Novel Texture Pattern Based Multi-level set Segmentation in Cervical Cancer Image Analysis

Arti Taneja¹, Dr. Priya Ranjan², Dr. Amit Ujlayan³

¹Research Scholar, Amity Institute of Information Technology, Uttar Pradesh, Noida -201303, India. Ph.: 0120-4392277
²Professor, Amity University, Uttar Pradesh, Noida, India. Ph.: 0120-4392277.
³Professor, Gautam Buddha University, Greater Noida, India.

Abstract - Computerized framework development is alternate to the Manual Method (MM) of cervical cancer analysis since MM suffers from the human errors, bulk quantities of Pap smear images, work loading and time complexities. Also, severity level prediction via counting of nucleus leads to incorrect prediction under geometric-based feature extraction. The cell-based segmentation evolved in research studies assures the automatic assistance with an assumption of a single cell. Practically, this assumption is not suitable since the image contains more cells. The multiple cell splitting for a number of cells is the challenging task. Also, the complex cell structure, poor contrast, and overlapping affect the cell segmentation performance. The cell contains a nucleus that describes the significant changes due to the disease. Hence, the cell boundary prediction is the complex task of geometric-based feature extraction techniques. This paper proposes the four new methods to update the cervical image processing tasks. First, the Neighborhood Concentric Filtering (NCF) is used to remove the noise present in the image and enhance the intensity level. Then, the cluster formation based on the intensity difference level estimation provides the multi-label output. Second, the Optimal Weight Updating with the Multi-Level set (OWU-ML) estimates the Region of Interest (ROI-Nucleus), extracts the nucleus texture features with an edge intensity information and form the window. Also, the intensity weight update by the OWU efficiently separates the layers that form the active contour over the image. Here, the Gray Level Co-occurrence Matrix (GLCM) extracts the texture pattern features of the nucleus portions in the form of angle variations. Finally, the Neural Network-based RVM classifier predicts the classes of (class 1 and class 2) cervical images. The optimal weight update, the GLCM features based multi-level set segmented output and the NNRVM classification improves the performance of severity level prediction in cancer treatment provision.

Keywords: Cervical Image Analysis, Gray Level Co-occurrence Matrix, Multi-Label output, Neighborhood-Concentric Filtering, Neural Network-Relevance Vector Machine, Optimal Weight update.

1 Introduction

Global cancer statistics reports show that the death rate due to the cancer is more in the developing countries and need to be reduced. The observation of cancer incidence and the mortality patterns is the preventive action for cancer burden. Automation-assisted methods evolved in research studies are based on cervical cytology screening and this is the labor-oriented process with high-intensity. The cytological preparation is the prerequisite for the screening operation. The Integration of Manual Liquid-based Cytology (MLBC) with the hematoxylin and eosin (H&E) stain effectively prepares the cytology for cervical screening. The morbidity and the mortality prevention from the cervical images are the major objectives of cervical screening. Age-approximate and the maintenance of abnormal screening results requires an optimal strategy that has the capability to avoid the unnecessary treatment and makes the screening as a sensible one.

Research studies lie in two aspects namely, the nuclei detection in either single or overlapping cells and the nuclei detection in both single and overlapping cells. The contour enhancement approach for boundary detection employs the probability estimation of Gradient Vector Flows (GVF) for each image pixel. The integration of edge map computation method and the stack-based refinement makes the GVFs are robust. The GVF implementation and the Active Contour Model (ACM) integration requires the proper initial contour and this is not suitable for practical implementation. The Nucleus Cytoplast Contour (NCC), Adaptive Threshold Decision (ATD) and the Gray Level Gradient Difference (GLGD) extract the nucleus from the images and segment the required objects from various images.

The Cytological Testing-based cervical screening has the limitations of less sensitivity, reproducibility, and specificity. Human Papillomavirus (HPV) diagnosis is an alternative to CT that has the ability of a prior indicator of cervical cancer stage and it is preferable in clinical analysis. The combination of Pap smear and the DNA images plays the major role in HPV diagnosis. The Pap smear image analysis effectively reduces the death rate. But, the evolution of stable H&E stain method provides the accurate analysis.
The less operational speed of image screening and diagnosis leads to optical imaging techniques development. Fluorescence lifetime imaging microscopy (FLIM) provides the additional information regarding the lifetime independent of excitation power fluctuations and the fluorophores concentration. The FLIM-based screening provides the better screening and diagnosis than the H & E imaging method. But, the huge time-consuming nature of H & E stain methods leads to the automatic assistance methods generation. Automated methods applicable in segmentation offers the higher accuracy than the manual processing. The attractive stage in nuclei segmentation is the Graph Cut (GC) approaches employment. The prior knowledge regarding the nuclei shape and the incorporation of manual annotation provide the more robust segmentation performance. The basic assumption of these methods is the image contains only one cell is not suitable in practical situations. Hence, the multiple touching-based cell splitting in cervical image analysis plays the major role in cell-segmentation techniques.

