1. INTRODUCTION

Breast cancer (BC) identification and diagnosis has for all time been a foremost concern for the pathologists and even for medical practitioners. 32 per cent of Indian inhabitants get cancer at some time in their life1. For Precise detection of BC, experts and medical practitioners prefer microscopic biopsy images collected under the microscope. In histopathology, BC biopsy images will be characterised into cancerous one or normal one2,3. Highly (40 x/100 x) magnified biopsy image provides consistent information about abnormal and normal tissues. Later the segmentation and categorisation application can be continued with other defence and military applications where we considered IRS satellite images for segmentation of required areas.

Plissiti4, et al. proposed color gradient watershed transform using 90 pap-stained cervical images of resolution 1536 x 2048 pixels and obtained 6 shape, 8 texture and 3 intensity features. They utilised maximum-relevance with minimum-redundancy (MR-MR) criterion for feature selection. They handled cell level diagnosis using image Processing methodologies. Bergmeir6, et al. presented a model for obtaining the local histograms and GLCM texture features. Huang and Lai7 explained a methodology for segmentation and categorisation methods for histology images basing on texture features and with help of SVM the highest Categorisation accuracy obtained is 92.8 per cent.

Adem Kalinli8, et al. considered otsu thresholding approach with the classifiers namely k-nearest neighbours, radial basis neural networks, support vector machines and k-means clustering, naïve bayes and functional trees for object extraction followed by classification. Kasmin10, et al. obtained the features of BC tissue images possessing area, perimeter, solidity, convex area, orientation filled area, major axis length, ratio of cell and nucleus area, eccentricity, mean intensity of cytoplasm, and circularity. The efficacy of other classifiers such as SVM, random forest, and fuzzy k-means is also examined. Proposed work uses Ductal Carcinoma (DC) BC images, as >80 per cent of BC is because of ducts. Also 40 X magnified Hematoxylin & Eosin (HE) DC images were chosen for clear cell segmentation.

Ali5, et al. considered Active contour models using multiple level sets for segmentation of 14 BC histology images with resolution 512 x 512 pixels and generated shape features to obtain an accuracy more than 90 per cent. Fatakdawala9, et al. proposed expectation maximisation driven geodesic active contour (EMaGAC) without and with overlap resolution using 100 breast histology images with resolution of 200 x 200
pixels and generated texture features to attain sensitivity of 86 per cent, Positive-Predicted Value of 64 per cent and overlap resolution of 90 per cent.

Kanchanamani, et al. discussed the various classifiers namely Support Vector Machine (SVM), naive bayes, k-nearest neighbour, multi-layer perceptron, and linear discriminant analysis and obtained best classification Sensitivity, Specificity and Accuracy values for SVM as 89.2 per cent, 96.4 per cent, and 92.5 per cent. Mouelhi, et al., defined modified geometric active contour model and touching nuclei method for nuclei detection. In the paper, they extracted colour and shape features and obtained segmentation accuracy of 97 per cent.

George, et al. defined marker-controlled watersheds transform using 92 breast cytological images of 640×480 pixels and extracted 12 statistical, 10 textural and 2 intensity features to evaluate classifier sensitivity and specificity as 95.49 per cent and 83.16 per cent, respectively.

Till now, less research is engaged on segmentation of noise corrupted, less contrast BC histopathological microscopic biopsy images. Those images lead to poor and erroneous segmentation of local image intensity non-homogeneities, and ultimately failing to extract true edges. Without the optimal segmentation, the processed result cannot be further utilised and preferred for classification. This paper proposes two new segmentation methodologies for attaining optimum segmentation with highest accuracy.

2. METHODOLOGY

Automatic detection, classification of BC biopsy images from microscope is quite challenging as the images retain clusters and overlapping cells. Distinctive stages involved in methodology include the tissue image enhancement using non-subsampled contourlet (NSC) transform for detection of edge levels are presented. Proposed methodology is presented in Fig.1. The details of stages utilised in methodology are discussed as follows.

2.1 Image enhancement

Preprocessing necessity is to reduce noise, improve processing speed, for contrast (quality) improvement of area of interest. Biopsy images gathered from light microscope may lack in some respect like uneven staining and poor contrast, so enhancement improves the contrast between the foreground (area of interest) and background. Circular Hough (CH) transform described by Hrebien, et al. is preferred for detection of edges its result is presented in Fig.3(b). For image enhancement initially contrast-limited adaptive histogram (CLAH) equalisation is used. CLAH equalised output is presented in Fig. 3(c).

