Modified Segmentation Algorithm for the Cancer Affected White Blood Cells and its Classification

1Kalaiselvi Chinnathambi, 2Asokan Ramasamy and 3Premkumar Rajendran
1,2Department of Electronics and Communication Engineering, Kongu Engineering College, Erode, Tamilnadu, India. 3Department of Electronics and Communication Engineering, Kongu Engineering College, Trichy, Tamilnadu, India.

ARTICLE INFO
Article history:
Received 23 December 2013
Received in revised form 25 February 2014
Accepted 26 February 2014
Available online 15 March 2014

Keywords:
Active Contour, Level Set Method, Heaviside Function, Segmentation, Gaussian Kernel, Feature Extraction, Feature Selection, Classification.

ABSTRACT
The cancer cells are multiplicative in nature. Doctors face difficulties in counting the white blood cells (WBCs) at a particular stage due to crowding of cells. This paper proposes the robust segmentation algorithm that can reliably separate touching cells. Segmentation is the main important step in medical image processing. Precisely locating the area of interest in an image, in the presence of inherent uncertainty and ambiguity, is a challenging problem in medical imaging. Hence, one is often faced with a situation that demands proper segmentation. The algorithm is composed of two steps. It begins with a detecting and finding the cells in the region that utilizes level set algorithm. Next, the contour of big cell is obtained using modified level set active contour based on a piecewise smooth function. Feature extraction process follows Segmentation. The required information from the Geometry and Texture features were obtained. The purpose of feature selection is to reduce the dimensionality for the input of classifier. The best features are selected according to the maximal statistical dependency criterion based on mutual information. Because of the difficulty in directly implementing the maximal dependency condition, first derive an equivalent form, called minimal-Redundancy-Maximal-Relevance criterion (mRMR), for first-order incremental feature selection. The Feature Selection process is carried out by using Minimum-Redundancy And Maximum-Relevance (MRMR) technique. BPN is used as a classifier for the classification process. Back propagation provides a computationally efficient method for changing the weights in a feed forward networks, with differentiable activation function units, to learn a training set of inputs-outputs. The aim of this network is to train the net to achieve a balance between the ability to respond correctly to the input patterns that are used for training and the ability to provide good response to the input that are similar. Finally, the proposed algorithm is compared with several images which aids in applications such as locating and identifying the tumours and other pathologies.

INTRODUCTION

People affected by Cancer in India are increasing day by day. A status shows that around 25lakh cases have been reported so far and this also predicts that about 0.25% of new cases were reported each year and nearly 0.1% of this population die every year due to this deadly disease. This diagnosis causes emotional trauma and its treatment is too costly to bear (www.chillibreeze.com/articles_various/cancer-in-India).

Segmenting individual cells in blood cancer is usually the first step that is required in automatic image analysis. And it is a challenging problem due to the complex nature of the cells. Image segmentation is the process of building a partition of the image into connected regions, such that pixels of the region are homogenous according to some criterion (gray value, motion, etc). Automatically identifying and multiplying the cells by exploiting the shape and intensity characteristics of the cells was proposed by Dipti Prasad Mukherjee et al. An energy functional dependent upon the gradient magnitude along the cell boundary, the region homogeneity within the cell boundary and spatial overlap of the detected cells is minimized. Bjorn Nilsson at al uses the model based combinational optimization scheme to segment cluster based problems. They concentrate on clustering image not an overlapping region.

Each of these segmentation methods produced good results on regions exhibiting little or no cell crowding; however, they often failed to separate touching cells accurately. The watershed family of algorithms has become one of the most commonly used segmentation methods to address the challenge of touching cells. However, the primary limitation of the watershed approaches is that they often result in over segmentation. Some algorithms
such as marker controlled watershed (Mandeep Kaur, 2011), Otsu method (Hima Bindu. Ch, 2009), rule based strategies (Wahbey, C. et al., 2002; Lin, G., Yu, W., H. Lee et al, 2009) were developed to address this problem. When the intensity of overlapping regions is brighter (or darker) than the non overlapping regions within individual cells, a set of false seeds will be created in the overlapping regions (Pravin B., et al, 2007). This is not surprising because the voting schema in (Pravin B., et al, 2007), is biased toward the boundary of the object. The edges of overlapping regions contribute to the creation of false seeds within the overlapping regions. The significant improvement of the new algorithm that applies a shifted Gaussian kernel (Kaiksha Zhang et al, 2009) and mean shift onto single-pass voting to generate more accurate and quicker seed detection was proposed later.

When the input data to an algorithm is too large to be processed and it is suspected to be notoriously redundant (e.g. the same measurement in both feet and meters) then the input data will be transformed into a reduced representation set of features (Pradeep, N. et al, 2012). Transforming the input data into the set of features is called feature extraction (Manavalan Radhakrishnan, 2012). If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input.

In machine learning and statistics, feature selection, also known as variable selection, attribute selection or variable subset selection, is the process of selecting a subset of relevant features for use in model construction. The central assumption when using a feature selection technique is that the data contains many redundant or irrelevant features. Redundant features are those which provide no more information than the currently selected features, and irrelevant features provide no useful information in any context. Feature selection techniques (Hanchuan Peng et al, 2005) are a subset of the more general field of feature extraction.

