SEGMENTATION AND ABNORMALITY DETECTION OF CERVICAL CANCER CELLS USING FAST ELM WITH PARTICLE SWARM OPTIMIZATION

P. SUKUMAR\(^1\) and R.K. GNANAMURTHY\(^2\)

\(^1\) Faculty of Electronics and Communication Engineering, Nandha Engineering College, Erode, India.
\(^2\) Principal, SKP Engineering College, Tiruvannamalai, India


Cervical cancer arises when the anomalous cells on the cervix mature unmanageable obviously in the renovation sector. The most probably used methods to detect abnormal cervical cells are the routine and there is no difference between the abnormal and normal nuclei. So that the abnormal nuclei found are brown in color while normal nuclei are blue in color. The spread or cells are examined and the image denoising is performed based on the Iterative Decision Based Algorithm. Image Segmentation is the method of paneling a digital image into compound sections. The major utilize of segmentation is to abridge or modify the demonstration of an image. The images are segmented by applying anisotropic diffusion on the Denoised image. Image can be enhanced using dark stretching to increase the quality of the image. It separates the cells into all nuclei region and abnormal nuclei region. The abnormal nuclei regions are further classified into touching and non-touching regions and touching regions undergoes feature selection process. The existing Support Vector Machines (SVM) is classified few nuclei regions but the time to taken for execution is high. The abnormality detected from the image is calculated as 45% from the total abnormal nuclei. Thus the proposed method of Fast Particle Swarm Optimization with Extreme Learning Machines (Fast PSO-ELM) to classify all nuclei regions further into touching region and separated region. The iterative method for to training the ELM and make it more efficient than the SVM method. In experimental result, the proposed method of Fast PSO-ELM may shows the accuracy as above 90% and execution time is calculated based on the abnormality (ratio of abnormal

*Corresponding author:* P. SUKUMAR, Faculty of Electronics and Communication Engineering, Nandha Engineering College, Erode, India, sucumarnandha@gmail.com
nuclei regions to all nuclei regions) image. Therefore, Fast PSO-ELM helps to detect the cervical cancer cells with maximum accuracy.

Key words: Cervical Cancer, Image Denoising, Extreme Learning Machine, White Blood Cells, Particle Swarm Optimization, Fast Extreme Learning Machine.

INTRODUCTION
Cancer is sourced through the gathering of inherited and epigenetic alterations in genes that usually cooperate a function in the directive of cell propagation (as explained in the Cell Proliferation lecture), thus foremost to unrestrained cell development. Cells obtain mutations in these genes as an effect of impulsive and environmentally-induced DNA injure. Those cells with alterations that encourage a growth and endurance benefit over ordinary cells are preferred for during a Darwinian method, leading to the development of a cancer. Human papilloma viruses (transmitted through sexual contact) occur a pathogenic function in the majority cases of cervical cancer (observe information on cervical cancer in the Pathological Characteristics of Benign and Malignant Neoplasms lab, the Cancer Screening lecture, and the Clinical and Translational Research online module). There are at least 77 subtypes of HPV that are distinguished by variations in their DNA progressions. HPV-16 or HPV-18 DNA is found in 70% of cervical tumors. An additional 20% of tumors contain HPV DNA matching to one of 20 other cancer-correlated subtypes. HPVs have a double-stranded DNA genome. Also, the selection of appropriate training algorithms is important, which enables the considered application to not experience the local or global minima problem. The above addressed criteria for improving the training performance, so as to result in better classification accuracy has been noted in neural network architecture, Fast Particle Swarm Optimization with Extreme Learning Machine (Fast PSO-ELM) Classifier proposed. White Blood Cells (WBCs), also known as leukocytes or leucocytes. These are cells of the protected structure that are concerned in the body not in favor of both contagious infection and unfamiliar resources. These cells help fight infections by attacking bacteria, viruses, and germs that invade the body. All leukocytes are produced and derived from a multipotent cell. This is also known as a hematopoietic stanch cell. They subsist for concerning 3 or 4 days in the human body. The microscopy images of cells in cervix uteri are marked, so that the abnormal nuclei found are brown in color while normal nuclei are blue in color which is shown in the Figure 1.