![Figure 2. NSC transform filter bank structure.](image)

![Figure 3. (a) BC biopsy image, (b) CH cell detected output, (c) CLAH enhanced image, and (d) edge detected image using NSC transform.](image)
2.1.1 Non Subsampled Contourlet Transform
Wavelets became efficient alternative to Fourier methodologies in microscopic image applications because of differing window size (wide for slow and narrow for high frequencies i.e., optimised time, frequency resolution). Resultants are \( \frac{1}{4} \) size of original tissue. Smoothed image can further decompose into sublevels for further iterations. Wavelet transform being 1-D, smoothness of curves in tissue image is limited to 1-D, which leads to implementation of contourlets (a 2-D transform). Contour let is non-shift invariant because it utilises up and down samplings in Directional Filter Bank (DFB) and Laplacian Pyramid (LP) construction. It leads to NSC transform implementation\(^2\). Its filter bank structure is presented in Fig. 2.

Contrast with contour let, NSC transform possesses (a) non-subsampled LP obtained by non-subsampled 2-channel 2-D filter banks and (b) non-subsampled DFB obtained by switching of down/up samplers in 2-channel DFB structure while filter up sampling accordingly. Non-subsampled LP (NSLP) preserves multiscale feature and Non-Subsampled DFB (NSDFB) possess directionality. Let the edge outcomes attained by NSC transform using NSLP and NSDFB be \( e(p, q) \). The outcomes are better than the DWT based method as presented in Fig. 3(d). NSC decomposes BC image into approximation sub-band (that preserves the content of input image) and detail sub-bands (that stores intensity changes in all directions). The energy coefficients from all decomposed detail sub-band levels is formulated as follows.

\[
e(p, q) = \frac{1}{N \times N} \sum_{i,j=1}^{N} |W_L(p,q)|^2
\]

And/or its standard deviation (SD) from NSC decomposed BC tissue image on every directional sub-band is formulated as

\[
\sigma_L(p,q) = \left( \frac{1}{N \times N} \sum_{i,j=1}^{N} |W_L(p,q)|^2 \right)^{\frac{1}{2}} - \mu_L
\]

Where \( \mu_L \) is the mean of \( L^p \) sub-band defined as

\[
\mu_L = \frac{1}{N \times N} \sum_{i,j=1}^{N} |W_L(p,q)|
\]

where \( p, q \) are the available total rows, columns of decomposed tissue image respectively, \( W_L \) is the \( L^p \) coefficient of the NSC decomposed sub-band, and \( L \times L \) is the NSC decomposed sub-band. It results with energy and SD feature vectors as \( f_L = [E_1, E_2, \ldots, E_n] \) and/or \( f_{\sigma L} = [\sigma_1, \sigma_2, \ldots, \sigma_n] \) respectively, where ‘\( n \)’ defines the number of obtained directional sub-bands. Pixel energy defines its information regarding the intensity variation. NSC energies of every pixel in BC image are evaluated (considering all direction average values) to construct the energy map, which gives the distinguished energy variations contained in it.

2.2 Cell Segmentation

2.2.1 K-Means Partitioning
Pre-processing is required for separation of nuclei from its false positives. K-means\(^2\) clustering is used for splitting the nuclei from false positives i.e. to divide N observations \((i_1, i_2, \ldots, i_n)\) of image I into K areas \((K \leq N)\), \((A_1, A_2, \ldots, A_k)\) for all \( A_i \) and mean \( m_i \) as

\[
KM = \arg \min_{j=1}^{K} \sum_{i=1}^{N} \left| \| I - m_j \|_F \right|
\]

The resultant of k-means is as presented in Fig. 4(b).