The rest of the paper is organized as follows. The introduction to image segmentation and the related methods used in this work is described in section 2 and also we formulated the derivation of level set in this section. The modified level set algorithm is described in section 3. In section 4 and 5 deals with feature extraction and feature selection processes. Then section 6, we validate our method by various experiments on cancer and normal microscopic blood smear images.

Image Segmentation:
In image processing and computer vision, partitioning an image into integrated regions is called segmentation (Jayaram, S., 2009). Getting a more meaningful and analyzable image is the main aim of segmentation (Bjorn Nilsson et al, 2005). Apart from this, locating boundaries is the main objective of segmentation.

More exactly, segmentation deals with assigning an unique label to each pixel in the object of interest such that they show a common characteristics. This reduces the pixel data to region based information.

Need for Segmentation:
In image processing it can be number of pixels with the same intensity in general. Segmentation is to separate the homogeneous area (Xin Qi et al, 2012). The analysis of blood slides is a powerful tool in determining the health status of an individual and could detect several diseases. The count and shape, lineage and maturity level of white and red blood cells could aid in the diagnosis of diseases that range from inflammatory to leukemia. Many automated techniques were proposed to overcome the tedious and time consuming task of human experts in counting and classifying white blood cells. Various techniques were used for the segmentation stage including mean shift algorithm, histogram equalization, thresholding, watershed algorithm.

Level Set Method:
The level set method (LSM) is a numerical technique for tracking interfaces and shapes. The advantage of the level set method is that, it doesn’t need variable parameters (Dipti Prasad Mukherjee et al, 2004). Also, the level set method easily keeps track of topological changes. When an object of interest merges or splits. The advantage of level set methods are implicit, parameter free method and provides a direct way to estimate the geometric properties of evolving structure and the applications include edge extraction, code tracking and contour tracking.

Consider the most general case the following form of curve propagation is

\[ C(p,t) = F(k) \times N \] (1)

where C is a closed curve propagation

F(k) is a force and N is normal to curve/surface

The level set method represents the curve in the form of an implicit surface is

\[ \emptyset(x, y, t) : R^2 x (0, t) \rightarrow R \] (2)
The level set method accounts to representing a closed curve using an auxiliary function $\phi$ called the level set function. This is derived from the initial contour according to the following condition

$$C(p, 0) = \{(x, y) : \phi(x, y, 0) = 0\} \quad (3)$$

and the level set method manipulates $C$ implicitly, through the function $\phi$. $\phi$ is assumed to take positive values inside the region and negative values outside the region which is determined by the curve propagation $C$.

The level set flow can be re-written in the following form

$$\phi_t + H(\phi_x, \phi_y) = 0 \quad (4)$$

where $H$ is a Hamiltonian.

**Mathematics of Level Set Method:**

Consider the image that has $N$ cells. Assume that the image $u_0$ is formed by two regions of approximatively piecewise-constant intensities, of distinct values $u_{oi}$ and $u_{o}$. Assume further that the object to be detected is represented by the region with the value $u_{oi}$. Let denote its boundary by $C_0$. Then $u_{i} \approx u_{oi}$ we have inside the object [or inside ($C_0$)], and $u_{o} \approx u_{o}$ outside the object [or outside ($C_0$)]. Now let us consider the following “fitting” term or level set energy functional term:

$$F_1(C) + F_2(C) = \int_{\text{inside} (C)} |I - c_1|^2 \, dx \, dy + \int_{\text{outside} (C)} |I - c_2|^2 \, dx \, dy \quad (5)$$

Where $c_1$ and $c_2$ are constants depending on $C$ are the averages of $u_o$ inside $C$ and respectively outside $C$. The energy term can be written as the sum of two fitting terms. In this simple case, it is obvious that, the boundary of the object, is the minimize of the fitting term

$$\min_{C} F_1(C) + F_2(C) = 0 \approx F_1(C_0) + F_2(C_0) \quad (6)$$

The constants $c_1$ and $c_2$ can be denoted as

$$C_1 = \int H(x,y)K(x,y)\,dx \, dy \quad (7)$$

$$C_2 = \int (1 - H(x,y))K(x,y)\,dx \, dy \quad (8)$$

where $H(x,y)$ is Heaviside function which will be discussed in next section and $K(x,y)$ is defined as 2D Gaussian kernel function:

$$K = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \quad (9)$$

$\sigma$ is standard deviation.

For instance, if the curve $C$ is outside the object, then $F_1(C) > 0$ and $F_2(C) \approx 0$. If the curve is inside the object, then $F_1(C) = 0$ but $F_2(C) > 0$. If the curve is both inside and outside the object, then $F_1(C) > 0$ and $F_2(C) > 0$. Finally, the fitting energy is minimized if $C=C_0$, i.e., if the curve is on the boundary of the object. These basic remarks are illustrated in Figure 1.