Fig 1: Input Image
First the image undergoes denoising process. Normally noises like Salt and Pepper Noise is formed in medical images. The main goal of an image denoising algorithm is then to reduce the noise level, while preserving the image features (such as edges, textures, etc.).

The rest of this research is structured as follows. Section II summarizes the concepts and related works. Section III details the proposed method, and Section IV discusses the experiments and the achieved results. Finally, Section V presents the conclusions of the work.

The article presented by YI et al. (2005) tackles a fast White Blood Cell (WBC) image segmentation method realized by on-line trained neural network. A pre-selecting performance, derived from mean shift algorithm and standardized sampling, is exploited as an initialization device to mainly diminish the training set while protecting the mainly expensive sharing information. Additionally, particle swarm optimization (PSO) is approved to train the network for a sooner junction and avoidance from a local optimum. WU et al. (2006), the segmentation scheme works successfully for classification of white blood cells. Along with the information that the H part in HSI color space encloses the majority of WBC information, and the S part encloses the structure information of the WBC nucleus, they expand an iterative Otsu's method founded on circular histogram for the leukocyte segmentation by enchanting full benefit of that knowledge.

JIANG et al. (2003) their article presents White Blood Cell (WBC) segmentation design using scale-Space filtering and watershed clustering is suggested. In their proposal, nucleus and cytoplasm, both of these mechanisms of WBC, are removed correspondingly using dissimilar techniques. Initially, a sub image including WBC is estranged from cell image. Through using feature space clustering procedure, this method productively evades the mixture and complication in image space, and can efficiently haul out various WBC sections from cell images of tangential blood spread. Their research exposes several common segmentation techniques that have established application in classification in biomedical-image dispensation particularly in blood cell image processing. Principally, segmentation of the image splits the entire image into some exclusive disjoint regions. The reality that the segmented image should keep greatest helpful information and abandon surplus information creates the complete procedures by ADOLAH et al. (2008).

LI et al. (2012) presents an effectual and proficient Computer Aided Diagnosis (CAD) scheme derived from Principle Component Analysis (PCA) and Extreme Learning Machine (ELM) to support the mission of thyroid disease diagnosis. The CAD method is included of three phases. Focusing on dimension reduction, the primary points apply PCA to create the majority discriminative latest feature set. After then, the method switches to the next phase whose goal is sculpt structure. ELM classifier is discovered to train a best prognostic model whose strictures are optimized.

SHANAVAS and GNANAMURTHY (2014) helps to perform the optimization of the benchmark circuits with the above said components of physical design using hierarchical approach of evolutionary algorithms. This hybrid approach can quickly produce optimal solutions for the popular benchmarks. A new learning algorithm called Extreme Learning Machine (ELM) for single-hidden layer feed forward neural networks (SLFNs) which arbitrarily decides concealed nodes and logically decides the output weights of SLFNs by HUANG et al. (2006). A hybrid learning algorithm is planned which utilizes the discrepancy evolutionary algorithm to choose the input weights and Moore–Penrose (MP) widespread contrary to logically decide the output weights by ZHU et al. (2005).
HUANG et al. (2010) investigates more revises ELM for sorting in the characteristic of the ordinary optimization process and widen ELM to a precise category of “generalized” SLFNs—support vector network. PIMENTA et al. (2013) estimates the proportion of ADC in persistent cervical cancer, the universal number of cases of cervical ADC in 2015, the effect of cervical screening on ADC, the number of ADC cases attributable to high-risk HPV types -16, -18, -45, -31 and -33, and the possible collision of HPV immunization using different data sources counting: GLOBOCAN (2008), Cancer Incidence in Five Continents (CI5) Volume IX, cervical showing information from the World Health Organization/Institut Català d’Oncologia Information Centre on HPV and cervical cancer, and available journalism. The HUCL is a hybrid of static and dynamic clustering approaches is proposed by MALATHI et al. (2015). In HUCL, the network is divided into layers and clusters of various sizes. The cluster heads are selected based on available energy, the distance to the sink and the number of neighbors.