2.2.2 Proposed Model 1 (PM-I)
Reliable unsupervised Watersheds concept in image processing depends on assumption of an image I as topographic 3-D space, with intensity (elevation information) versus two spatial coordinates. To get watershed lines of I with local minima \( L_0 = \{I_1, I_2, \ldots, I_M\} \) flooding process is initiated from \( n_{\min} = I_{\min} + 1 \) to \( n_{\max} = I_{\max} + 1 \), where \( I_{\min}, I_{\max} \) are the respective lower, higher intensity values and \( c_2(L) \) be the points within catchment basin due to local minimum \( L_i \). During flooding process, the union of catchment basins at \( n^{th} \) stage, \( c_2(n) \), can be related as \( c_2(n) = c_2(n) + 1 = T(I_{\max} + 1) \), Where \( T(n) = \{x | I(x) < n\} \) is the image points set with intensity less than \( n \) and its set of connected intensity components be \( R \). For each connected component \( c_2 \) in \( R \), the intersection \( k \) with \( c_2(n) \) is evaluated as \( k = r \cap c_2(n) \). Further \( c_2(n) \) set is derived sequentially using \( c_2(n) \) based on possibilities of \( k \) as:

- If \( k \) is vacant or if \( k \) possess one component of \( c_2(n) \), then \( k \) belongs to an existing basin of local minimum i.e.

where \( n_{\min} \) and \( n_{\max} \) are \( \{L_1, L_2, \ldots, L_M\} \) respectively.

Application of traditional watershed using image gradient produces over segmentation The image gradient \( g(p, q) \) is obtained using gray scale dilatation and erosion, with structuring element (SE) S. So, Marker Controlled Watershed (MCW) method is introduced. Xiaopeng\(^4\), et al. discussed that due to low SE size, opening, and closing morphology will detach bright, dark details and large SE eradicate large contours. So in PM-I altering SE size is constructed for every pixel in the tissue image \( I(p, q) \) which is defined as Adaptive Marker Controlled Watershed (AMCW) method. Initially SE map, \( M(p, q) \) is evaluated using the weighted variance, and weighting coefficient discussed by Rajyalakshmi\(^10\), et al. Then all pels in \( I(p, q) \) are altered using \( M(p, q) \) with opening (\( \circ \)) and closing (\( \bullet \)) operations as

\[
M(p, q) = [g(p, q) \circ M(p, q)] \bullet M(p, q)
\]

In this PM-I, the edge details attained by NSC transform \( e(p, q) \) are combined with \( M(p, q) \) using fusion technique and the observed results are as depicted in Fig.4(c).
The component \( q \) obtained from equation (1) is integrated:

\[
D_{EF} = I_{FE} \text{ SCI. J., VOL. 4, NO. 1, JANUARY 2019, D}
\]

\[
\text{else zero. Where } F_q \text{ is the membership function, that relies on number of LS}
\]

\[
\text{functions initiated. For energy minimisation, it is formulated by integrating e(p,q) with MLS-LCIF to overcome local in homogeneities problem. And the initial front for MLS is provided by NSC transform. Energy formulated by integrating e(p,q) with MLS-LCIF is}
\]

\[
e_q = \sum_{j=1}^{N} \int_{\Omega \cap O_j} F_G(q-p) e(p,q) - b(q)c_j \| u_k(p) \| dp
\]

\[
\text{The above equation (8), defines the integrated LCIF due to the response attained from NSC transform. Now the integrated LCIF is applied for the whole image to extract global function. Where, } u_q(p) \text{ is the membership function, } u_q(p) = 1, \text{for } p \epsilon \Omega_q \text{ else zero. } F_G(q-p) = 0 \text{, for } p \not\epsilon O_q.
\]

- Integration of NSC edge details with LCIF

NSC edge detected results are combined with MLS results, to obtain nuclei boundaries which removes complete intensity in homogeneities presented in Fig. 4(d). NSC solves the problem of coarser details and also it achieves all fine details of the image. So, for cell extraction in H & E images the NSC edge details \( e(p,q) \) obtained from equation (1) is integrated with LCIF to overcome local in homogeneities problem. And the initial front for MLS is provided by NSC transform. Energy formulated by integrating \( e(p,q) \) with MLS-LCIF is

\[
e_q = \sum_{j=1}^{N} \int_{\Omega \cap O_j} F_G(q-p) e(p,q) - b(q)c_j \| u_k(p) \| dp
\]
For $N=2$, let $\phi$ be the LS function used to represent $\hat{\Omega}$ as $\hat{\Omega} = \rho$ for $\phi(\rho) > 0$, and $\Omega_2 = \rho$, for $\phi(\rho) < 0$. To reduce $e_q$ in $\hat{\Omega} \eta q$, $e_q$ integral must be minimised with respect to $q$ over the image domain $\hat{\Omega}$. In PM-II formulation, total energy $e$ using dependent parameters $(\phi, b_c)$ is obtained as,

$$ e(\Phi, b_c) = \left[ \frac{1}{\hat{\Omega}} \sum_{j=1}^{N} \left[ F_G(q-p) \right] e(p.q) - b(q)c_j |^2 dq \right] M^F \left( \phi(p) \right) dp $$