Research works turns to segmentation of overlapping cells that depends on the appropriate nuclei and background marker selection. The overlapping cell segmentation is the combination of nuclei boundary characteristics and the prior knowledge regarding the expected shape. The cervical cancer treatment provision requires the prior cancer detection. The increase in mean scattering coefficient due to the acidic acid agents leads to the abnormal tissue appearance within the normal regions. The evolution of optical imaging techniques in research studies assures the effective diagnosis. The coordinated and cytology screening mechanisms implementation results show that the cervical disease is the major problem compared to other. The microscopic-based liquid cytology evaluation offers the high efficiency and the sensitivity values. But, the false rates are also high during the practical implementation. The automated cytology and Human Papilloma Virus (HPV) analysis reduce the false rates. The abnormal Pap smear image utilization in HPV analysis leads to slowness and low morphologic perceptibility. The statistical results from the AAR methods convey that the part in the cell (nuclei) describes the variations due to the disease affection. Hence, an accurate nuclei boundary description is the necessary task and it is crucial in PAP smear image analysis. The normal/abnormal cell discrimination in PAP smear images includes the quantification of nuclei changes that lead to the evolution of structure element based segmentation.

The cell structure complexities for high cytoplasm variations are high for the unknown cells and location. The generic and parametric methods utilization offers the better cells delineation than the existing shape segmentation methods. But, the poor contrast and the inconsistent staining leads to the morphological feature extraction, watershed segmentation, and the nuclei separation. The super pixel algorithms evolution in research studies contributes the image boundaries prediction, speed, memory, and efficiency improvement. The simple linear iterative clustering adapts the k-means clustering algorithms for the super pixels generation. The modern super pixel generation algorithms offer the better results compared to the traditional superpixel segmentation algorithms. The structural parts segmentation based on template fitting, edge detectors, and the active contour models has the high performance compared to the other segmentation algorithms. The large cells, overlapping and the non-appropriate image artifacts cause the limitation in recognition of boundaries that contains only one cell. The energy function optimization depends upon the elliptical shape, area overlap and intensity ratio. But, the cervical images are not exact elliptical and the adaptation of elliptical shape also not offered the solution to the accurate segmentation problem. The provision of controlling the overlapping cells enhances the accuracy of segmentation. The novel technical contributions of proposed Novel Texture Pattern-based Multi-level set Segmentation (NTPMS):

- Nucleus texture pattern extraction through the optimized cell segmentation techniques is proposed.
- The multi-level set algorithm proposed in this paper extracts the exact nucleus portion in the cervical images irrespective of the geometrical features.
- Periodical estimation of intensity difference values forms the cluster that provides the multi-label output.
- The optimal weight update in proposed algorithm separates the layers which form the active contour over the image.
- Nucleus texture pattern extraction by the GLCM provided the clear image analysis and normal/abnormal classification

This paper is organized as follows: Section II describes the related works on cervical cancer image analysis. Section III discusses the implementation process of proposed Novel Texture Pattern-based Multi-level set Segmentation (NTPMS) algorithm. Section IV presents the performance analysis of NTPMS regarding the accuracy, precision, recall, sensitivity, specificity and the coefficient metrics. Finally, section V presents the conclusion.

2 Related Work

This section describes the traditional feature extraction and segmentation techniques influenced in the cervical image analysis. The reports from the different medical societies declare that the cancer is the second leading cause of death. There are several cancer treatments such as radiotherapy, surgery etc. has the limitations such as drugs selection related to diseases and toxicities possession in cancer drugs. Saslow et al. [1] reviewed the update of the American Cancer Society (ACS) reports and provided the new screening recommendations that addressed the age-appropriate screening strategies with the high-risk human papillomavirus (HPV) basis. Jemal et al. [2] presented the statistics about the cancer incidence and mortality rates that conveyed overall incidence rates in developing world are half of the rates in developed world. The application of cancer control knowledge and the physical activity monitoring significantly reduced the global cancer burden.
The cervical cancer diagnosis through the screening process is the labor consuming and time complexity. Govindaraju et al [3] investigated the green synthesis that explored the silver nitrate (Ag NO3) influence on cervical cell analysis. With the considerations of selective toxicity and therapeutic index, the new horizon is declared for cell nuclei splitting. The Photo Acoustic diagnosis provides the solution to Cervical Cancer screening problems. Peng et al [4] carried out the PAI experiments for the Depth Maximum Amplitude Projection (DMAP) analysis. They employed the paired-t-test for an indication of Mean Optical Absorption (MOP) difference that exists in between normal and cervical tissue variation for high confidence value. The molecules, cells and tissues identification and characterization require a unique biomedical fingerprint and it is the major tool in cervical cancer analysis with the ability for malignancy and premalignancy stages. Ramos et al [5] summarized the research areas and assured the HPV detection and monitoring response compared to diagnosis perspective. They also presented the comprehensive studies to suggest the Raman spectroscopy validation for molecular diagnostics and cancer analysis.