WANG et al. (2013) presented as vary of total HPV, solitary HPV and numerous HPV disease were analogous through the five existence. They suggested that HPV disease is common with HPV16 and HPV 58 as the main subtypes in women in Shenzhen city. The occurrence of HPV 18 infection is increasing earlier than any others, which will direct it to be single of the major subtypes in this city in the prospect. Cai organized for efficiently support the researchers conduct quantitative analysis of ceramic micro-structures, a segmentation algorithm derived from mean shift is utilized for the clay micro-structure representation. As the collection and relocate procedure of microscopic image will inexorably be subject to patchy distribution of light, electronic noise and other meddling factors which create the image quality worsening, it is essential to reduce noises and enhance edges for ceramic microscopic image processing at first. So, the median filter is used to take away the noises in the clay microstructure images. Then the part with comparable feature is alienated and fused by the mean shift segmentation algorithm.

HUANG et al. (2005) examined the consequence of eight PSO topologies on presentation of the PSO-ELM. The outcomes showed empirically that the Global topology was additional shows potential than every other topologies in optimizing the PSO-ELM according to the Root Mean Squared Error (RMSE) on the validation set in the majority of the appraised datasets. Though, no correlation was detected between this excellent routine on the RMSE and the testing accurateness. Two-dimensional principal component analysis (2D PCA) is compared with other algorithms like 1D PCA, Fisher discriminant analysis (FDA), independent component analysis (ICA) and Kernel PCA (KPCA) which are used for image representation and face recognition by Senthilkumar and Gnanamurthy (2015). SARASWATHI et al. (2011) appraises the presentation of ICGA-PSO-ELM and contrast their results with existing techniques in the prose. An exploration into the purposes of the chosen genes, using an organism’s biology method, exposed that several of the recognized genes are concerned in cell signaling and propagation. An investigation of these gene sets demonstrates a better illustration of genes that program concealed proteins than found in haphazardly preferred gene sets. A novel hybrid approach founded on clustering and Particle Swarm Optimization (PSO) is suggested for gene selection and categorization of microarray data by YANG et al. (2013). In those approaches, PSO merging with clustering technique are used to achieve gene selection to decrease redundancy. Owing to its improved simplification presentation with much faster junction tempo than further learning algorithms for neural networks, Extreme Learning Machine (ELM) is preferred to perform sample classification in the hybrid method.
MATERIALS AND METHODS

In this section demonstrated that the proposed Fast PSO-ELM method for defect detection in cervical cancer cells. The main aspects of the algorithm namely, Extreme Learning Machine, Particle swarm Optimization is discussed as follows.

**Extreme Learning Machine**

Extreme learning machine (ELM) suggested is a single-hidden layer feed forward networks (SLFNs) which haphazardly preferred the input influences and systematically decides the output influences of SLFNs. One input rule of the ELM is that one could haphazardly decide and attach the hidden nodule parameters. After the hidden nodules parameters are selected accidentally, SLFN suits a linear system where the output influences of the network can be methodically established using simple simplified inverse process of the hidden layer output matrices. For an surveillance data set with $N$ nodes in the hidden layer and the excitation utility $G$, the extreme learning machine model can be expressed as

\[
f(x) = \sum_{i=1}^{N} \beta_i G(a_i, b_i, x_i) = \beta . h(x),
\]

Where $\beta_i$ is the output weight of the $i$th hidden layer node and the output neuron, $a_i$ is the input weight of the input neuron and the $i$th hidden layer node, and $b_i$ is the offset of the $i$th hidden layer node. Consider $h(x) = [G(a_1,b_1,x_1),\ldots,G(a_N,b_N,x_N)]$ denotes the output matrix of hidden layer. $a_i$ and $b_i$ are randomly selected before training and remain the same in the training procedure. The output weights $\beta_i$ can be obtained by solving the least-squares solutions of the following linear equation:

\[
\min_{\beta} \sum_{i=1}^{N} \| \beta . h(x_i) - y_i \|
\]

It obtain a best solution and density when compared to ELM, adapted PSO-ELM, etc., then ELM has been successfully used in the following applications: Biometrics, Bioinformatics, Image processing, Signal processing etc., ELM is a simple tuning-free three-step Algorithm. The machine learning rapidity of ELM is tremendously rapid. The conventional standard learning algorithms which only work for differentiable establishment functions however it also facing some issues like local minima, over fitting, etc.,
Given a training set \( \mathcal{N} = \{ (x_i, t_i) | x_i \in \mathbb{R}^n, t_i \in \mathbb{R}^m, i = 1, \ldots, N \} \), activation function \( g(x) \), and hidden node number \( N \),

**Step 1:** Randomly assign input weight \( w_i \) and bias \( b_i \), \( i = 1, \ldots, N \).

**Step 2:** Calculate the hidden layer output matrix \( H \).

**Step 3:** Calculate the output weight \( \beta \)

\[
\beta = H + T
\]

Where \( T = [t_1, \ldots, t_N]^T \).

