ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 4, Issue 11, November 2016**

# Diagnosis of Cervical Cancer using Global and Local Shape Features

B.Ashok[1], S.Anu.H.Nair[2], N.Puviarasan[3], P.Aruna[4]

Assistant Professor/Programmer, Dept. of C.S.E, Annamalai University, India[1]

Assistant Professor, Dept. of C.S.E, Annamalai University, India[2]

Associate Professor, Dept. of C.S.E, Annamalai University, India[3]

Professor, Dept. of C.S.E, Annamalai University, India[4]

**ABSTRACT**: The main objective of this paper is to diagnose the cervical cancer at earlier stage using pap smear images. It is performed by explicit use of shape features. Segmentation of pap smear cell image is carried out using active contour model. The segmented pap smear images are encoded with global and local shape features. The global and local shape features are combined and then feature selection is performed by applying sequential floating forward selection method. BPNN and SVM classifiers are utilized. Results show that SVM classifier yields better accuracy than BPNN classifier for diagnosis of cervical cancer.

**KEYWORDS**: Cervical cancer, Classification, Feature selection, Segmentation, Shape features

## I. INTRODUCTION

Cervical cancer is the second common type of cancer among women all over the world. It is believed that the human papilloma virus (HPV), more specifically HPV 16 and HPV 18 type of viruses are the main constraints for the formation of cervical cancer. Screening tests are used to diagnose the cervical cells whether they are normal or abnormal. Life of the patients can be saved by early diagnosis. For early diagnosis, regular screening tests are required for women. Developing countries like India still have lake of awareness. So, the cervical cancer affected patients are more in India while comparing with developed nations. In order to conduct more screening test, an automated diagnosis method is necessary because manual diagnosis may take more time and more skilled man power. Recent years, lots of computer aided diagnosis (CAD) methods are developed rapidly to assist doctors in confirmative diagnosis. All these methods have its own merits and demerits. Still more researches are required to improve the reliability and accuracy of diagnosis. Shape features can be divided as local and global shape features. Global features represent the whole image and local features encode the specific region of an image. In this work, cervical cancer diagnosis method is presented using both shape features.

The paper is organized as follows. Section 2 briefs about related works. Section 3 presents the methodology. Section 4 deals with results and discussion. Section 5 concludes the work.

## II. RELATED WORK

Dibet Garcia-Gonzalez et al (2016) presented a new method to segment nucleus in pap smear images which combines segmentation and classification algorithms. Loris Nanni et al (2015), Edwin Jayasingh Mariarputham et al (2015) and Sukumar P et al (2015) discussed SVM based classification methods for cervical cancer. Yessi Jusman et al (2014) utilized ANN, SVM, kNN and LDA classifiers. Thanatip Chankong et al (2014) and Lili Zhao et al (2016) suggested new methods for cervical cancer cell segmentation. Siti Noraini Sulaiman et al (2015) proposed clustering algorithm to segment pap smear images into the nucleus, cytoplasm and background regions. Sayeda, Ahmed M et al (2016) implemented classification of breast tumors using magnetic resonance images. Ayubu Hassan Mbaga et al (2014) applied support vector machine based on recursive feature elimination.

Nagarajan,G. et al (2016) presented about medical image feature extraction and selection. Bolon-Canedo, V., et al (2014) reviewed the most recent feature selection methods developed in this field and presented medical binary dataset

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

contains various cancers. Jothi G et al (2016) performed supervised feature selection for MRI brain tumor image classification. Meenakshi M. Kavitha R et al (2014) implemented segmentation and classification for cervical cells. Rabindra Kumar Singh et al (2015) reviewed the feature selection techniques based cancer classification and also the predominant role of SVM for cancer classification. Barnali Sahua et al (2012) performed feature selection using particle swarm optimization for cancer microarray data. Anousouya Devi,M. et al (2016) and Xiaoping Qiu et al (2007) implemented classification of normal or abnormal cells of cervical cancer using ANN. Konstantina Kourou et al (2015) conducted a reviewed machine learning algorithms such as ANN and SVM for cancer prognosis.