(9)

Replacing,

$$ e(\Phi, b_c) = \left[ \sum_{j=1}^{N} e_j(p) \right] M^F \left( \phi(p) \right) dp $$

(10)

where $e_j(p) = \left[ F_G(q-p) \right] e(p.q) - b(q)c_j |^2 dq$ and

$$ M^F \left( \phi \right) = M^F \left( \phi_1(q),q \right), \phi = \phi_1, \phi_2(q) $$

$M^F \left( \phi \right)$ is the membership function described as

$$ M^F_1(\Phi) = H^F(\Phi) $$

and $M^F_2(\Phi) = 1 - H^F(\Phi)$ with Heaviside smoothed function, $H^F(\Phi) = \frac{1}{2} \left[ 1 + \frac{2}{\pi} \text{arctan}(\Phi) \right]$. For $N > 2$, two or more LS functions are used to represent $N$ disjoint regions $\hat{\Omega}_1, \hat{\Omega}_2, \ldots, \hat{\Omega}_N$. When $N=3$, two LS functions $\phi_1$ and $\phi_2$ are defined using $M^F_1(\phi) = H^F(\phi)$.

$$ M^F_1(\phi_1, \phi_2) = H^F(\phi_1)H^F(\phi_2) $$

and

$$ M^F_2(\phi_1, \phi_2) = (1 - H^F(\phi_1))H^F(\phi_2) $$

Let LS functions be represented by a vector function $\Phi = (\phi_1, \phi_2, \ldots, \phi_j)$.

The energy functional $E^\Phi$ is

$$ E^\Phi(\Phi, b_c) = e(\Phi, b_c) + R(\Phi) $$

(11)

where energy regularisation term $R(\Phi)$, defined as

$$ R(\Phi) = \sum_{\omega=1}^{N} R(\phi_\omega) $$

for LS function $\Phi = (\phi_1, \phi_2, \ldots, \phi_j)$. Considering regularisation $R(\Phi) = \sum_{\omega=1}^{N} |\rho| \ V \phi_\omega |$ term into account, the term $E^\Phi(\Phi, b_c)$ in equation (11) can be minimised using the principle discussed in Rajyalakshmi, et al. The integrated response of NSC edge details with MLS extracts clearly the nuclei shown in Fig. 4(d). Performance evaluations for traditional K-means, ACMC, MLS, PM-I, PM-II were presented in Table 1.

### 2.3 Feature Extraction

Mouelhi, et al. extracted Haralick’s, oriented gradients histogram, 4 colour based moment features to categorise cancerous One over BC biopsy images. A number of comprehensive discussions were reviewed in the survey for segmentation and classification. Tables 2 (a) and 2 (b) show the features extracted in the proposed model.

These features set act as input for categorisation using supervised classifiers.

#### 2.4 Supervised Classification using Features Set

Using the segmented features attained from PM-II, the testing data set is classified using K-nearest neighbour (KNN) classifier studied in Altman’s multi-class support vector machine (MC-SVM) classifiers studied in Hsu, et al. The entire classification is done using the shape, intensity and texture features elevated from 96 trained images. The response of the ROC plot for KNN and M-SVM are compared and proved that M-SVM provides better classifier accuracy over KNN.

##### 2.4.1 KNN Classifier

KNN classifies images based on closest trained samples in feature space. Classification also relies on highest vote of neighboring points. Problem occurs while selecting the neighbors, which are closest to each sample. Shortest distance is chosen for classifying BC tissue images. Euclidean distance in-between pixel $(p_r, q_r)$ and pixel $(p_s, q_s)$ is

$$ D(x, y) = \sqrt{(p_r - p_s)^2 + (q_r - q_s)^2} $$

(12)

##### 2.4.2 MC-SVM Classifier

Let $F$ be feature set, $w$ be vector splitting hyperplane, then optimal minimised function to be minimised for MC-SVM is:

$$ O_F = \min \left[ \frac{1}{2} w^T C \left( \sum_{j=1}^{N} \epsilon_j \right) \right] $$

(13)

Where $C$ is the user defined parameter, and

$$ 0 \leq \epsilon_j \leq 1 - y_j (f_j(w + b)) $$

label of $f_j$ \{1, 1\}, Hyperplane normal splitting vector $w$ is evaluated as

$$ w = \sum_{j=1}^{N} \epsilon_j y_j f_j $$

(14)