The ELM learning algorithm looks much simpler and it gives accurate result when compare to other algorithms. Machine Learning is about building programs with tunable parameters. Extreme Learning Machines (ELM) provides efficient solutions. ELM possesses unique features to deal with regression and (multi-class) classification tasks. Consequently, ELM offers advantages such as fast learning speed, ease of implementation.

**Standard Particle Swarm Optimization:**

PSO suggests the group performance of classification, for instance flora and fauna in a gather or fish in a school, and preserve be expressed as a mechanically sprouting method. PSO works through initializing a flock of flora and fauna haphazardly over the penetrating space, where each bird is known as a “particle”. These “particles” fly through a convinced pace and locate the large-scale best position after various iteration. At every iteration, every particle adjusts its pace vector, founded on its impetus and the influence of its top position \( (P_b) \) with the finest position of its neighbors \( (P_p) \), and then a new position the “particle” to fly is obtained. Guessing the element of searching space is \( D \), the whole number of particles is \( n \), the position of the \( i \)-th particle can be stated as vector \( X_i = (x_{i1}, x_{i2}, \ldots, x_{iD}) \); the best position of the \( i \)-th particle penetrating in anticipation of presently is indicated as \( P_{ib} = (p_{i1}, p_{i2}, \ldots, p_{iD}) \) and the best position of all particles penetrating awaiting at present is denoted as vector as \( P_g = (p_{g1}, p_{g2}, \ldots, p_{gD}) \); the velocity of the \( i \)-th particle is indicated as vector as \( V_i = (v_{i1}, v_{i2}, \ldots, v_{iD}) \). Then the original PSO is illustrated as:

\[
V_{id}(t+1) = V_{id}(t) + C_1 \cdot \text{rand}(0,1) \cdot [p_{id}(t) + x_{id}(t)] + C_2 \cdot \text{rand}(0,1) \cdot [p_{gd}(t) + x_{id}(t)] \\
X_{id}(t+1) = X_{id}(t) + V_{id}(t+1)
\]

Then, the parameters (all weights and biases) are restructured with the help of Extreme Learning Machine algorithm. In Figure 2 illustrates that the proposed methodological process is used to detect the cervical cancer cells.
Fig 2: Methodology Used To Detect Cervical Cancer Cells

Fast Particle Swarm Optimization with ELM Method:

This research work integrates to progress Fast PSO-ELM by combining an improved PSO with ELM. The Fast PSO-ELM to select the input influences to boost the generalization presentation and the provisioning of the SLFN. The details of the proposed method are as follows:

Step 1: Initialize a population array of swarm particles with of a set of input influences and secreted favoritisms: \( P_i = [W_{11}, W_{12}, ..., W_{1n}, W_{21}, W_{22}, ..., W_{2n}, ..., W_{H1}, W_{H2}, ..., W_{Hn}, b_1, b_2, ..., b_H] \) with random initialized within the range of \([-1, 1]\) on \( D \) dimensions in the search space.

Step 2: for every group particle, the matching harvest weights are calculated according to ELM as in Equation (1).

Step 3: Then the fitness of each particle \( f(x) \) is evaluated as in Equation (5). In order to avoid over fitting of the SLFN, the fitness of every particle is accepted as the Root Mean Squared Error (RMSE) on the validation locate only in preference to the complete training set.
where \( f(P_i), f(P_{i,best}) \) and \( f(g_{i,best}) \) are the corresponding fitness for the \( i \)-th particle, the best position of the \( i \)-th particle and global best position of all particles, respectively. \( w_0P_i, w_0P_{i,best} \) and \( w_0g_{i,best} \) are the corresponding output weights obtained by MP generalized inverse when the input weights are set as the \( i \)-th particle, the best position of the \( i \)-th particle and global best position of all particles, respectively. The parameter \( \eta > 0 \) is tolerance rate.