## III. METHODOLOGY

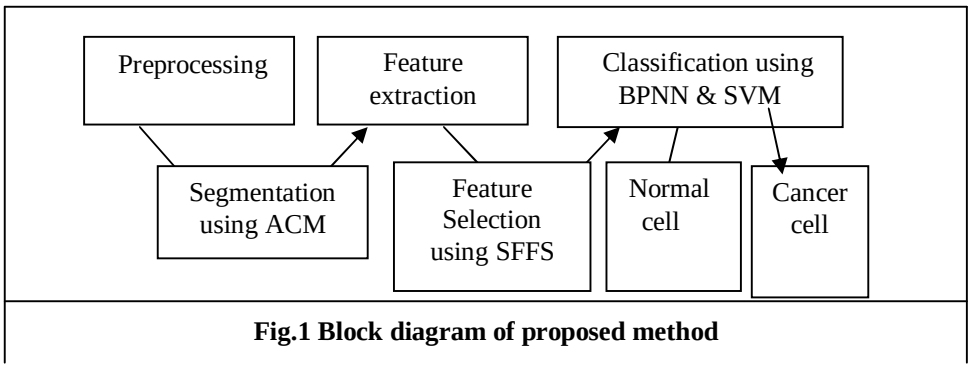



**Fig.1 Block diagram of proposed method**

Fig.1 depicts the block diagram of the proposed method.

### *A. DATASET*

To implement the method of this research work, the dataset is obtained from Rajah Muthiah Medical College & Hospital, Annamalai University, Annamalainagar. The dataset contains 480 cervical cell images.

### *B. SEGMENTATION*

Image preprocessing is the first procedure in image processing which involves gray scale conversion, noise or artifacts removal and image enhancement. Segmentation is a method which subdivides an image into regions. Segmentation accuracy decides the eventual success or failure of computerized analysis procedures. In this work, active contour model is applied for segmentation. Active contour model is a process to obtain a segmented object. Active contour is a energy-minimizing spline guided by external constraint forces. Active contour is also known as snakes which is a framework for getting object. The framework minimizes an energy associated to the current contour as a sum of internal and external energies. External energy expression is derived such that it is low at object boundary. Internal energy regulates the shape of contour controlling its curvature and shape regularity. Autonomous and self-adapting in their search for minimal energy is the main advantage of snakes. The contour is defined in the $(x, y)$ plane of an image and position of the snake $V(S) = (x(s) + y(s))$ energy function is given as

$$E^*_{snake} = \int_0^1 E_{snake(V(S))} \mathrm{ds} \quad (1)$$

$$= \int_0^1 E_{int}\left(V(S)\right) + E_{image}\left(V(S)\right) + E_{con}\left(V(S)\right) ds \quad (2)$$

where $E_{int}$ is internal energy of the spline due to bending, $E_{image}$ is rise to the image forces and $E_{con}$ is rise to the external constraint forces.

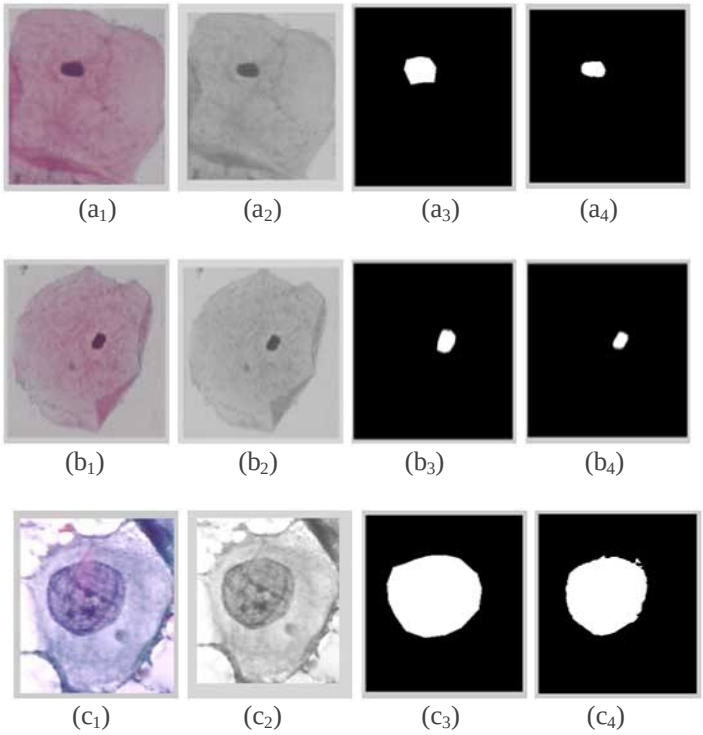
It is shown in the Fig.2, images ($a_1$ - $c_1$) show the original image. Images ($a_2$ - $c_2$) are the gray scale image of ($a_1$ - $c_1$) respectively. Image ($a_3$) depicts the user made curve which is very near to nucleus boundary. The image ($a_4$) displays the segmented image of nucleus. In the same manner, the images ($b_3$ - $b_4$) and ($c_3$ - $c_4$) are shown the various stages in nucleus segmentation. In Fig.3, the images $d_1$ and $e_1$ are the original images. $d_2$ and $e_2$ are the respective gray scale images. $d_3$ and $e_3$ show the curvature very close to the cytoplasm boundary. $d_4$ and $e_4$ in Fig.3 are the segmentation of cytoplasm. The contours slide while minimizing their energy, high level computation can activate with the contour model by forcing it toward an appropriate local minimum. The curvature then starts deforming and moving towards the desired object boundary. At the end, it completely shrink-wraps around the object. The segmentation quality highly depends on the image. Clear gray level objects are suitable for the best quality of segmentation.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