In active contour model we will minimize the above fitting term and we will add some regularizing terms, like the length of the curve $C$, and or the area of the region inside $C$ proposed by M.Kass et al, which is based on the Mumford-Shah model (Chan, T., L. Vese, 2001). Therefore, we introduce the energy functional $F(c_1, c_2, C)$, defined by

$$F(c_1, c_2, C) = \mu \ast \text{length}(C) + \nu \ast \text{Area (inside}(C)) + \lambda_1\int_{\text{inside} (C)} |I - c_1|^2 \, dx \, dy + \lambda_2\int_{\text{outside} (C)} |I - c_2|^2 \, dx \, dy \quad (10)$$

$F_1(C) > 0, F_2(C) = 0, \text{Fitting} > 0; \ F_1(C) = 0, F_2(C) > 0, \text{Fitting} > 0; \ F_1(C) > 0, F_2(C) > 0, \text{Fitting} > 0; \ F_1(C) = 0, F_2(C) = 0, \text{Fitting} = 0$

![Fig. 1: Boundary condition.](image)
µ ≥ 0, v ≥ 0, λ₁, λ₂ > 0 are fixed parameters. In numerical calculations, λ₁ = λ₂ = 1 and v = 0. If this value is small enough, then it segments smaller objects otherwise it segments larger objects in the cell region. Therefore, consider the minimization problem is

\[ \inf_{c_1, c_2, C} F(c_1, c_2, C) \]  

(11)

In the isoperimetric inequality the length is comparable with area is given by

\[ \text{Area (inside(C))} \leq c^* \text{length (C)} \]  

(12)

where c is a constant.

**Relation with the Mumford - Shah Function:**

The Mumford–Shah functional [15] for segmentation is

\[ F_{MS}(u, C) = \mu \ast \text{length}(C) + \lambda \int_{\Omega} |I - u(x, y)|^2 \, dx \, dy + \int_{\Omega/C} |\nabla u(x, y)|^2 \, dx \, dy \]  

(13)

where \( \mu \) and \( \lambda \) are positive parameters. A reduced form of this problem is simply the restriction of \( F_{MS} \) to piecewise constant functions \( u \), i.e., \( u = \text{constant} \, c_i \) on each connected component of \( \Omega/C \). Therefore, as it was pointed out by D. Mumford and J. Shah, \( c_i = \text{average} (u_o) \) on each connected component. The reduced case is called the minimal partition problem.

The active contour model with \( v=0 \) and \( \lambda_1 = \lambda_2 = \lambda \) is a particular case of the minimal partition problem, in which we look for the best approximation \( u \) of \( u_o \), as a function taking only two values, namely

\[ u = \begin{cases} \text{average} (u_o) \text{ inside } C \\ \text{average} (u_o) \text{ outside } C \end{cases} \]  

(14)

and with one edge \( C \), represented by the snake or the active contour.

**Level Set Formulation:**

In the level set method (Mumford, D., J. Shah, 1989; Zhu, G.P. *et al.*, 2007), the closed curve \( C \), inside\( (C) \) and outside\( (C) \) which is represented by the zero level set of a Lipschitz function as in figure 2.

\[ C = \partial \omega = \{ (x,y) \in \Omega : \emptyset (x,y) = 0 \} \]

Inside \( (C) = \omega = \{ (x,y) \in \Omega : \emptyset (x,y) > 0 \} \)  

(15)

Outside \( (C) = \Omega / \partial C = \{ (x,y) \in \Omega : \emptyset (x,y) < 0 \} \)

For the level set formulation of our variational active contour model, we replace the unknown variable \( C \) by the unknown variable \( \phi \).

![Fig. 2: Lipschitz Function.](image)

Using the Heaviside function \( H \), and the one-dimensional Dirac measure \( \delta_o \), and defined, respectively, by

\[ H(z) = \begin{cases} 1, & z \geq \varepsilon \\ 0, & z \leq -\varepsilon \\ \frac{1}{2} \left[ 1 + \frac{z}{\varepsilon} + \frac{1}{\pi} \sin \left( \frac{\pi z}{\varepsilon} \right) \right], & |z| \leq \varepsilon \end{cases} \]  

(16)

Differentiate Heaviside function with respect to \( z \), we get

\[ \frac{dH}{dz} = \lim_{\Delta z \to 0} \frac{\Delta H}{\Delta z} \]  

(17)

where \( \Delta H/\Delta z \) is the slope of the closed curve. \( H \) should be a flat line in regions, therefore slope becomes zero. Then \( dH/dz \) is zero for \( z > \varepsilon \) and \( z < -\varepsilon \).
If \( z = \in \), \( z_+ = -a/2 \), \( z_- = +a/2 \), then slope gives
\[
\frac{dH}{dz} = \lim_{a \to 0} \frac{1}{a}
\]
Equation (18)

The delta function,
\[
\delta(z) = \frac{d}{dz} \left\{ \begin{array}{ll}
0, & z < \in \\
\infty, & z = \in \\
0, & z > \in
\end{array} \right.
\]
Equation (20)

The energy function can be minimized by iteratively employing the gradient descent method.
\[
\frac{\delta E}{\delta \theta} = \lim_{\varepsilon \to 0} \frac{d}{d\varepsilon} \left( \frac{1}{2} \int_{\Omega} |1 - c_1 H_\theta(\theta) - c_2 (1 - H_\theta(\theta))|^2 dx \right)
\]
Equation (21)

By this method, all the edges are detected. But some of the smaller cells are not detected. To detect those smaller areas, Euler Lagrange equation is employed. This equation acts locally to separate smaller objects. Finally, all the edges are detected by level set formulation.