In cervical screening, the morphological and biochemical properties alternation for malignant transformation approach in the optical imaging techniques improve the performance. Orfanoudaki et al [6] utilized the contrast agents to provide the three-dimensional cell clusters that provide an effective knowledge of biological events and the multi-center randomized cells utilization offers the high standardization. The Automated Assisted Reading (AAR) techniques deployment in cervical screening process has the less error rate and the maximum productivity which depends on the accurate segmentation of abnormal cells. Zhang et al [7] proposed the global/local scheme with the Graph-Cut (GC) approaches that utilized the combination of normal and abnormal cells. The tumor histopathology characterization defined the nuclear regions from the H & E tissue sections. Chang et al [8] performed the automated analysis by using the nuclear segmentation formulation within the graph framework. They presented the Multi-Reference Graph Cut (MRGC) with the prior knowledge regarding the reference and local image features. Zhang et al [9] introduced the autofocusing method that rejected the coverslip and the actual focal plan was extracted. The hybrid global and local scheme segmented the normal and the abnormal cells. The contextual and cytoplasmic information capture improved the specificity. Early detection of cervical cells is the challenging task in standardized treatment provision. The accurate detection and segmentation were limited by the cells overlapping and the less contrast. Happy et al [10] presented the unsupervised approach called Extended Depth of Field (EDF) for accurate cell nuclei segmentation.

The metabolism changes indication in EDF effectively detected the pre-cancer development status. Wang et al [11] used the Fluorescence lifetime imaging microscopy (FLIM) for metabolic changes detection. They studied the application of FLIM to the unstained tissues with the morphological features. The fluorescence lifetime analysis showed that the FLIM provided the accurate results than the H & E method. The adaptive GC method combines the intensity, texture, and shape and boundary information and offers the better abnormal cells segmentation. The accuracy enhancement in CC screening analysis involves the nuclei cell distribution and shape size analysis. Mahanta et al [12] provided the automated process to predict the cell abnormalities in the screening process. The interactive segmentation of normal/abnormal cells via fuzzy based histogram computation. The touching-based nuclei splitting involved the marker generation by the learning of texture, shape, and contextual information. Talukdar et al [13] proposed the validity measures for FCM such as Partition Coefficient (PC), Partition Entropy (PE), compactness and the separation function. Song et al [14] presented the combination of multi-scale convolution network and the graph partitioning for an accurate segmentation. The refinement via coarse-to-fine segmentation reduced the complexities effectively. They also provided the enhancement to establish the cut-off values between the normal/abnormal patterns and the abnormal values classification based on cancer stage. The true nuclei determination involves the shape, text and the image intensity characterization reduces the false positive findings. Plissiti et al [15] presented the automated method for boundary detection and cell nuclei determination. They examined the unsupervised and supervised classification techniques with the feature selection schemes with the minimal redundancy and maximal relevance.

The Pap smear image utilization suffers from the low value of sensitivity and the specificity leads to the HPV diagnosis. Tasoglu et al [16] presented the need of simple, easy HPV diagnosis methods and reviewed the existing methods to provide the future directions to cervical analysis. The manual screening-based Pap smear image test required the color and shape properties to make the automated one. But, the automated process was suffered from the cell structure complexities. Genctav et al [17] proposed the unsupervised approach for cervical cells classification/segmentation. The sequential processes such as automatic thresholding, hierarchical segmentation, and the binary classification maximized the similarities and offered the consistent staining without any parameters adjustment. The segmentation result depends upon the pixels selection belongs to a particular class. The complexities and non-obvious nature of Pap smear images affected the general assumption such that the pixels are distributed in nature. The pixels associated with the nuclei are noisy pixels that need an isolation process. Plissiti et al [18] presented the automation method for cell nuclei prediction in Pap smear images. They performed morphological analysis to detect the centroid locations of nuclei and it is incorporated with the nucleus circumference. The fully automated method had the ability to handle the overlapping images. Hyperthermia is a medical therapy for enhanced cancer treatments in which the external energy is applied to raise the tumor region temperature. The cell viabilities decrease immediately to the simulation starts requires an optimal labeling process. The concentration dependent toxicity effects required the
revision in magnetic labeling process. Huang et al [19] optimized the magnetic labeling process for the analysis of cytotoxicity’s evaluation. They also obtained the condition for magnetic label optimization that leads to a reduction in cell viabilities. The obscure boundaries locating and the noise sensitivity leads to an extension of Gradient vector Flow (GVF) models to the radiating modules (RGVF). Li et al [20] proposed RGVF that is used for segmentation of initial contours. The edge map computation and the stack based refinement in RGVF assured the robustness against the contaminations and located the obscure boundaries effectively.