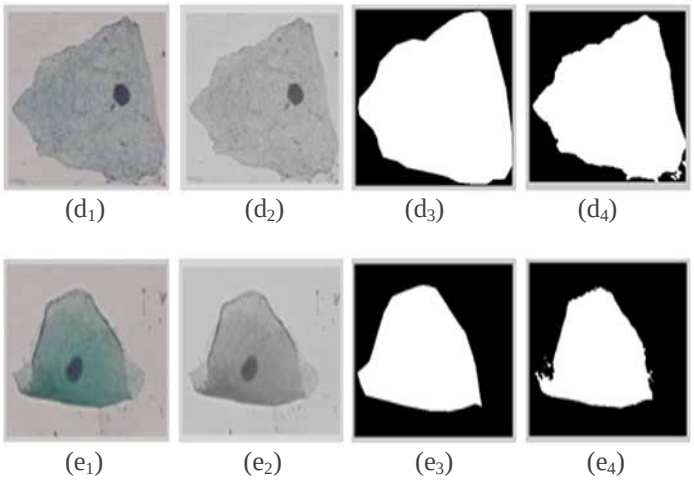


($a_1$) ($a_2$) ($a_3$) ($a_4$)

($b_1$) ($b_2$) ($b_3$) ($b_4$)

($c_1$) ($c_2$) ($c_3$) ($c_4$)

**Fig. 2 Nucleus Segmentation using Active Contour Model**



($d_1$) ($d_2$) ($d_3$) ($d_4$)

($e_1$) ($e_2$) ($e_3$) ($e_4$)

**Fig. 3 Cytoplasm segmentation using Active Contour model**

### *C. LOCAL SHAPE FEATURES*

Local shape feature is the representation or encoding of a specific region of an image. Several local shape features are required to encode an image. In this work, nucleus and cytoplasm based local shape features are extracted. The ratio of nucleus and cytoplasm is given as

$$\text{Nucleus-Cytoplasm ratio} = \frac{Nucleus_{area}}{Nucleus_{area} + Cytoplasm_{area}} \quad (3)$$

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

Nucleus longest diameter and Cytoplasm longest diameter is calculated as the longest distance between two pixels on the objects border. Nucleus shortest diameter and cytoplasm shortest diameter is defined as the biggest diameter a circle can have when the circle is totally inscribed in the object. Elongation of nucleus is given by

$$Nucleus_{elong} = \frac{Nucleus_{short}}{Nucleus_{long}} \quad (4)$$

Elongation of cytoplasm is given by

$$Cytoplasm_{elong} = \frac{Cytoplasm_{short}}{Cytoplasm_{long}} \quad (5)$$

Nucleus brightness is calculated as the average perceived brightness that is a function of the colors wavelength

$$Nucleus\ Brightness = 0.299\ Red\ \mu + 0.587\ Green\ \mu + 0.114\ Blue\ \mu \quad (6)$$

Nucleus roundness is computed as

$$Nucleus_{Circle} = \frac{\pi}{4} Nucleus_{long}^2 \quad (7)$$

$$Nucleus_{roundness} = \frac{Nucleus_{area}}{Nucleus_{circle}} \quad (8)$$

$$Nucleus_{position} = \frac{2a\sqrt{(x_{Nucleus} - x_{Cytoplasm})^2 + (y_{Nucleus} - y_{cytoplasm})^2}}{Cytoplasm_{long}} \quad (9)$$

Likewise, nucleus related features such as area, eccentricity, equivdiameter, major axis length, minor axis length, convex area, perimeter, solidity, euler number, extent, orientation, centroid and cytoplasm related features such as area, eccentricity, equivdiameter, major axis length, minor axis length, convex area, perimeter, solidity, euler number, extent, orientation, centroid and roundness are extracted. In this work, totally 32 local shape features are obtained.