**Modified Level Set algorithm:**

We modified the level set algorithm by incorporating the active contour into level set. We have applied the level set only within the region. This is achieved by the convolution of piecewise smooth function (active contour) and Heaviside function (level set). We incorporate this modification in step 4 which will discuss in later section.

We modified the level set algorithm by incorporating the active contour into level set. We have applied the level set only within the region. This is achieved by the convolution of piecewise smooth function (active contour) and Heaviside function (level set). We incorporate this modification in step 4 which will discuss in later section.

**Active Contour:**

Active contour model (Kass, M. et al, 1988), delineates the area of interest outline from a noisy image. The objective is energy function minimization i.e. Energy function of the present contour is represented (breakup into) as it is a summarization of external and internal energy. This model is also known as snake model.

Internal energy is defined within the contour itself to maintain the contour smooth and external energy is computed from the image data to move the contour toward an object boundary. When the external and internal energy becomes equal, the energy attains equilibrium and contour stabilizes.

Internal Energy of the snake is
\[
E_{\text{internal}} = E_{\text{cont}} + E_{\text{curv}}
\]
Equation (22)

where \( E_{\text{cont}} \) is the energy of the snake contour
\( E_{\text{curv}} \) is the energy of the spline curvature.

\[
E_{\text{internal}} = (\alpha(s) |v(s)|^2 + \beta(s) |v_n|^2) / 2
\]
\[
= (\alpha(s) \frac{dv}{ds} (s) |s|^2 + \beta(s) \| \frac{d^2v}{ds^2} (s) \|^2) / 2
\]
Equation (23)

The first-order term makes the snake act like a membrane and second-order term makes it act like a thin plate. Large values of \( \alpha(s) \) will increase the internal energy of the snake as it stretches more and more, whereas small values of \( \alpha(s) \) will make the energy function insensitive to the amount of stretch. Similarly, large values of \( \beta(s) \) will increase the internal energy of the snake as it develops more curves, whereas small values of \( \beta(s) \) will make the energy function insensitive to curves in the snake. Smaller values of both \( \alpha(s) \) and \( \beta(s) \) will place fewer constraints on the size and shape of the snake.

The 2D active uses the point distribution model to minimize the shape range to a well known domain as a result of training. The advantages over classical techniques are such that snakes are autonomous and self-adapting in their search for a minimal energy state. They can be easily manipulated using external image forces. They can be made sensitive to image scale by incorporating Gaussian smoothing in the image energy function.
They can be used to track dynamic objects in temporal as well as the spatial dimensions (Shi, Y., W.C. Karl, 2005).

This model is used to segment white blood cell areas by embedding the local image information. The energy functional area is given by

$$E(C, f_1, f_2) = \lambda_1 \int \int_{\Omega_{\text{inside}}} K_{\sigma}(x - y) ||I - f_1||^2 dy \, dx + \lambda_2 \int \int_{\Omega_{\text{outside}}} K_{\sigma}(x - y) ||I - f_2||^2 dy \, dx$$  \hspace{1cm} (24)$$

Where $\lambda_1$ and $\lambda_2 > 0$ are fixed parameters, $K_{\sigma}$ is Gaussian kernel with standard deviation $\sigma$, $f_1$ and $f_2$ are two smooth functions that approximate the local image intensities inside and outside of contour $C$ respectively.

$$f_1(x) = \frac{K_{\sigma} \ast [H_{\epsilon}(\emptyset)I(x)]}{K_{\sigma} \ast H_{\epsilon}(\emptyset)}$$  \hspace{1cm} (25)$$

$$f_2(x) = \frac{K_{\sigma} \ast [(1 - H_{\epsilon}(\emptyset))I(x)]}{K_{\sigma} \ast (1 - H_{\epsilon}(\emptyset))}$$  \hspace{1cm} (26)$$

The standard deviation $\sigma$ plays an important role in practical applications and this value varies for several images.

In the above equation, the regularized parameter of Heaviside function $H$ and dirac function $\partial$ as follows:

$$H_{\epsilon}(z) = \frac{1}{2} \left[ 1 + \frac{2}{\pi} \arctan \left( \frac{z}{\epsilon} \right) \right]$$  \hspace{1cm} (27)$$

$$\partial_{\epsilon}(z) = \frac{1}{\pi \epsilon^2} \frac{\epsilon}{x^2 + \epsilon^2}$$  \hspace{1cm} (28)$$

The flow of level set using ACM is as follows:

**STEP 1:** Initialize the parameters: $\epsilon$, $\sigma$, $\sigma_{\phi}$.  
**STEP 2:** Find distance between the center and radius.  
**STEP 3:** Apply Heaviside function and Delta function.  
**STEP 4:** Apply piecewise smooth function to energy functional terms.  
**STEP 5:** Start the iteration value.  
**STEP 6:** Obtain the convolution of distance and $K$-Phi.  
**STEP 7:** Update the iteration value in order to segment the interested region.

**Feature Extraction:**

In image processing, feature extraction is a special form of dimensionality reduction. When the input data to an algorithm is too large to be processed, then the input data will be transformed into a reduced representation set of features. Transforming the input data into the set of features is called feature extraction. Feature extraction, is used to extract the individual cells from the input image. The cell feature extraction is based on four main groups.

