix and all the performance measures in point values were depicted in Fig. 6. All these values were estimated by using equations (13)–(18).

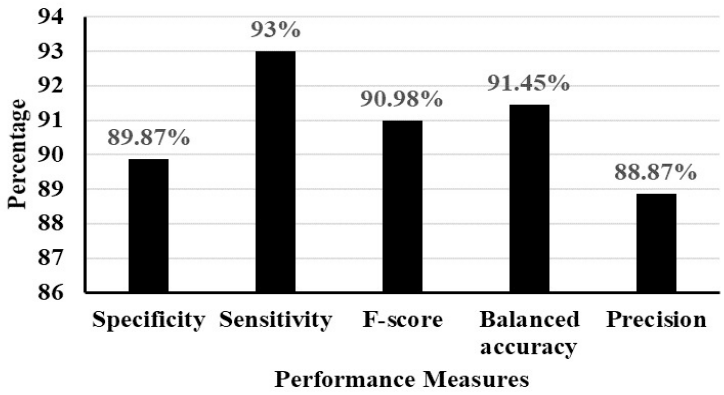

Fig. 6. SVM classifier performance on the basis of various parameters (Specificity, Sensitivity, F-score, Balanced accuracy and Precision).

The main contributions of this paper are as follows:

- A comparison on three machine learning algorithms viz: *k*-means clustering algorithm, active contour model and fuzzy *c*-means clustering algorithm.
- Experimental evaluation is done using various quantitative measures to show the performance of three segmentation algorithms.
- A support vector machine classifier is used to classify images into benign and malignant and performance evaluation has been done and demonstrated in a graphical format.

6. Conclusion

The main objective of this study was to enhance the breast cancer detection accuracy using computer aided diagnosis approach. In view of this goal, the paper presented the contribution, its outline, flowchart and quantitative parameters in a simulation environment. Based on the proposed methodology, this study used a publicly available dataset known as UCSB, with normal (benign) and abnormal (malignant) images for performance analysis of segmentation algorithm for breast cancer detection in a simulation environment. The segmentation has been done using three algorithms viz: fuzzy *c*-means clustering, *k*-means clustering and ACM. Among these, *k*-means clustering has the maximum segmentation accuracy of 93%. Other performance metrics are also superior for *k*-means clustering algorithm. The automated classification of breast cancer images is based on SVM model on GLCM features and achieved an accuracy of 91.1%. The SVM classification approach suggested in this study permits to achieve high sensitivity also. The performance of our study is comparable or inferior to the up-to-date approaches. The future scope of this study would be the application of proposed method on different histopathology images in wide sense. Moreover, this study can be implemented in real-time detection of breast cancer.

References

[1] F. Bray, J. Ferlay, I. Soerjomataram, M. Ervik R.L. Siegel, L.A. Torre, A. Jemal. (2015) "Global Cancer Statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries." *A Cancer Journal for Clinicians* **68(6)**: 394–424.

[2] L.A. Torre, F. Islami, R.L. Siegel, E.M. Ward, A. Jemal. (2017) "Global cancer in women: burden and trends." *Cancer Epidemiology, Biomarkers & Prevention* **26(4)**: 444–457.

[3] Kumar, Rajesh, Rajeev Srivastava, and Subodh Srivastava. (2015) "Detection and classification of cancer from microscopic biopsy images using clinically significant and biologically interpretable features." *Journal of Medical Engineering* **2015**: 1–14, Article ID 457906.

[4] M. M. R. Krishnan, C. Chakraborty, R. R. Paul, and A. K. Ray. (2011) "Quantitative analysis of sub-epithelial connective tissue cell population of oral submucous fibrosis using support vector machine." *Journal of Medical Imaging and Health Informatics* **1(1)**: 4–12.

[5] Ali, W. A. Wani, and K. Saleem. (2011) "Cancer scenario in India with future perspectives." *Cancer Therapy* **8**: 56–70.

[6] S. Bhattacharjee, J. Mukherjee, S. Nag, I. K. Maitra, and S. K. Bandyopadhyay. (2014) "Review on histopathological slide analysis using digital microscopy." *International Journal of Advanced Science and Technology* **62**: 65–96.

[7] John M. Eisenberg. (2016) "Core-Needle Biopsy for Breast Abnormalities." In: Comparative Effectiveness Review Summary Guides for Clinicians [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US), 2007. Available from: https://www.ncbi.nlm.nih.gov/books/NBK368367/

[8] M. N. Gurcan, L. Boucheron, A. Can, A. Madabhushi, N. Rajpoot, and B. Yener, (2009) "Histopathological image analysis: A review." *IEEE Reviews in Biomedical Engineering* **2(1)**: 147–171.