**Step 4:** Velocity update - Update the velocities of all particles at time \( k \) (current iteration) using the particles objective or fitness values which are functions of the particles current positions in the design space at time \( k \). At each iterations, the rapidity of the entire particles is rationalized as:

\[
v_i(k + 1) = \beta [v_i(k) + c_1 \gamma_1 (P_{i,best} - p_i) + c_2 \gamma_2 (g_{i,best} - p_i)]
\]  

**Step 5:** Position update - The position of each particle is updated using velocity vector as follows:

\[
p_i(k + 1) = x_i + v_i(k + 1)
\]  

**Step 6:** Memory update - Update \( p_{i,best} \) and \( g_{i,best} \) when circumstance is convened and new-fangled population is generated.

**Step 7:** Stopping Criteria - The algorithm repeats steps 3 to 6 until certain criteria are met, along with hard threshold value as maximum number of iterations. Once stopped, the algorithm describes values of \( g_{best} \) and \( f(g_{best}) \) as its resolution.

Thus improved PSO with ELM finds the best optimal weights \( W \) and bias \( b \) so that the fitness reaches the minimum to achieves better generalization performance, with minimum number of hidden neurons. It makes advantage of both ELM and PSO. The procedure describing proposed Fast PSO-ELM approach is as follows.

1. Initializing FPSO with population size, inertia weight and creations without improvable.
2. Estimating the robustness of each particle.
3. Comparing the robustness values and establishes the restricted best and inclusive best particle.
4. Updating the velocity and position of each particle till value of the robustness
5. After touching, the large-scale best particle in the swarm is fed to ELM classifier for training.

6. Training the ELM classifier.

The Fast PSO-ELM takes the benefit of minimum structural risk of ELM, where Structure Risk Minimization (SRM) is an inductive principle of use in machine learning and the quick global optimizing ability of FPSO.

The function of algorithm of segmentation by particulate swarm, like any evolutionary algorithm, is influenced by factors such as the criterion of stop, the structure of particle, the objective function.

RESULTS AND DISCUSSION

The experimental results are performed on assorted blood cell images using MATLAB. Since that the WBC are only segmented and its number of WBC noticed through a variety of techniques is evaluated with essentially present in the image which is physically gained. Collected 50 images from pathologist and 40 used for training data and remaining 10 for testing data. To classify abnormal nuclei regions and all nuclei regions, fast PSO-ELM technique is used in 'a*b*' regions based on the color information. By using appropriate n-values (Assume n=6), it is possible to separate abnormal nuclei regions and all nuclei regions respectively with the help of proposed technique. It is shown in the figure 3.

Fig 3: Abnormal Nuclei Region

The main aim of Fast PSO-ELM is to locate the boundary which is most distant from the nearest vectors of both of the two categories. Using proposed Fast PSO-ELM, the touching region and separated regions are classified exactly which is represented in figure 4 and 5.
Fig 4: Non Touching Nuclei Region

Fig 5: Touching Nuclei Region

The accuracy is calculated by combining final segmented image obtained by extracting the all nuclei region and the enhanced image nuclei are indicates the abnormality count detected in abnormal nuclei region (shown in figure 6)

Fig 6: Detected Abnormal Nuclei
Testing Precision and Recall

Abnormality is defined as the ratio of abnormal nuclei regions to that of all nuclei regions. The accuracy calculations for abnormal and normal nuclei to provide prevalence are as follow.

\[
\text{Accuracy} = (\text{sensitivity}) \times (\text{prevalence}) + (\text{specificity}) \times (1-\text{prevalence})
\]

The accuracy may be determined from sensitivity and specificity. The sensitivity relates to the tests ability to identify the condition correctly and specificity relates to the tests ability to exclude the condition correctly. These both are statistical measures for the performance of binary classification tests. The abnormal accuracy and nuclei accuracy is calculated as follows.

\[
\text{Sensitivity} = \frac{\text{No. of true positives}}{\text{No. of true positives} + \text{No. of false negatives}}
\]
\[
= \frac{\text{total No. of sick individuals in population}}{\text{probability of a positive test, given that the patient is ill}}
\]

\[
\text{Specificity} = \frac{\text{No. of true negatives}}{\text{No. of true negatives} + \text{No. of false positives}}
\]
\[
= \frac{\text{total No. of well individuals in population}}{\text{probability of a negative test, given that the patient is well}}
\]

By using above equations in proposed method, it is measured abnormal and nuclei accuracy. Thus the proposed Fast PSO-ELM technique finds most of the abnormal nuclei better than existing method and quite less result for all nuclei present in the cervical cancer cells.