They are:

- Textural
- Geometrical
- Statistical
- Morphological

**Textural:**

The textural features (Selvarajah, S. and S.R. Kodituwakku, 2011) reflect the statistical arrangement of the pixels in the image. The sum and difference histograms of the gray levels of the neighboring pixels are analyzed at different directions, and based on such analysis, the mean value, angular second momentum, contrast, and entropy are treated as the features. They are determined for the red, green, and blue colors, independent of the nucleus and cytoplasm. We have applied the Unser method of textural features, generating them for the original image and the image of reduced resolution.

**Geometrical:**

The geometrical features describe different aspects of the geometry of the cell and use parameters describing the area, radius, perimeter, symmetry (the difference between lines that are perpendicular to the major axis to the cell boundary in both directions), compactness (perimeter^2/area), concavity (the extent to which the actual boundary of a cell lies on the inside of each chord between nonadjacent boundary points), the length of the major and minor axes, etc. Up to 19 geometrical features have been created this way. No color information is used here.
Statistical:
The statistical parameters refer to the color distribution contained in the cell image. We have created the features based on the histograms (Manavalan Radhakrishnan, 2012) of the image matrix and the gradient matrix of the image for three color components red, green, and blue. The mean value, variance, skewness, and kurtosis of both histograms are used as the features (Pradeep, N. et al, 2012).

Morphological:
The last set of features refers to the results of the morphological operations, such as erosion and dilation, performed a few times on the cell image. Such features include the relative size and number of cells before and after applying these morphological operations.

Feature Selection:
Feature selection techniques are often used in domains where there are many features and comparatively few samples. Variable selection is a method for feature selection, model construction uses this process for relevant feature selection. Redundant features are discarded as they provide no information. Irrelevant features provide no useful information in any context.

Minimum - Redundancy and Maximum - Relevance (MRMR):
Peng et al. proposed an MRMR feature-selection method (Hanchuan Peng et al, 2005) which uses mutual information, correlation, distance/similarity scores for feature selection. With mutual information, relevant features and redundant features are considered simultaneously. For the class c, the relevance of a feature set S is defined by the average of all mutual information between the individual feature xi and the class c as follows:

$$D = \frac{1}{|S|} \sum_{x_i \in S} I(x_i; c)$$  (29)

The redundancy of all features in the set S is the average value of all mutual information values between the feature xi and the feature x

$$R = \frac{1}{|S|^2} \sum_{x_i, x_j \in S} I(x_i; x_j)$$  (30)

The MRMR criterion is the difference between the above as follows:

$$mRMR = \max_S \left[ \frac{1}{|S|} \sum_{x_i \in S} I(x_i; c) - \frac{1}{|S|^2} \sum_{x_i, x_j \in S} I(x_i; x_j) \right]$$  (31)

In the globally optimal feature set there are n full-set features, let f_i be the set indicator function for feature x_i so that f_i = 1 indicates presence and f_i = 0 indicates absence of the feature x_i. Let c_i = I(x_i; c) and a_ij = I(x_i; x_j). The optimization problem can be defined as:

$$mRMR = \max_{f \in \{0,1\}^n} \left[ \frac{\sum_{i=1}^{n} c_i f_i}{\sum_{i=1}^{n} f_i} - \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} f_i f_j}{\sum_{i=1}^{n} f_i} \right]$$  (32)

The above equation shows that MRMR feature selection is an approximation for optimal maximum-dependency feature selection. MRMR maximizes the mutual information between the joint distribution of the selected features and the classification variable. Overall the algorithm is more efficient (in terms of the amount of data required) than the theoretically optimal max-dependency selection, yet produces a low redundancy feature set.

Classification:
Introduction to BPN:
A well known supervised learning method for training artificial neural network is Back propagation (Sivanandam, S.N., 2006), an abbreviation for "backward propagation of errors". A training set is made up of many inputs to get the desired output, uses the generalized delta rule. The goal of any supervised learning algorithm is to find a function that best maps a set of inputs to its correct output. A single-layer neural network however, learns a function that outputs a label solely using the intensity of the pixels in the image. The goal and motivation for developing the back propagation algorithm is to find a way to train multi-layered neural networks such that it can learn the appropriate internal representations to allow it to learn any arbitrary mapping of input to output.

Architecture of BPN:
An elementary neuron with R inputs is shown below, with each input weighted an appropriate w. The transfer function f is formed by the sum of the weighted inputs and the bias. Neurons can use any differentiable transfer function f to generate their output.
Figure 3 shows the standard characteristics of back propagation networks. Perceptrons are arranged in layers, with the first layer taking in inputs and the last layer producing outputs. The middle layers have no connection with the external world, and hence are called hidden layers. Each perceptron in one layer is connected to every perceptron on the next layer. Hence information is constantly “fed forward” from one layer to the next.

![Back Propagation Network Diagram](image)

**Fig. 3: A Back Propagation network.**

**Learning as an optimization problem:**

A simple neuron is considered with two input units, one output unit and no hidden units. Each neuron uses a linear output that is the weighted sum of its input. Figure 4 shows the simple neural network with two input units and one output unit.