The small size nucleus compared to the cytoplasm and the background leads to a poor threshold in Adaptive Threshold Decision (ATD) for discrimination. Pai et al [21]presented the Maximal Gray Level Gradient Difference (MGLGD) for nuclear extraction. The contour formation in NCC dependent upon the gray level difference between the nucleus and cytoplasm that offered the superior performance than the existing GVF-ACM-ATD without requiring proper contour initialization. The compact representation of shape and the active shape model utilization are necessary for prior knowledge regarding the expected shape. Plissiti et al [22] detected and described the unknown nuclei boundaries in the images. By using the weight parameters that controls the force and energy of deformable models. The problem inaccurate boundary detection was resolved by the weight parameter-based overlapping nuclei prediction. The noise and cell occlusion in overlapping images degrade the performance. Nosrati and Hamarneh [23] proposed the variational segmentation using the star-shaped prior using the directional derivatives for overlapping cells segmentation in Pap smear images. They introduced the Voronoi energy term that controls the neighbor cells overlapping. The accuracy and computational time of optimized approaches better than the non-optimized approaches. The reduction in preprocessing complexities requires the fast super pixel computation and the easy utilization. The segmentation quality and the operational speed improvement depends on the extracted superpixels. Achanta et al [24] introduced the linear iterative algorithm called SLIC that includes the $k$-means clustering approach for an efficient pixels generation. The SLIC implementation also offered the better segmentation performance and the maximum operational speed. The cluster of distracting pixels affected the super-voxels segmentation process and hence the proper algorithm is required for computational complexity reduction. Lucchi et al [25] proposed the automated graph partitioning methods that incorporate the shape and distinctive shape learning for better recognition. They demonstrated the better computational efficiency and segmentation quality. The review concluded that the suitable weight estimation and the multi-level set formulation are required to provide the optimal trade-off between the accuracy improvement and clear image analysis.

3 Novel Texture Pattern-based Multi-level set Segmentation

This section illustrates the methods involved in proposed work. The flow diagram of proposed Novel Texture Pattern-based Multi-level Segmentation (NTPMS) consists of successive processes such as Neighborhood Concentric Filtering (NCF), Optimal Weight Update-Multi-level set segmentation, Gray Level Co-occurrence Matrix (GLCM) feature extraction, and the Neural Network- Relevance Vector Machine (NN-RVM) classification as in Fig. 1.

Fig. 1 Flow diagram of NTPMS.

Initially, Neighborhood Concentric Filtering (NCF) removes the noise through the connected component analysis. The interpretation of information in the images is the necessary prior task in the segmentation process. Hence, the Gaussian model is used to enhance the image quality for better human perception. Then, the energy, direction and weight update in diverse directions from the active contour. The unique form of image representation depends on the performance of feature extraction. Here, the Gray Level Co-occurrence Matrix (GLCM) is used to extract the various features for unique representation. Finally, the Neural Network- Relevance Vector Machine (NN-RVM) classifier provides the necessary labeling.

3.1 Neighborhood Concentric Filtering

The good quality images are the major requirement for feature extraction and the segmentation. The noise removal and the image enhancement are the initial stages of proposed work. The window formation by using the connected components extraction removes the noise present in the images. The algorithm for Neighborhood Concentric Filtering (NCF) is as follows:

Initially, the matrix window with the size $3 \times 3$ is initialized for the input cervical image. Then, the Connected Components (CC) are extracted from the matrix form and maximum CC is computed and erased. Finally, the noise-free
image \( Y(i, j) \) is formed by comparing the differences center pixel/boundary with the center pixel \( \text{temp}(5) \). If the difference is greater than the center pixel value, then the matrix cell is replaced with the average value otherwise they are replaced with the median value. The interpretation of the information in the images requires the enhanced image form. Hence, the Gaussian model with the standard deviation (\( \sigma \)) and the image gradient (\( I_g \)) values are used to find the enhanced image (\( I_e \)) as follows:

\[
\sigma = \sqrt{\frac{1}{N+M} \sum_{i=1}^{M+N} \left( I_g(i) - \frac{2}{N} I_e \right)^2}
\]

\[
I_e = \frac{I_g}{\max(I_g)}
\]

Where, \( M, N \) – Row and column size of the image. Fig. 2 (a), (b) and (c) shows the input image, NCF filtered image and the enhanced image. The clear shape boundary description with the minimum energy consumption is the major requirement to track the dynamic objects.