### 3. RESULTS AND DISCUSSIONS

Performance evaluation of classifiers is compared using average sensitivity, classifier accuracy, and average specificity. Average Balanced Classifier Rate (BCR), geometric mean of average values of specificity and selectivity as well as Area below curve (AUC) of classifier response are chosen for evaluation. Overall proposed method performance is assessed
Table 2(a). Textural features extracted

<table>
<thead>
<tr>
<th>Feature type</th>
<th>Feature type</th>
</tr>
</thead>
<tbody>
<tr>
<td>33 textural</td>
<td>6 histogram, 22 Haralick’s, 3 Tamura’s, 2 GRLM features.</td>
</tr>
<tr>
<td>(includes 6</td>
<td></td>
</tr>
<tr>
<td>histogram)</td>
<td></td>
</tr>
<tr>
<td>22 Haralick’s features: Uniformity (Energy/Angular Second Moment), Dissimilarity, Entropy, Inertia, Contrast, Inverse difference, Homogeneity / Inverse difference moment, correlation, Cluster Prominence, Cluster Shade, Autocorrelation, Sum of Squares, Maximum probability, Sum Variance, Sum Entropy, Sum Average, Difference entropy, Difference variance, Information measures of correlation, Information measures of correlation, Inverse difference normalized, Maximal correlation coefficient, Inverse difference moment normalized.</td>
<td></td>
</tr>
<tr>
<td>3 Tamura’s features: Coarseness, Contrast, Directionality.</td>
<td></td>
</tr>
<tr>
<td>2 GRLM features: Short and Long Run Emphasis</td>
<td></td>
</tr>
</tbody>
</table>

Table 2(b). Shape, intensity features extracted

<table>
<thead>
<tr>
<th>Type</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shape based features</td>
<td>Area, Perimeter, Compactness, Equivalent Diameter, Major axis length, Minor axis length, Orientation, Centroid, Nucleus to Cytoplasm ratio, Eccentricity, Convex area, Solidity, Extent.</td>
</tr>
<tr>
<td>Intensity feature{2}</td>
<td>Intensity_{max} and Intensity_{min}, Intensity_{mean}</td>
</tr>
</tbody>
</table>

Table 3. Performance assessment of classifiers

<table>
<thead>
<tr>
<th>Parameter</th>
<th>KNN</th>
<th>MC-SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average sensitivity (%)</td>
<td>97.13</td>
<td>98.24</td>
</tr>
<tr>
<td>Average selectivity (%)</td>
<td>93.76</td>
<td>97.57</td>
</tr>
<tr>
<td>BCR (%)</td>
<td>95.43</td>
<td>97.90</td>
</tr>
<tr>
<td>Classifier accuracy (%)</td>
<td>99.01</td>
<td>99.76</td>
</tr>
<tr>
<td>AUC</td>
<td>0.9703</td>
<td>0.9743</td>
</tr>
</tbody>
</table>

Table 4. Stage wise execution time

<table>
<thead>
<tr>
<th>Individual stage</th>
<th>Time in sec (mean±deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-Processing</td>
<td>2.35±0.21</td>
</tr>
<tr>
<td>Nuclei detection</td>
<td>3.26±1.56</td>
</tr>
<tr>
<td>K-means</td>
<td>8.53±0.83</td>
</tr>
<tr>
<td>MCW</td>
<td>10.62±1.34</td>
</tr>
<tr>
<td>Nuclei Segmentation</td>
<td>15.64±2.34</td>
</tr>
<tr>
<td>PM-I</td>
<td>6.83±0.76</td>
</tr>
<tr>
<td>PM-II</td>
<td>4.75±0.69</td>
</tr>
<tr>
<td>Feature Extraction</td>
<td>4.89±1.34</td>
</tr>
<tr>
<td>Tissue</td>
<td>5.21±1.39</td>
</tr>
<tr>
<td>Classification</td>
<td>1.17±0.67</td>
</tr>
</tbody>
</table>

By traditional parameters namely

Average Sensitivity $S_{EN} = \frac{T_P}{(T_P + F_N)}$

Average Specificity $S_{FP} = \frac{T_N}{(F_P + T_N)}$

Segmentation Accuracy $A_{CC} = \frac{T_P + T_N}{(T_P + T_N + F_P + F_N)}$

Computational Time ($C_t$): Difference between the return time of the algorithm at start-up of the process ($t_{start}$) and the return time at the end of the process ($t_{end}$). i.e., $C_t = t_{start} - t_{end}$