Initially, before training, the weights will be set to random. Then the neuron learns from training examples, which in this case consists of a set of tuples \((x_1, x_2, t)\) where \(x_1\) and \(x_2\) are the inputs to the network and \(t\) is the correct output. Since the weights are initially random, the network given \(x_1\) and \(x_2\) will compute an output \(y\) which very likely differs from \(t\). A common method for measuring the discrepancy between the expected output \(t\) and the actual output \(y\) is using the squared error measure:

\[
E = (t - y)^2
\]  

where \(E\) is the discrepancy or error.

The back propagation algorithm aims to find the set of weights that minimizes the non-linear error using Gradient Descent error.

![Simple Neural Network Diagram](image)

**Fig. 4: A simple neural network with two input units and one output unit.**

**Procedure of BPN:**

1. **Initialization of weights**
   
   During this step, some random values are assigned.

2. **Feed Forward Networks**
During this stage, each input receives an input signal and transmits this signal to each of the hidden units. Each hidden unit then calculates the activation function and sends its signal to each output unit. The output unit calculates the activation function to form the response of the net for the given input pattern.

Step 3: Back Propagation of errors.
During this step, each output unit compares its computed activation function with target value to determine the associated error for that pattern with the unit.

Step 4: Updating of weights and biases.
During final stage, weights and biases are updated using delta rule and activation.

RESULTS AND DISCUSSION

Microscopic blood smear images and the cancer affected blood images were taken to validate the algorithm. Normal blood cells and cancer affected blood cells were segmented individually using modified level set algorithm and the parameters were calculated and listed in Table 1.

Figure 5 shows that input images which contains Red Blood Cells (RBC)&WBC. Image C&D is affected by cancer and remaining cells normal blood cells. Cancer cells are cells that grow and divide at an unregulated, quickened pace. However, more recent research has that the failure to recognize cancer cells is caused by the lack of particular co-stimulated molecules.

![Input Images](image1)

**Fig. 5: Input Images.**

![Level set algorithm result without adding active contour](image2)

**Fig. 6: Level set algorithm result without adding active contour.**

That aid in the way antigens react with lymphocytes (WBC). WBC is thought to use a dual receptor system when they determine whether or not to kill human cells. Figure 6 shows that the existing algorithm results in segmenting both RBC and WBC. For image segmentation, we have to find the radius of the cell region. Figure 7 show the radius for various input images.

**Modified Level Set algorithm:**

![Radius of the images](image3)

**Fig. 7: Radius of the images.**
After finding the radius, it starts segmenting the white blood cells. The parameters to obtain the optimum result are as follows in Table 2. This parameter will be tuned properly for various images as shown in figure 8.

**Parameter Validation:**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Image A</th>
<th>Image B</th>
<th>Image C</th>
<th>Image D</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>1.3</td>
<td>0.5</td>
<td>1</td>
<td>1.4</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Radius, $r$</td>
<td>30</td>
<td>25</td>
<td>40</td>
<td>45</td>
</tr>
<tr>
<td>$\Sigma$</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

**Fig. 8:** Parameter Validation for our method.

Here, for different values of $\epsilon$ (epsilon) and $\sigma$ (sigma) segmentation is done to normal images. From that we can examine that segmentation occurs best for $\epsilon = 5$ and $\sigma = 20$. So we fix these values for the remaining segmentation evaluation.

**Table 1:** Parameter Validation for Input Images

<table>
<thead>
<tr>
<th>Measure</th>
<th>Image A</th>
<th>Image B</th>
<th>Image C</th>
<th>Image D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iteration</td>
<td>3300</td>
<td>2400</td>
<td>4000</td>
<td>4000</td>
</tr>
<tr>
<td>Mean</td>
<td>35.2128</td>
<td>29.1133</td>
<td>1.3009</td>
<td>1.6395</td>
</tr>
<tr>
<td>Variance</td>
<td>241.3513</td>
<td>193.9551</td>
<td>9.1203</td>
<td>8.3557</td>
</tr>
<tr>
<td>Output Value</td>
<td>57.5403</td>
<td>51.1224</td>
<td>10.2471</td>
<td>10.5306</td>
</tr>
<tr>
<td>Heaviside Function</td>
<td>0.9724</td>
<td>0.9938</td>
<td>8.8555</td>
<td>0.8589</td>
</tr>
</tbody>
</table>

The segmented result for the constant values of $\epsilon$, $\sigma$ is as follows in figure 9.

**Fig. 9:** Segmented Result for our method.

The parameters will vary for different images. The parameter $\sigma_{\phi}$ plays an important role in segmentation and it must be tuned very properly.

**Table 2:** Values for Various Bloods Smear Images.
The following Table 3 illustrates feature extracted values from the segmented image. The feature selection was carried out by MRMR method and the values are shown in Table 4 from extracted value.