![Fig. 2](image)

Fig. 2 (a) Input image, (b) Filtered image and (c) Enhanced image

The model that is helpful for clear boundary analysis is called active contour model. The segmentation accuracy depends on the convergence state of energy minimization that requires the optimal weight update.

### 3.2 Optimal Weight Update-Multi-level Set Segmentation

The important stage in proposed NTPMS is the multi-level segmentation process. The energy, weight, and direction updates are forms the active contour over the region. The algorithmic steps to perform the multi-level segmentation based on optimal weight as follows:

#### Multi-level set segmentation

**Input** – Enhanced image, ‘\( I_e \)’, ROI Mask, ‘\( R_{out} \)’, Texture Pattern, ‘ATP’

**Output** – Clustered Output, ‘\( I_e \)’

**Step 1: Initial Masking**

\( M = \sqrt{(R_{out}(x))^2 - (R_{out}(y))^2} - \sqrt{(1 - R_{out}(x))^2 - (1 - R_{out}(y))^2} + R_{out} - \frac{1}{2} \)

**Step 2: For iteration = 1 to N**

**Step 3: Idx = Index(M)**

**Step 4: Curvature**

\( \frac{\partial u}{\partial t} = \lambda \mu M(Idx) + k \)

**Step 5: Energy Update**

\( I_p = \begin{cases} M & \text{if } (M \leq 0) \\ 0 & \text{else} \end{cases} \)

\( E_p = \begin{cases} M & \text{if } (M > 0) \\ 0 & \text{else} \end{cases} \)

**Internal Energy**

\( \text{Energy, } I_E = \frac{(l_p(Idx) + \sum I_p)}{\max((l_p(Idx) + \sum I_p))} + \alpha * \frac{\partial u}{\partial t} \)

**Step 6: Difference in energy update**

\( dt = \frac{\lambda}{\max(I_E)} \)

**Step 7: Direction Update**

\( D_{Pos} = \sqrt{\max(ap^2, bn^2) + \max(cp^2, dp^2)} \)

**Step 8: Contour Weight Update**

\( \Omega_{i+1} = \frac{\Omega_i - dt * \sqrt{\frac{\Omega_i}{20} + 1}}{M + dt * I_E} \)

**Step 9: End for**

Initially, the Region of Interest (ROI) masking is the prior computation in the segmentation process. Then, the following processes are performed for each iteration.

- The index corresponds to the masked form is computed and based on the difference of pixel variations in eight directions with the spacing of \( 0^\circ, 30^\circ, 45^\circ, 60^\circ, 90^\circ, 120^\circ, 135^\circ, 180^\circ \) and the opposite values, the curvature (\( \frac{\partial u}{\partial t} \)) is computed for the images in step 4.
The energy of contour formation comprises the internal and external energy values. Based on the masked value (either > 0 or <=0), the internal and the external energy values are updated with the mask value in step 5:

$$I_p = \begin{cases} M & \text{if } (M \leq 0) \\ 0 & \text{else} \end{cases}$$ (3)

$$E_p = \begin{cases} M & \text{if } (M > 0) \\ 0 & \text{else} \end{cases}$$ (4)

- The optimized energy is computed by using the following formula in step 6.

$$I_e = \frac{(i_d(i_d)+\Sigma i)^a}{\max((i_d(i_d)+\Sigma i)^a)} + \alpha * \frac{\partial u}{\partial t}$$ (5)

- The existing energy level is updated with the computed maximum energy and the optimized new energy level has the great impact of weight update.

- The diverse directional coefficients (positive, negative, forward, backward, left and right-ap, an, bp, bn, cp, cn, dp, dn)-based directional update supports the clear image analysis in step 7.

- Finally, the contour weight is updated with the computed mask, energy and the updated energy level in step 8.

Fig. 3 (a) and (b) shows the binary mask image and the multi-level segmented output.