### *D. MOMENT INVARIANT FEATURES*

Hu, M.K., (1962) proposed seven properties related to connected region that are invariant to rotation, scaling, and translation. Hu obtained six absolute orthogonal invariants and one skew orthogonal invariant. They are independent of orientation, position, size and parallel projection. The moment invariants have been accepted to be the sufficient measures for analyzing image patterns concerned with the images translation, rotation and scaling under the assumption of images with continuous functions and noise-free.

$$\emptyset_1 = \eta_{20} + \eta_{02} \quad (10)$$

$$\emptyset_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (11)$$

$$\emptyset_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \mu_{03})^2 \quad (12)$$

$$\emptyset_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \mu_{03})^2 \quad (13)$$

$$\emptyset_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (14)$$

$$\emptyset_6 = \eta_{20} - \eta_{02}[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \quad (15)$$

$$\emptyset_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - (\eta_{30} - 3\eta_{12})^2(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (16)$$

$\emptyset_7$ is the skew moment and this invariant distinguish the mirror images. The seven moment invariants are unchanged under image scaling, translation and rotation.

In this work, 32 local shape features and 7 moment invariant features, totally 39 shape features have been extracted.

### *E. FEATURE SELECTION*

The sequential floating forward selection (SFFS) method is one of the most efficient feature selection techniques. This algorithm starts with a null feature subset. Then, the best feature in the feature set that satisfies some criterion function is added with the current feature subset. This is the main process of the sequential forward selection. Further, any improvement in the criterion is searched if some feature is excluded. By this way, the worst feature according to the criterion is eliminated from the feature subset. This is the process of sequential backward selection.

In this manner, the algorithm proceeds by adding or subtracting the number of features until the optimal feature subset is selected. The word floating denotes the increase or decrease in dimensionality of features in the selected feature subset. Let full feature set is denoted as $F = \{f_1, f_2, \ldots f_n\}$. An optimal feature subset that has to be arrived is given as $S = \{s_1, s_2, \ldots s_m\}$ where $m < n$ . That is the number of selected features must be less than the full feature set. Initially the selected feature set is set to null and it is denoted as $S = \{\emptyset\}$.

Algorithm 1 shows the step by step procedure of sequential floating forward selection method.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

**Algorithm 1 : Sequential Floating Forward Selection method**

**Input** : 39 number of input features
**Output** : 18 number of selected features

Step 1: Input the set of all features, $F = \{f_1, f_2, \ldots f_n\}$.

Step 2: Initialize the empty feature subset $S = \{\emptyset\}$, and counter variable $k = 0$

Step 3: Select the best significant feature

$$f^+ = \text{argmax}[J(S_k + f)] \; ; \quad f \notin S_k$$

$$S_k = S_k + f^+$$

$$k = k + 1$$

Step 4: Select the least significant feature

$$f^- = \text{argmax}[J(S_k - f)] \; ; \quad f \in S_k$$

Step 5: If $J(S_k - f^-) > (S_k)$ then

$$S_{k+1} = S_k - f$$

$$k = k + 1$$

go to Step 4

else

go to Step 3

Step 6: Get a selected optimal subset of features $S$.

### *F. BACKPROPAGATION NEURAL NETWORK (BPNN) CLASSIFIER*

Backpropagation is a training algorithm for multilayer supervised artificial neural network. Before to start training an artificial neural network, the network is built. It means that the number of nodes in the input layer, output layer and the hidden layer is defined.

In the backpropagation network, the input layer receives the input and the output layer produces the output. The network structure is given in Fig. 4.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

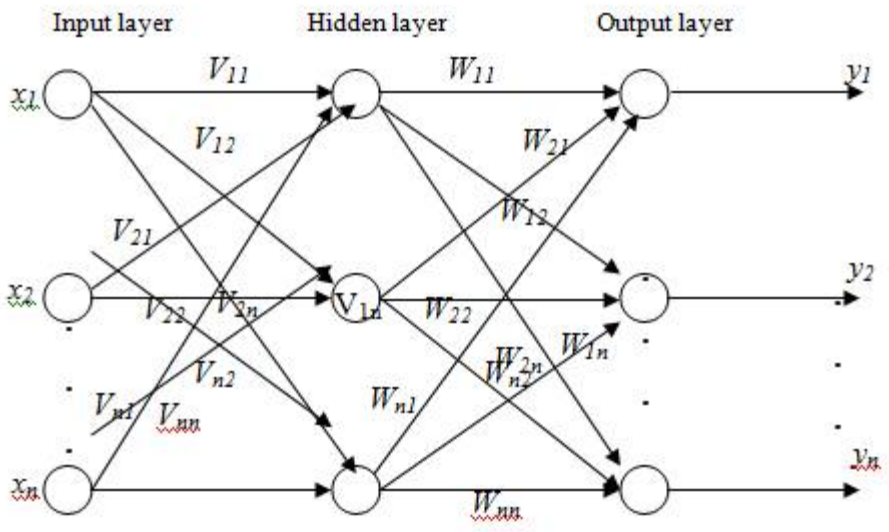



Fig. 4 Two layer backpropagation neural network.