Table 1 shows the accuracy and execution time for Proposed Fast PSO-ELM technique

<table>
<thead>
<tr>
<th>Techniques</th>
<th>Accuracy (%)</th>
<th>Execution Time (Seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELM</td>
<td>67.55</td>
<td>49</td>
</tr>
<tr>
<td>PSO-ELM</td>
<td>79.22</td>
<td>38</td>
</tr>
<tr>
<td>Fast PSO-ELM</td>
<td>94.76</td>
<td>21</td>
</tr>
</tbody>
</table>
The Table 1. gives the accuracy and execution time comparison of the proposed method of abnormality detection using Fast PSO-ELM. That the abnormality detection for cancer cells through touching and non touching nuclei for enhancing the process using Fast PSO-ELM gives 94.76% accuracy and less execution time for proposed method of Fast PSO-ELM. The calculation are provided the potential solution for time concerning in proposed method.

CONCLUSION

Cervical cancer is the second most common cancer among the entire world. Numerous cells are present in noise spread image and it is complicated to find abnormal nuclei exactly using the image. The proposed technique provides early diagnostics treatment for cancer cells. Also, this research work presents information to detect the abnormal cells using various algorithms. When the image is affected by noise, it can be easily detected and removed by using denoising process. In tradition, there is no denoising techniques are implemented for cell image defects. The excellence of the image is improved using diffusion and image stretching. Fast PSO-ELM algorithm is used to separate the touching and separated regions from all nuclei and abnormal nuclei. The accuracy is calculated based on abnormality and final segmented image. And this method achieves an abnormal detection accuracy of 96%. The main beneficial of the proposed method is that it is appropriate for images tarnished by high degree of noise, as it can effectively isolate abnormal and all nuclei regions with higher quantity of accuracy. In future Fast PSO-ELM technique can be executed in numerous cells for identifying better accuracy. The minute cell extraction can be done with the assist of this technique.

List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>Support Vector Machine</td>
</tr>
<tr>
<td>PSO</td>
<td>Particle Swarm Optimization</td>
</tr>
<tr>
<td>ELM</td>
<td>Extreme Learning Machine</td>
</tr>
<tr>
<td>SLFN</td>
<td>Single-hidden Layer Feed-forward neural Network</td>
</tr>
<tr>
<td>HPV</td>
<td>Human PapillomaVirus</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribo-Nucleic Acid</td>
</tr>
<tr>
<td>WBC</td>
<td>White Blood Cells</td>
</tr>
</tbody>
</table>

REFERENCES


SEGMENTACIJA I DETEKCIJA ABNORMALNIH ĆELIJA CERVIKALNOG KANCERA KORIŠĆENJEM METODA BRZOG ELM-a

P.SUKUMAR1 i R.K.GNAMURTHY2

1 Fakultet elektroničkih i komunikacionih inženjeringa Nandha Inženjering koledž, 
2 Glavni SKP Koledž za inženjering, Tiruvannamalai, India

Izvod

Najviše korišćeni metodi za utvrđivanje abnormalnih cervikalnih ćelija su rutinski i ne mogu da utvrde razlike normalnih i abnormalnih jedara. Predmet rada je uvođenje novih metoda koje omogućavaju utvrđivanje razlika normalnih i abnormalnih jedara. U radu su dati detalji materijala i metoda rada i detaljno prikazani rezultati. Eksperimentalni podaci su potvrdili da predložena metoda brzog PSO-ELM pokazuje 90 % pouzdanosti i vreme operacije se izračunava pomoću kalkulacije zasnovane na odnosu abnormalnih i ostalih regiona jedra. Na taj način brzogPSO-ELM metoda pomaže da se brzo detekuje kancer grlića sa maksimalnom pouzdanošću.

Primljeno 04. IV 2015. 
Odobreno 05.IX. 2015.