![Fig. 3 (a) Binary Mask image, (b) Multi-level set segmented output](image)

From the segmented output, the features extraction is the next stage to provide the unique form of the image. The proposed NTPMS utilizes the Gray Level Co-occurrence Matrix (GLCM) for the feature extraction process.

### 3.3 GLCM feature extraction

The matrix representation in which the number of rows and columns represent the gray levels of the clustered output image ($I_C$) is referred as Gray Level Co-occurrence Matrix (GLCM). The relative frequency variations of the pixels separated by the distance, intensity and the angle in the matrix element includes the statistical probability values $P(i, j|d, \theta)$. The increase in the number of pixel levels will increase the dimensionality of the GLCM features storage. Hence, the gray level reduction is preferred for storage minimization. Table 1 shows the features extracted from the GLCM for unique representation. For the number of neighboring gray levels (N), the statistical probability (p) values of two gray levels (a, b) and their mean ($\mu_a, \mu_b$) standard deviations ($\sigma_a, \sigma_b$) are used to estimate the texture features.

<table>
<thead>
<tr>
<th>S.no</th>
<th>Feature</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Contrast ($C_c$)</td>
<td>$C_c = \frac{\sum_{a=b}^{N} (a-b)^2 \cdot p(a,b)}{\sum_{a=b}^{N} p(a,b)}$</td>
</tr>
<tr>
<td>2</td>
<td>Correlation ($C_r$)</td>
<td>$C_r = \frac{\sum_{a,b} (a-\mu_a) \cdot (b-\mu_b) \cdot p(a,b)}{\sqrt{\sum_{a} (a-\mu_a)^2 \cdot p(a,b)} \cdot \sqrt{\sum_{b} (b-\mu_b)^2 \cdot p(b,b)}}$</td>
</tr>
<tr>
<td>3</td>
<td>Cluster Prominence ($C_P$)</td>
<td>$C_P = \frac{\sum_{a,b} (a+b) \cdot (a+b) \cdot p(a,b)}{\sum_{a} \sum_{b} p(a,b)}$</td>
</tr>
<tr>
<td>4</td>
<td>Cluster shade ($C_s$)</td>
<td>$C_s = \frac{\sum_{a,b} (a+b) \cdot (a+b) \cdot (a+b) \cdot p(a,b)}{\sum_{a} \sum_{b} p(a,b)}$</td>
</tr>
<tr>
<td>5</td>
<td>Dissimilarity (D)</td>
<td>$D = \frac{\sum_{a,b} (a+b) \cdot (a+b) \cdot p(a,b)}{\sum_{a} \sum_{b} p(a,b)}$</td>
</tr>
<tr>
<td>6</td>
<td>Energy (E)</td>
<td>$E = \frac{\sum_{a} (a-\mu_a)^2 \cdot p(a,b)}{\sum_{a} p(a,b)}$</td>
</tr>
<tr>
<td>7</td>
<td>Entropy (Ent)</td>
<td>$Ent = \sum_{a} \sum_{b} p(a,b) \cdot \log(p(a,b))$</td>
</tr>
<tr>
<td>8</td>
<td>Cluster homogeneity ($C_h$)</td>
<td>$C_h = \frac{\sum_{a,b} \sum_{i,j} (a+b) \cdot (a+b) \cdot p(a,b)}{\sum_{a} \sum_{b} p(a,b)}$</td>
</tr>
<tr>
<td>9</td>
<td>Overall homogeneity (Homop)</td>
<td>$Homop = \sum_{a} \sum_{b} \frac{1}{1+abs(a-b)}$</td>
</tr>
<tr>
<td>10</td>
<td>Maximum Probability ($P_{max}$)</td>
<td>$P_{max} = \max_{a,b} \sum_{i} \sum_{j} p(a,b)$</td>
</tr>
<tr>
<td>11</td>
<td>Variance (Var)</td>
<td>$Var = \frac{\sum_{a} \sum_{b} (a+b)^2 \cdot p(a,b) + \sum_{a} \sum_{b} (a+b)^2 \cdot p(a,b)}{\sum_{a} \sum_{b} p(a,b)}$</td>
</tr>
<tr>
<td>12</td>
<td>Auto Correlation (AC)</td>
<td>$AC = \frac{\sum_{a} \sum_{b} (a+b) \cdot (a+b) \cdot p(a,b)}{\sum_{a} \sum_{b} p(a,b)}$</td>
</tr>
<tr>
<td>13</td>
<td>Average Kurtosis (AK)</td>
<td>$AK = \frac{\sum_{a} \sum_{b} (a+b) \cdot (a+b) \cdot (a+b) \cdot p(a,b)}{\sum_{a} \sum_{b} p(a,b)}$</td>
</tr>
<tr>
<td>14</td>
<td>Average Skewness (AS)</td>
<td>$AS = \frac{\sum_{a} \sum_{b} (a+b) \cdot (a+b) \cdot (a+b) \cdot p(a,b)}{\sum_{a} \sum_{b} p(a,b)}$</td>
</tr>
</tbody>
</table>

The extracted features from GLCM computation have the great impact on the classification and the pattern recognition like either normal or abnormal.