*1) Backpropagation Training Algorithm*

The first step in training process is, all the weights must be assigned to small random numbers. In order to train the backpropagation network, the following steps are required:

1. The training pair from the training set is selected and the input vector is applied to the network input.
2. The output of the network is calculated using activation functions.
3. The error between the network output and the desired output is calculated.
4. The weights of the network are adjusted in a way that minimizes the error.
5. Steps 1 to 4 is repeated for each training pair in the training set until the error for the entire set is preferably low.

The *s* value of each neuron in the first layer is calculated as the weighted sum of its neurons inputs. The activation function *f* then forces *s* to produce the output value *o* for each neuron in that layer. Once the set of outputs for a layer is calculated, it serves as an input to the next layer. This process is continuously repeated until the final layer of network output is produced. This is expressed mathematically as follows:

$$s_j = \sum_{i=1}^{n} x_i V_{ij} \tag{17}$$

$$o_j = f(s_j) = \frac{1}{(1+\exp(-s_j))} \tag{18}$$

where $x_i$ is $i^{th}$ element of the input pattern, $V_{ij}$ is weight from $i^{th}$ input to $j^{th}$ neuron in the hidden layer and $s_j$ is the activation value of the $j^{th}$ hidden neuron. The output of the hidden layer is given as an input to the output layer.

*2) Adjusting the Weights of the Output Layer*

The training process for a single weight from neuron *p* in the hidden layer *j* to neuron *q* in the output layer *k* is as follows: The output of a neuron in the output layer *k* is subtracted from the target value *u* to produce an error signal ($\delta$). This is multiplied by the derivative of the squashing function [*o*(1-*o*)] thereby producing the $\delta$ value as

$$\delta = o(1-o)(u-o) \tag{19}$$

The error value $\delta$ is multiplied by the output value from the neuron *p*. This product is in turn multiplied by learning rate coefficient $\eta$ and the result is added to the weight. The same process is repeated for each weight proceeding from neuron in the hidden layer to a neuron in the output layer.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

$$\Delta W_{pq,k} = \eta \delta_{q,k} o_{p,j} \quad (20)$$

$$W_{pq,k}(n+1) = W_{pq,k}(n) + \Delta W_{pq,k} \quad (21)$$

where $\eta$ is the training rate co-efficient, $W_{pq,k}(n)$ is the value of a weight from neuron $p$ in the hidden layer to neuron $q$ in the output layer at step $n$ and the subscript $k$ indicates that the weight is associated with its destination layer, $W_{pq,k}(n+1)$ is the value of the weight at step $n+1$ (after adjustment), $\delta_{q,k}$ is the value of the $\delta$ for neuron $q$ in the output layer $k$ and $o_{p,j}$ is the value of the output for neuron $p$ in the hidden layer $j$.

*3) Adjusting the Weights of the Hidden Layers*

The Hidden layers have no target vector, so the training process described above cannot be used. Backpropagation trains the hidden layers by propagating the output error back through the network layer by layer, adjusting weights at each layer. Equations (20) and (21) are used for both output and hidden layers. During training, weights in the hidden layer operate in the reverse, passing the value of $\delta$ from the output layer to the hidden layer. Each of these weights are multiplied by the $\delta$ value of the neuron to which it connects in the output layer. The value of $\delta$ needed for the hidden layer neuron is produced by summing all such products and multiplying by the derivative of the squashing function

$$\delta_{p,j} = o_{p,j}(1 - o_{p,j}) \sum_{q=1}^{n} \delta_{q,k} W_{pq,k} \quad (22)$$

Knowing the weights ( $W_{ij}$'s ) after training, the input patterns are fed so that ANN have the capability to recognize.

*6) SVM classifier*

SVM is developed in COLT-92 by Boser, Guyon & Vapnik at Bell laboratories from statistical learning theory since in 1960s. Support vectors are simply the co-ordinates in a multidimensional space. The very important training points are support vectors. They decide the hyperplane.

In Fig. 5, the SVM classification process is illustrated. The rounded objects are support vectors.