**Table 3:** Feature Extracted Values for Various Images.

<table>
<thead>
<tr>
<th>FEATURES</th>
<th>IMAGE-1</th>
<th>IMAGE-2</th>
<th>IMAGE-3</th>
<th>IMAGE-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEAN</td>
<td>0.9387</td>
<td>1.7905</td>
<td>31.6939</td>
<td>33.6423</td>
</tr>
<tr>
<td>SD</td>
<td>2.6893</td>
<td>3.2521</td>
<td>12.2643</td>
<td>14.0738</td>
</tr>
<tr>
<td>VARIANCE</td>
<td>10.5763</td>
<td>150.4125</td>
<td>198.0711</td>
<td>195.3266</td>
</tr>
<tr>
<td>ENTROPY</td>
<td>1.2557</td>
<td>5.1732</td>
<td>5.0565</td>
<td>5.0565</td>
</tr>
<tr>
<td>CORRELATION</td>
<td>-0.8279</td>
<td>-0.8333</td>
<td>0.3441</td>
<td>0.3441</td>
</tr>
<tr>
<td>AREA</td>
<td>644 - 1315</td>
<td>1805 - 1520</td>
<td>18730</td>
<td>21227</td>
</tr>
<tr>
<td>CENTROID</td>
<td>20 - 50</td>
<td>4 - 135.5</td>
<td>76.3 - 66</td>
<td>82.6 - 67.3</td>
</tr>
<tr>
<td>PERIMETER</td>
<td>77.8 - 87.1</td>
<td>180.5 - 228.40</td>
<td>562</td>
<td>596</td>
</tr>
<tr>
<td>MAL</td>
<td>31.9 - 59.6</td>
<td>59.8 - 65.6</td>
<td>180.7161</td>
<td>195.3266</td>
</tr>
<tr>
<td>MIL</td>
<td>25.9 - 30.2</td>
<td>40.5 - 0.3</td>
<td>155.5882</td>
<td>157.6771</td>
</tr>
<tr>
<td>SOLIDITY</td>
<td>0.97 - 0.87</td>
<td>0.9 - 0.7</td>
<td>0.9406</td>
<td>0.9543</td>
</tr>
<tr>
<td>ECCENTRICITY</td>
<td>0.58 - 0.86</td>
<td>0.7 - 0.7</td>
<td>0.5087</td>
<td>0.5902</td>
</tr>
<tr>
<td>EQUIVOLUMETER</td>
<td>28.6 - 40.9</td>
<td>47.93 - 3.99</td>
<td>154.4273</td>
<td>164.3991</td>
</tr>
<tr>
<td>EXTREMA</td>
<td>0.7880</td>
<td>0.4998</td>
<td>0.7452</td>
<td>0.7509</td>
</tr>
<tr>
<td>SRE</td>
<td>0</td>
<td>0</td>
<td>0.0016</td>
<td>0.0016</td>
</tr>
<tr>
<td>LRE</td>
<td>0</td>
<td>0</td>
<td>877952</td>
<td>1184908</td>
</tr>
<tr>
<td>GLNU</td>
<td>0</td>
<td>0</td>
<td>1823.7</td>
<td>2018.1</td>
</tr>
<tr>
<td>RP</td>
<td>0.0083</td>
<td>0.0076</td>
<td>0.0076</td>
<td>0.0075</td>
</tr>
<tr>
<td>LGRE</td>
<td>0</td>
<td>0</td>
<td>0.4955</td>
<td>0.4932</td>
</tr>
<tr>
<td>HGLRE</td>
<td>0</td>
<td>0</td>
<td>652118</td>
<td>772108</td>
</tr>
<tr>
<td>SRLGGE</td>
<td>0</td>
<td>0</td>
<td>0.6316</td>
<td>0.6265</td>
</tr>
<tr>
<td>SRHGLE</td>
<td>8.9833</td>
<td>78.1761</td>
<td>350.2303</td>
<td>328.3554</td>
</tr>
<tr>
<td>LRLGGE</td>
<td>1.4333</td>
<td>0</td>
<td>406.8750</td>
<td>407.3916</td>
</tr>
<tr>
<td>HRGGE</td>
<td>4162.9</td>
<td>5332.5</td>
<td>31419</td>
<td>32606</td>
</tr>
<tr>
<td>CONTRAST</td>
<td>0</td>
<td>0</td>
<td>62</td>
<td>81</td>
</tr>
<tr>
<td>DISSIMILARITY</td>
<td>0</td>
<td>0</td>
<td>62</td>
<td>81</td>
</tr>
<tr>
<td>HOMOGENITY</td>
<td>0</td>
<td>0</td>
<td>17053</td>
<td>18777</td>
</tr>
<tr>
<td>IDM</td>
<td>0</td>
<td>0</td>
<td>62</td>
<td>81</td>
</tr>
<tr>
<td>GLCM MEAN</td>
<td>223,1250</td>
<td>323,4063</td>
<td>309,0781</td>
<td>345,4688</td>
</tr>
<tr>
<td>GLCM SD</td>
<td>1785</td>
<td>2587.3</td>
<td>2105.0</td>
<td>2317</td>
</tr>
<tr>
<td>SSD</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SAD</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 3 abbreviations are Standard Deviation(SD), Major Axis Length(MAL), Minor Axis Length(MIL), Short Run Emphasis(SRE), Long Run Emphasis(LRE), Gray Level Non-Uniformity(GLNU), Run Percentage(RP), Low Gray-Level Run Emphasis(LGRE), High Gray-Level Run Emphasis(HGRE), Short Run Low Gray-Level Emphasis(SRLGGE), Short Run High Gray-Level Emphasis(SRHGLE), Long Run Low Gray-Level Emphasis(LRLGGE), Long Run High Gray-Level Emphasis (LRGGE), Inverse Difference Moment(IDM), Gray-Level Co-occurrence Matrix Mean(GLCM MEAN), Gray-Level Co-occurrence Matrix Standard Deviation (GLCM SD), Sum of Squared Difference(SSD), Sum of Absolute Difference(SAD).