### 3.4 Neural-Network Relevance Vector Machine

The dependency investigation of targets on the input requires the supervised learning approach. The supervised model includes the input vectors $\{x_n\}_{n=1}^{N}$ and associated target values $\{t_n\}_{n=1}^{N}$. The target values represents the class labels obtained from the classification process. The overfitting in training set due to the real data requires the suitable overlap analysis. The probabilistic sparse kernel that adopts Bayesian approach for learning the overfitting samples refers Relevance Vector Machine (RVM). The number of predictions from the RVM is based on the function described as $y(x) = \sum_{i=1}^{N} \omega_i K(x, x_n) + \omega_0$ (6) The function defined in (6) represents the relationship between the model weights ($\omega_i$) and kernel function $K(\ldots)$ in terms of input samples. The non-associated input samples with the non-zero weights are close to the decision boundary refers “relevance” vectors. The prediction of posterior
membership for the given class and the optimal solution are the objectives of the RVM. The logistic sigmoidal function generalizes the linear model and the corresponding likelihood is computed for the class instances (c) as

\[ P(c/w) = \prod_{i=1}^{n} \sigma(c(y(x_i)))^{c_i}[1-\sigma(c(y(x_i)))]^{1-c_i} \]  

(7)

Where, \( \sigma(y) \) is the logistic sigmoid function

\[ \sigma(y(x)) = \frac{1}{1+\exp(-y(x))} \]  

(8)

The most probable weights computation, iterative reweighted least square algorithm utilization and Gaussian approximation are repeated until the convergence criteria is satisfied. The iterative processes are modelled as the neural network training to provide the effective class labels. The features extracted from the GLCM are responsible for training of Neural Network (NN). The contour to be segmented is defined by the sequence of control points equal to the number of neurons in the network. The weights associated with each neuron denoted by the gain value. The membership function and repetitive weight update to classify the samples as either normal or abnormal.

NN-RVM


Output: Y – Labeled Output

Initialize map radius size, radius sample updating rate, radius decay rate of Neurons for cluster extraction.

i, j – Row and Column iteration of Feature Matrix ‘T’ respectively.

Step 1: Input\textsubscript{Neuron} = T; Particles = V;

Step 2: For t=1 to number of iteration

Step 3: Compute the summation

\[ \text{Sum}_{\text{Particle}} = \text{Sum}_{\text{Particle}} + \sqrt{(\text{map} - \text{Input}_{\text{Neuron}})^2}; \]

Step 4: Dis = \( \sqrt{\Delta x^2 + \Delta y^2}; \)

// Difference from original cluster size to extracted window size and calculate distance of \( \Delta x \) and \( \Delta y \)

Step 5: Update gain parameter for each sample gain = \( \text{update}_{\text{rate}} \times \exp \left(-\frac{\text{dis}}{2 \times \text{update}_{\text{radius}}} \right); \)

Step 6: Update map matrix according to gain. Where \( \text{wx}, \text{wy} - \) window size

\[ \text{mw}(x, y) = \text{mw}(x, y, a) + \text{gain} \times (\text{mw}(\text{wx}, \text{wy}) - \text{mw}(x, y)); \]

Step 7: Update radius sample rate, radius of Neuron cluster area

\[ \text{update}_{\text{radius}} = 1.0 + (\text{update}_{\text{radius}} - 1.0) \times \text{radius}_{\text{decay}}; \]

End loop

Step 8: dis\text{(j,k)} = \left( \left( \left( \text{sp}(j) - \text{sp}(k) \right)^2 \right)^{0.5} \right); \)

Step 9: \( \text{sp}_{\text{update}} = (\text{dis}_{\text{min}}(\text{sp}) + \text{sp}_{\text{previous}})/2; \)

// Update minimum distance level of particle and update particle

Step 10: dif \text{ (i)} = (x - \text{cent}(i))^2;

// Find the magnitude of the difference between particles and center of the window.