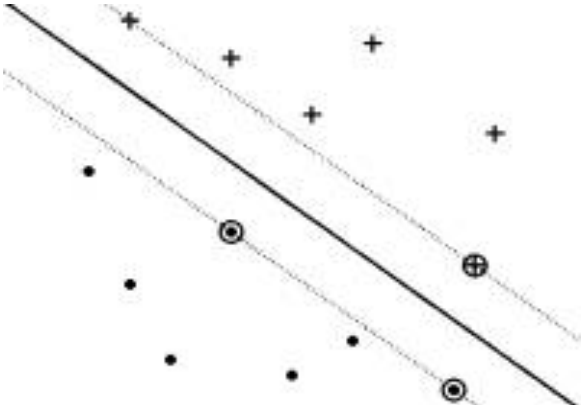


Fig. 5 SVM classifier

Fig. 5 shows the linear classification of SVM. Classifier separates the support vectors with their corresponding groups by a hyperplane. Kernels are similarity functions for machine learning algorithms. Support vector machines belong to member of kernel function. There are general purpose kernel functions and problem oriented kernel functions. In linear separable problems, the kernel functions map the training data with the high dimensional feature space. Decision function is a hyperplane in the feature space. The hyperplane for linearly separable pattern is given as

$$w^T x + b = 0 \quad (23)$$

where training set $T = \{(x_i, y_i)\}$, $x_i$ is the input variables and $y_i$ is the classification label. $w$ is the weight vector and $b$ is the perpendicular distance from the orgin to the hyperplane. Separation of positive training data from negative is given as in the mathematical form

$$w^T x + b \geq 1 \text{ for } y_i = +1 \quad (24)$$

$$w^T x + b < -1 \text{ for } y_i = -1 \quad (25)$$

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

Maximization of the Margin M is given as

$$M = \frac{2}{|w|} \tag{26}$$

Minimization of the margin is given as

$$M = \frac{1}{2} w^T w \tag{27}$$

In this work, BPNN and SVM is used to classify normal or abnormal cervical cancer from pap smear images.

## IV. RESULTS & DISCUSSION

Feature selection using SFFS method is performed for local shape features. Out of 32 features 18 features are chosen. The number of features in the global shape features is only 7. So, feature selection is not performed for these features. Further, the local and global shape features are combined as single feature vector. By applying feature selection using SFFS method, out of 39 features 29 features are selected. Classification is performed by BPNN and SVM classifiers using the local shape features, global shape features and combined shape features with reduced features separately. The performance measure of the classifiers are analyzed in terms of accuracy, sensitivity and specificity. The accuracy is given as

1) Accuracy: Accuracy is obtained by correctly classified images divided by the classified images.

$$\text{Accuracy} = \frac{TP + TN}{(TN + TP + FP + FN)} \tag{28}$$

where TP is positive cases that are correctly identified in the dataset, TN is negative cases that are correctly classified as negative in the dataset, FP is negative cases that are incorrectly classified as positive in the dataset and FN is positive cases that are wrongly classified as negative in the dataset.

Table 1 depicts the accuracy of BPNN and SVM classifiers for local, global and combined shape features.

Table 1 Accuracy of BPNN and SVM classifiers

| Sl.no | Type of features | BPNN(%) | SVM(%) |
|---|---|---|---|
| 1 | Local shape features | 86.79 | 89.25 |
| 2 | Global shape features | 84.44 | 87.42 |
| 3 | Combined shape features | 88.49 | 91.16 |

Fig. 6 shows the accuracy level of local, global and combined features.

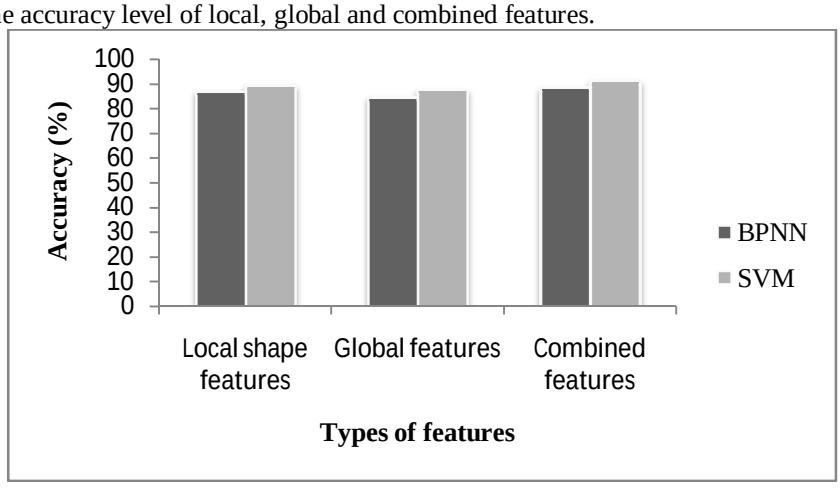



Fig. 6 Accuracy of features

2) Sensitivity: Sensitivity is obtained as correctly classified true positive rate divided by true positive and false negative samples.