[9] O. Sertel, J. Kong, H. Shimada, U. Catalyurek, J. Saltz, and M.N. Gurcan. (2009) "Computer-aided prognosis of neuroblastoma on whole-slide images: Classification of stromal development." *Pattern Recognition* **42(6)**: 1093–1103.

[10] S. Doyle, M. Feldman, J. Tomaszewski, and A. Madabhushi. (2012) "A boosted Bayesian multiresolution classifier for prostate cancer detection from digitized needle biopsies." *IEEE Transactions on Biomedical Engineering* **59(5)**: 1205–1218.

[11] Basavanhally, S. Ganesan, M. Feldman, N. Shih, C. Mies, J. Tomaszewski, and A. Madabhushi. (2013) "Multi-field-of-view framework for distinguishing tumor grade In ER+ breast cancer from entire histopathology slides." *IEEE Transactions on Biomedical Engineering* **60(8)**: 3052–3055.

[12] M. Balazsi, P. Blanco, P. Zoroquiain, M.D. Levine, and Jr.M.N. Burnier. (2016) "Invasive ductal breast carcinoma detector that is robust to image magnification in whole digital slides." *Journal of Medical Imaging* **3(2)**: 027501.

[13] C. Mercan, S. Aksoy, E. Mercan, L.G. Shapiro, D.L. Weaver, and J.G. Elmore. (2018) "Multi-instance multi-label learning for multi-class classification of whole slide breast histopathology images." *IEEE Transactions of Medical Imaging* **37(1)**: 316–325.

[14] M. Kowal, P. Filipczuk, A. Obuchowicz, J. Korbicz, and R. Monczak. (2013) "Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images." *Computers in Biology and Medicine* **43(10)**: 1563–1572.

[15] P. Filipczuk, T. Fevens, A. Krzyzak, and R. Monczak. (2013) "Computer-aided breast cancer diagnosis based on the analysis of cytological images of fine needle biopsies." *IEEE Transactions on Medical Imaging* **32(12)**: 2169–2178.

[16] Y.M. George, H.H. Zayed, M.I. Roushdy, and B.M. Elbagoury. (2014) "Remote computer-aided breast cancer detection and diagnosis system based on cytological images." *IEEE Systems Journal* **8(3)**: 949–964.

[17] A.D. Belsare, M.M. Mushrif, M.A. Pangarkar, and N. Meshram. (2015) "Classification of breast cancer histopathology images using texture feature analysis." In *Proceedings of the TENCON 2015—2015 IEEE Region 10 Conference,* p. 1–5.

[18] P.S. Pawar, and D.R. Patil. (2013) "Breast cancer detection using neural network models." In *Proceedings of the International Conference on Communication Systems and Network Technologies,* p. 568–572.

[19] M. Sameti, R.K. Ward, J. Morgan-Parkes, B. Palcic. (2009) "Image feature extraction in the last screening mammograms prior to detection of breast cancer." *IEEE Journal of Signal Processing* **3(1)**: 46–52.

[20] M. Reza Ali. (2004) "Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement." *Journal of VLSI Signal Processing* **38(1):** 35–44.

[21] E. Gelasca, B. Jiyun, O. Boguslaw, and B.S. Manjunath. (2008) "Evaluation and benchmark for biological image segmentation." In *Proceedings of the IEEE International Conference on Image Processing*, p. 1816–1819.

[22] Na, Shi, Liu Xumin, and Guan Yong. (2010) "Research on *k*-means clustering algorithm: An improved *k*-Means clustering algorithm." In *Proceedings of 3rd IEEE International Symposium on Intelligent Information Technology and Security Informatics,* p. 63–67.

[23] D. Zarpalas, P. Gkontra, P. Daras, And N. Maglaveras. (2014) "Gradient-based reliability maps for ACM-based segmentation of hippocampus." *IEEE Transactions of Biomedical Engineering* **61(4):** 1015–1026.

[24] H. Zhou, G. Schaefer, and C. Shi. (2009) "Fuzzy c-means techniques for medical image segmentation." *Fuzzy Systems in Bioinformatics and Computational Biology* **242**: 257–271.

[25] Chattopadhyay, Subhagata, Dilip Kumar Pratihar, and Sanjib Chandra De Sarkar. (2012) "A comparative study of fuzzy c-means algorithm and entropy-based fuzzy clustering algorithms." *Computing and Informatics* **30(4)**: 701–720.

[26] R.M. Haraclick. (1973) "Texture features for image classification." *IEEE Transactions on Systems, Man, and Cybernetics* **3(6)**: 610–621.

[27] M.A. Aswathy, and M. Jagannath. (2017) "Detection of breast cancer on digital histopathology images: present status and future possibilities." Informatics in Medicine Unlocked **8**: 74–79.