Precision $P_{PPV} = \frac{T_P}{(T_P + F_P)}$

Dice Similarity Coefficient $D = \frac{2(A \cap B)}{(A + B)}$

$F_1$ score $= \frac{2 \times P_x \times R_x}{(P_x + R_x)}$

where A and B respectively define the sets of algorithm and ground-truth segmented pixels,

True Positive $T_P$: The total number of correctly detected true tumor pixels. A detected pixel is considered as $T_P$ when its Euclidean distance with ground truth pixel location is < 7.5 μm(30 pixels),

True Negative $T_N$: The total number of correctly identified false tumor pixels as non tumor pixels,

False Positive $F_P$: The total number of detected pixels which are not appeared in the ground-truth,

False negative $F_N$: The total number of non detected pixels that are appeared in the ground-truth.

Classifier accuracy and performance assessment of classifiers as represented in Table 3 and observed that the integrated segmentation approach (PM-II) of Contourlet and MLS with MC-SVM classifier provides high system accuracy and best classifier rate over traditional K-means partitioning, NSC, MCW, MLS, PM-I segmentation methodologies. The execution time for different steps performed in the discussed model was as represented in Table 4. On part of execution time also PM-II provides better evaluation when merged with MC-SVM approach.

The proposed method is compared on basis of subjective performance measures using the literature techniques as described in Table 5. Using PM-II, we obtained better results than traditional approaches. The results were also compared to ground-truth results provided by the pathologist.
### Table 5. Quantitative state-of-art evaluation of proposed method using the literature techniques

<table>
<thead>
<tr>
<th>Reference</th>
<th>Number of images</th>
<th>Resolution (pixels)</th>
<th>Segmentation algorithms used</th>
<th>Features extracted</th>
<th>Quantitative performance measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ali5, et al.</td>
<td>100 histology</td>
<td>512 × 512</td>
<td>Active contour models using multiple level sets.</td>
<td>Shape</td>
<td>Hausdorff distance (HD) 5 pixels over 95 %, Mean Absolute Distance (MAD) A&lt;sub&gt;cc&lt;/sub&gt; &gt; 90 %</td>
</tr>
<tr>
<td>Fatakdawala6, et al.</td>
<td>100 histology</td>
<td>200 × 200</td>
<td>Expectation maximisation driven geodesic active contour (EMaGAC) without and with overlap resolution</td>
<td>Texture</td>
<td>S&lt;sub&gt;en&lt;/sub&gt; : 86 %, PPV : 64 %, HD : 2.1, MAD : 0.9, Overlap resolution : 90 %</td>
</tr>
<tr>
<td>Plissiti4, et al.</td>
<td>90 images</td>
<td>1536 × 2048</td>
<td>Color gradient watershed transform</td>
<td>6 shape, 8 texture, and 3 Intensity</td>
<td>HD: 1.71 (mean) ± 0.54 (std).</td>
</tr>
<tr>
<td>Mouelhi7, et al.</td>
<td>18 images with 3-amino-9-ethylcarbazole chromogen and 42 images 3-diaminobenzidine chromogen</td>
<td>2048 × 1360</td>
<td>Modified geometric active contour model and touching nuclei method</td>
<td>Color and shape</td>
<td>A&lt;sub&gt;cc&lt;/sub&gt; : 97 %</td>
</tr>
<tr>
<td>George8, et al.</td>
<td>92 breast cytological images</td>
<td>640 × 480</td>
<td>Marker-controlled watersheds transform.</td>
<td>12 statistical, 10 textural and 2 intensity</td>
<td>S&lt;sub&gt;en&lt;/sub&gt; : 95.49, S&lt;sub&gt;pe&lt;/sub&gt; : 83.16</td>
</tr>
<tr>
<td>Proposed method</td>
<td>96 training and 24 testing images</td>
<td>760 × 570</td>
<td>NSC integrated with MLS.</td>
<td>13 shape 33 textural and 2 intensity</td>
<td>S&lt;sub&gt;en&lt;/sub&gt; : 98.77 %, S&lt;sub&gt;pe&lt;/sub&gt; : 97.81 %, A&lt;sub&gt;cc&lt;/sub&gt; : 99.02 %, C&lt;sub&gt;r&lt;/sub&gt; : 16.42±4.47 s, PPV = 89.74, D&lt;sub&gt;1&lt;/sub&gt; = 0.95, F&lt;sub&gt;1&lt;/sub&gt; = 0.91</td>
</tr>
</tbody>
</table>