$$Sensitivity = \frac{TP}{TP+FN} \tag{29}$$

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

**Table 2 Sensitivity of BPNN and SVM classifiers**

| Sl.no | Type of features | BPNN(%) | SVM(%) |
|---|---|---|---|
| 1 | Local shape features | 85.29 | 87.24 |
| 2 | Global features | 84.78 | 86.53 |
| 3 | Combined features | 88.46 | 90.88 |

Table 2 show the sensitivity of BPNN and SVM classifiers. The higher value of sensitivity is obtained as 90.88% from SVM classifier using combined shape features.

3) Specificity*:* Specificity is calculated as correctly classified true negative rate divided by the true negative and false positive samples.

$$Specificity = \frac{TN}{TN+FP} \tag{30}$$

**Table 3 Specificity of BPNN and SVM classifiers**

| Sl.no | Type of features | BPNN(%) | SVM(%) |
|---|---|---|---|
| 1 | Local shape features | 86.49 | 89.29 |
| 2 | Global features | 84.91 | 86.74 |
| 3 | Combined features | 88.64 | 90.24 |

Table 3 displays the specificity of BPNN and SVM classifiers. The SVM classifier yields better specificity as 90.24% using combined shape features than using the features explicitly.

## V.CONCLUSION

In this work, diagnosis of cervical cancer using shape features have been implemented. Local and global shape features are combined to yield better results. Feature selection is performed using SFFS method. The SVM and BPNN classifiers are applied. Performance measures such as accuracy, sensitivity and specificity are calculated. The results reveal that the SVM yields better performance than BPNN classifier as accuracy is 91.16%, sensitivity is 90.88% and specificity is 90.24%.

## REFERENCES

1. Anousouya Devi, M., Ravi, S., Vaishnavi, J. and Punitha.S (2016), Classification of Cervical Cancer Using Artificial Neural Networks, Procedia Computer Science, (Elsevier),Vol. 89, pp 465-472.
2. Ayubu Hassan Mbaga and Pei ZhiJun (2015), Pap Smear Images Classification for Early Detection of Cervical Cancer, International Journal of Computer Applications, Vol. 118, Issue No. 7, pp 10-16.
3. Barnali Sahua and Debahuti Mishrab (2012), A Novel Feature Selection Algorithm using Particle Swarm Optimization for Cancer Microarray Data, Procedia Engineering (Elsevier), Vol. 38, pp 27 – 31.
4. Bolon-Canedo, V., Sanchez-Marono, N., Alonso-Betanzos, A., Benitez, J.M. and Herrera, F. (2014), A review of microarray datasets and applied feature selection methods, Information Sciences, (Elsevier), Vol.282, pp 111-135.
5. Dibet Garcia-Gonzalez, Miguel Garcia-Silvente and Eugenio Aguirre (2016), A multiscale algorithm for nuclei extraction in pap smear images, Expert Systems with Applications, (Elsevier), Vol.64, pp 512-522.
6. Edwin Jayasingh Mariarputham and Allwin Stephen (2015), Nominated Texture Based Cervical Cancer Classification, Computational and Mathematical Methods in Medicine, Vol. 2015, Article ID 586928.
7. Hu, M.K., (1962), Visual Pattern Recognition by Moment Invariants, IRE Transactions on Information Theory, Vol. 8, Issue 2, pp. 179-187.
8. Jothi G. and Hannah Inbarani H. (2016), Hybrid Tolerance Rough Set–Firefly based supervised feature selection for MRI brain tumor image classification, Applied Soft Computing, (Elsevier),Vol. 46, pp 639-651.
9. Kavitha R and Vijayakumar B (2014), Identification of cervical malignancy and cervical cell categorization, International Journal of Management, IT and Engineering, Vol. 4, Issue 4, pp 99-107.
10. Konstantina Kourou, Themis P. Exarchos, Konstantinos P. Exarchos, Michalis V. Karamouzis and Dimitrios I. Fotiadis (2015), Machine learning applications in cancer prognosis and prediction, Computational and Structural Biotechnology Journal (Elsevier), Vol.13, pp 8–17.
11. Lili Zhao, Kuan Li, Mao Wang, Jianping Yin, En Zhu, Chengkun Wu, Siqi Wang and Chengzhang Zhu (2016), Automatic cytoplasm and nuclei segmentation for color cervical smear image using an efficient gap-search MRF, Computers in Biology and Medicine, (Elsevier), Vol.71, pp 46-56.
12. Loris Nanni, Sheryl Brahnam, Stefano Ghidoni and Emanuele Menegatti (2015), Improving the descriptors extracted from the co-occurrence matrix using preprocessing approaches, Expert Systems with Applications, (Elsevier),Vol. 42, Issue 22, pp 8989-9000.
13. Nagarajan,G., Minu, R.I., Muthukumar, B., Vedanarayanan, V. and Sundarsingh, S.D. (2016), Hybrid Genetic Algorithm for Medical Image Feature Extraction and Selection, Procedia Computer Science,(Elsevier), Vol. 85, pp 455-462.
14. Rabindra Kumar Singh and Dr. M. Sivabalakrishnan (2015), Feature Selection of Gene Expression Data for Cancer Classification: A Review, Procedia Computer Science (Elsevier), Vol.50, pp 52 – 57.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 4, Issue 11, November 2016**