**Table 4:** Feature Selected Values for Various Image.

<table>
<thead>
<tr>
<th>BEST FEATURES</th>
<th>IMAGE-1</th>
<th>IMAGE-2</th>
<th>IMAGE-3</th>
<th>IMAGE-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEAN</td>
<td>0.9387</td>
<td>1.7905</td>
<td>31.6939</td>
<td>33.6423</td>
</tr>
<tr>
<td>SD</td>
<td>2.6893</td>
<td>3.2521</td>
<td>12.2643</td>
<td>14.0738</td>
</tr>
<tr>
<td>VARIANCE</td>
<td>10.5763</td>
<td>150.4125</td>
<td>198.0711</td>
<td>195.3266</td>
</tr>
<tr>
<td>ENTROPY</td>
<td>1.2557</td>
<td>5.1732</td>
<td>5.0565</td>
<td>5.0565</td>
</tr>
<tr>
<td>CORRELATION</td>
<td>-0.8279</td>
<td>-0.8333</td>
<td>0.3441</td>
<td>0.3441</td>
</tr>
<tr>
<td>AREA</td>
<td>644 - 1315</td>
<td>1805 - 1520</td>
<td>18730</td>
<td>21227</td>
</tr>
<tr>
<td>CENTROID</td>
<td>20 - 50</td>
<td>4 - 135.5</td>
<td>76.3 - 66</td>
<td>82.6 - 67.3</td>
</tr>
<tr>
<td>PERIMETER</td>
<td>77.8 - 87.1</td>
<td>180.5 - 228.40</td>
<td>562</td>
<td>596</td>
</tr>
<tr>
<td>MAL</td>
<td>31.9 - 59.6</td>
<td>59.8 - 65.6</td>
<td>180.7161</td>
<td>195.3266</td>
</tr>
<tr>
<td>MIL</td>
<td>25.9 - 30.2</td>
<td>40.5 - 0.3</td>
<td>155.5882</td>
<td>157.6771</td>
</tr>
</tbody>
</table>

**BPN result:**

Features are extracted from the segmentation result of the sample under consideration. The test input and target values are loaded in the classifier for training. The values for validation and testing for given sample is set in percentage. The hidden layers are chosen by considering the fact that computational complexity will increase and the number of iterations will decreases. Network is created and trained with extracted values. The network is tested with samples and features which gives MSE and percentage of error. The values for matrix figure 10. and ROC are plotted, from which T,P,F and N are determined. The values are plotted against actual and predicted.
Fig. 10: Confusion Matrix.

shows the confusion matrix for the hidden layer \( \rightarrow 1 \). The confusion matrix shows the values for True, False, Positive and Negative and it will show in figure 11.

\[
\begin{array}{c|c|c|c|c}
 & T & N \\
\hline
P & True Positive & True Negative \\
F & False Positive & False Negative \\
\end{array}
\]

Fig. 11: Values in 2x2 Matrix.

Fig. 12: ROC Graph.

Receiver Operating Characteristic (ROC) is the graph between FPR and TPR is shown in figure 12. The diagonal line is diagonally guess line which is a point of \((0.5, 0.5)\). The line above diagonal line is good and below is worse. At the point \((0, 1)\) is perfect classification which classifies cancer and non-cancer. For extracted values, graph has been plotted. From the above figure 12, it is observed that classification is better when compared to EM Algorithm.

**Accuracy:**
Table 5: Performance Measure of Accuracy for MLS and EM Algorithm.

<table>
<thead>
<tr>
<th>No. of Hidden Layers</th>
<th>BPN</th>
<th>MLS Algorithm</th>
<th>EM Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>95%</td>
<td>95%</td>
<td>94%</td>
</tr>
<tr>
<td>2</td>
<td>95.7%</td>
<td>93%</td>
<td>93%</td>
</tr>
<tr>
<td>3</td>
<td>96%</td>
<td>90%</td>
<td>90%</td>
</tr>
<tr>
<td>4</td>
<td>98%</td>
<td>84%</td>
<td>84%</td>
</tr>
<tr>
<td>5</td>
<td>99.2%</td>
<td>83%</td>
<td>83%</td>
</tr>
</tbody>
</table>

For every number of hidden layers, MLS Algorithm gives better result than EM Algorithm is shown in table 5.

**Conclusion:**

The modified level set using active contour algorithm for segmentation results in better classification accuracy than the EM algorithm. This algorithm tested for various blood cells. Extract features from the segmented result such as geometric and texture features is little higher, the WBC extraction of blood smear images followed by relevant feature extraction for immature cell detection is achieved and hence, the immature cell detection with the extracted features are classified with BPN classifier. The future work is to validate the results with medical persons and modify the neural network algorithm by weight updation and by reducing the number of iterations.

**REFERENCES**


www.chillibreeze.com/articles_various/cancer-in-India.