### 4. FUTURE SCOPE

Proposed Model-II is applied to electron microscopic (biopsy) images. PM-II proves as the best combination over the discussed traditional techniques. This module can also be preferred to all military applications, like remote sensing imagery, where initially need to acquire it and then understand the imagery, followed by segmentation of required area, and at last categorisation. Also applicable to ‘Autonomous vehicle’, an image processing advanced defence application. This vehicle possesses a computer vision module that grasps the 3D scenic images.

### 5. CONCLUSIONS

Automated characterisation approach was presented for BC recognition from microscopic tissue images with HE staining using clinically interpretable feature set. Optimised investigation on cell nuclei depends on enhancement and segmentation methodologies used. For image efficacy, CLAH equalisation method is chosen and biopsy image is enhanced using NSC transform for good recovery of edges. For nuclei segmentation, K-means, PM-I and PM-II techniques are utilised. AMCW retains all image details as SE size is adaptive. PM-I integrates AMCW method with NSC transform methodology; PM-II integrates NSC result with MLS. Among all categories, anticipated segmentation method performs better area extracted result than traditional MCW, region growing, MLS, PM-I approaches for nuclei extraction. After tissue image segmentation, obtained 49 features set for 96 training tissue images. Out of respective features, we choose 11 features for characterisation. Diverse sorting methods used were kNN, and MC-SVM classifiers. Local inhomogeneities problem is also solved using NSC integrated with AMCW transform. There is hike in performance measures i.e., average specificity, accuracy, sensitivity, BCR for MC-SVM to 97.57, 99.76, 98.24, 97.90 respectively over KNN classifier. Also there is an enhancement of AUC for MC-SVM to 0.9703 compared to KNN. The segmentation and classification results were tested for 24 images and compared with the manual results taken from the pathologist. The Average sensitivity, average specificity, PPV, average accuracy of all BC histopathological images using the proposed method were obtained as 98.77 per cent, 97.81 per cent, 87.94 per cent and 99.02 per cent, respectively. The computational time required for whole process including pre-processing, detection, nuclei segmentation, and nuclei classification is 16.42 ± 4.47 s. Higher the F1-score and dice index better is the classification, those are obtained as 0.91
and 0.95, respectively. The PM-II provides better subjective evaluation techniques compared to the literature techniques discussed.

REFERENCES


**ACKNOWLEDGEMENTS**

The authors acknowledge Department of Science & Technology, Government of India for financial support with their vide sanction reference No.100/ (IFD)/1283/2015-2016, under Women Scientist Scheme to carry out this work. The authors also thank Dr U. Ramesh, Asst Professor, Rangaraya Medical College for providing high resolution H&E stained breast cancer histopathological Images.

**CONTRIBUTORS**

Ms Rajyalakshmi Uppada is working as a Professor in Aditya Engineering College, Surampalem, Kakinada. She completed her Ph.D from JNTU Kakinada. She has published 12 journal paper and presented 6 international conference papers.

In the current study, she is responsible for the proposal of the technique, software implementation and testing the results.

Dr Koteswararao Sanagapallela is presently working as a Professor in KL University, Vijayawada, India. He has retired as a Scientist ‘G’, Associate Director, DRDO-NSTL. He has published 30 IEEE paper. With his thirty two year of design and development experience and expertise in the Anti Submarine warfare (ASW) Fire Control Systems for torpedoes and rocket launchers, he have created a strong edifice in the weapon technology in the country at NSTL.

In the current study, he has provided guidance for implementing the proposed technique.

Dr Satyaprasad Kodati is presently working as Sr. Professor in Electronics & Communication department & Rector, Vignan’s Foundation for Science, Technology & Research, Vadlamudi, Guntur. He retired as a Professor in Electronics & Communication department and Director-IST, JNTU Kakinada, India. He has held different positions as Head of ECE Department, vice-principal, principal ofUCEK, JNTUK, and as Director of Evaluation and Rector at JNTUK, Kakinada. He has published more than 70 Journal Publications and he has presented more than 50 conference proceeding.

In the current study, he has provided guidance for implementing the proposed technique.