15. Sayeda, Ahmed M., Eman Zaghloulb and Nassef, Tamer M. (2016), Automatic Classification of Breast Tumors Using Features Extracted from Magnetic Resonance Images,Procedia Computer Science(Elsevier) pp 392 – 398.
16. Siti Noraini Sulaiman, Nor Ashidi Mat-Isa, Nor Hayati Othman and Fadzil Ahmad (2015), Improvement of Features Extraction Process and Classification of Cervical Cancer for the NeuralPap System, Procedia Computer Science, (Elsevier), Vol.60, pp 750 – 759.
17. Sukumar P and Dr. Gnanamurthy R. K. (2015), Computer Aided Detection of Cervical Cancer Using Pap Smear Images Based on Hybrid Classifier, International Journal of Applied Engineering Research, Vol. 10, No. 8, pp. 21021-21032.
18. Thanatip Chankong, Nipon Theera-Umpon and Sansanee Auephanwiriyakul (2014), Automatic cervical cell segmentation and classification in Pap smears, Computer Methods and Programs in Biomedicine, (Elsevier), Vol. 113, Issue 2, pp 539-556.
19. Xiaoping Qiu, Ning Tao, Yun Tan and Xinxing Wu (2007), Constructing of the risk classification model of cervical cancer by artificial neural network, Expert Systems with Applications,(Elsevier), Vol. 32, Issue 4, pp 1094-1099.
20. Yessi Jusman, Siew Cheok Ng and Noor Azuan Abu Osman (2014), Intelligent Screening Systems for Cervical Cancer, The Scientific World Journal, Vol. 2014, Article ID 810368

## BIOGRAPHY

B. Ashok was born in 1969 in India. He received his M.Sc., (Computer Science) from Bharathidasan University,Tiruchirapalli and M.Phil (Computer Science) from Annamalai University. Presently he is working as a Assistant Professor/programmer in Annamalai University. His area of research is Image processing and Data Mining. He has twenty years of experience in Application Programming. He is involved in research activities for the past seven years.

Dr.S.Anu H Nair is an Assistant Professor in the Department of Computer Science and Engineering of Annamalai University. She received her Bachelor degree in 2002, Master Degree from Manonmaniam Sundaranar University in 2005 and the Ph.D degree from Annamalai University in 2016. Her main research areas include image processing, pattern classification techniques and neural networks. She has published ten research papers in International conferences and national conferences. She has published several research papers in International journals and a book chapter.

Dr.P.Aruna was born in 1968 in India. She received her B.E. from Madras University, M. Tech from IIT Delhi and the Ph.D degree from Annamalai University. Presently she is working as a Professor in the Department of Computer Science and Engineering of Annamalai University. She has published 80 research papers in International Journals and Conferences and 26 research papers in National Journals and Conferences. She has twenty five years of teaching experience and sixteen years of research experience. She has published 3 book chapters. Her area of specialization includes Neural networks & Fuzzy systems, Data Mining and Image processing. She has guided 6 Ph.D scholars.

N. Puviarasan was born in 1963 in India. He received his M.S(Soft Sys) from BITS, Pilani, M.S (Engg) from Annamalai University. Presently he is working as an Associate Professor in the Department of Computer Science and Engineering of Annamalai University. He has published 25 research papers in International journals and conferences and 13 research papers in National journals and conferences. He has twenty seven years of teaching experience and thirteen years of research experience. He has published one book chapter. His area of specialization includes Neural networks and Fuzzy systems, Data mining and Image Processing.