Step 11: new\text{Dist} = \text{minimum}{(\text{difference});

// Extract minimum difference and update that article as new fitness value in a cluster

Step 12: N=0; sum\text{y} = 0;

For length of particles,

\[ \text{sum}_{\text{y}} = \text{sum}_{\text{y}} + \text{new}_{\text{dist}}; \]

Step 13: If new\text{Dist} > 0,

\[ \text{N}=\text{N}+1; \]

End loop

Step 14: Update center as, cent\text{update} = \frac{\text{sum}_{\text{y}}}{\text{N}};

End loop

Step 15: If cent\text{update} = \text{cent}_{\text{old}}.

\[ Y = \text{Return Label}; \]

End if

The feature set from the GLCM and the testing vectors are extracted from the prior steps. The map reduces the size, sample update rate, and decaying rate are initialized as the first stage. The mean window size and the distance update from the step 3 to 11 offers the center update to the logistic function. The average center update value is equal to the old center, then the classified output is extracted from step 12 to 15. The output (Normal(class 1), Abnormal (class 2)) from multi-level set segmented form, GLCM extracted features offers the better severity level prediction in cervical images.

4 Performance Analysis

This section illustrates the performance validation of proposed Novel Texture Pattern-based Multi-level set Segmentation (NTPMS) regarding the sensitivity, specificity, accuracy, precision, and recall. Besides the comparative analysis of proposed NTPMS with the existing SVM and Neuro-fuzzy formulation on the coefficient metrics states the effectiveness in NTPMS in cervical image analysis.

Performance Metrics

The performance validation of the proposed NTPMS on the basic parameters and the comparative analysis of the existing SVM and the Neuro-Fuzzy formulation assures that the optimal weight update improves the segmentation accuracy. Table 2 presents the comparison of performance metrics for proposed NTPMS and the existing methods of SVM/Neuro-Fuzzy formulation.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>NTTPMS</th>
<th>SVM</th>
<th>Neuro Fuzzy</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP</td>
<td>105</td>
<td>108</td>
<td>52</td>
</tr>
<tr>
<td>TN</td>
<td>44</td>
<td>25</td>
<td>38</td>
</tr>
<tr>
<td>FP</td>
<td>4</td>
<td>23</td>
<td>86</td>
</tr>
<tr>
<td>FN</td>
<td>3</td>
<td>0</td>
<td>56</td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td>97.2222</td>
<td>100</td>
<td>48.1481</td>
</tr>
</tbody>
</table>
Fig. 4 shows the comparative analysis of sensitivity, specificity measures for NTPMS, SVM, Neuro-Fuzzy formulation. The optimal weight update by the multi-level set formulation and the novel texture patterns (NTP) improves both the sensitivity and specificity values. With Neuro-Fuzzy formulation, the NTPMS offers 49.07 and 69.41 % better in sensitivity and specificity values. Even though the SVM offers the highest sensitivity (100), the substantial reduction in specificity occurs in SVM. But, the NTPMS offers 88.24 % better specificity due to the clear image analysis.

Fig. 4 Sensitivity, specificity analysis for NTPMS, SVM and Neuro-fuzzy models

Fig. 5 shows the comparative analysis of accuracy, precision and recall values for NTPMS, SVM, and neuro-fuzzy formulation. The SVM offers 85.2564, 82.4427 and 100 % and the NTPMS offers 96.7949, 96.3303 and 97.2222 %. The optimal weight update unit by the ML formulation improves the accuracy and precision by 13.53 and 16.85 %.

Fig. 5 Accuracy, Precision and Recall analysis for NTPMS, SVM and Neuro-fuzzy models

Fig. 6 shows the comparative analysis of coefficient metrics (Jaccard, Dice, and Kappa) for proposed NTPMS and the existing SVM and neuro-fuzzy formulation. The SVM offers 0.8526, 0.9204 and 0.6008 and the NTPMS offers 0.9776, 0.9887 and 0.9904. The optimal weight update unit by the ML formulation improves the coefficient metrics by 14.66, 7.42 and 64.85 %.

Fig 6 Coefficient metrics analysis

5 Conclusion

This paper discussed the limitations in severity level prediction in cervical cancer images and the Manual Methods (MM) and their solution via computerized framework. Here, the nucleus texture pattern is extracted through the optimized cell segmentation techniques. The employment of multi-level set algorithm extracted the exact nucleus portion in the cervical images irrespective of the geometrical features. The periodical estimation of intensity difference values forms the cluster that provides the multi-label output. The optimal weight update based intensity weight estimation in proposed algorithm separated the layers which form the active contour over the image. The nucleus texture pattern extraction by the GLCM provided the clear image analysis and normal/abnormal classification. An accurate severity level prediction is achieved by using the proposed GLCM-OWU-ML combination compared to the geometrical feature extraction algorithms.

References


