RESEARCH ARTICLE

Severity Analysis of Cervical Cancer in Pap Smear Images by using EEETCM, ERSTCM & CFE method based Texture Features and Hybrid Kernel based Support Vector Machine Classifier.

S. Athinarayanan[1] and Dr. M. V. Srinath[2].
1.   Research Scholar, Manonmaniam Sundaranar University, Tirunelveli, Tamilnadu, India.
2.   Director, Department of MCA, STET Women's College, Mannargudi.

………………………………………………………………………………………………….....

Manuscript Info	*Abstract*
…………………..	………………………………………………………………………

Classification of medical imagery is a difficult and challenging process due to the intricacy of the images and lack of models of the anatomy that totally captures the probable distortions in each structure. Cervical cancer is one of the major causes of death among women in worldwide. Proper and timely diagnosis can prevent the life to some level. Due to its importance, the aim of the paper is to investigate about the classification of Abnormal Cell of the pap smear image by using individual and combining individual feature extraction method with the classification technique. In this paper three feature extraction methods were used: From that three, two were individual feature extraction methods namely Effective Extending Enriched Texton Co-Occurrence Matrix (EEETCM) & Enriched Rough Set Texton Co-Occurrence Matrix (ERSTCM) and remained one was combining individual feature extraction method named asConcatenated Feature Extraction (CFE). The CFE method represents all the individual feature extraction method of EEETCM & ERSTCM features are combining together as one feature to assess their joint performance. Then these three feature extraction methods are tested over Hybrid Kernel based Support Vector Machine (HKSVM) Classifier. This Examination was conducted over a set of single cervical cell based pap smear images. The dataset contains four classes of images, with a total of 512 images. The distribution of number of images per class is not uniform. Then the performance was evaluated inboth the individual and combining individual feature extraction method with the classification techniques by using the statistical parameters of sensitivity, specificity & accuracy. Hence the resultant value of the statistical parameters described in individual feature extraction method with the classification technique, proposed ERSTCM+HKSVM Classifierhad given the better results than the other EEETCM+HKSVM Classifier and combining individual feature extraction method with the classification technique described, proposed CFE+HKSVM Classifier had given the better results than other EEETCM+HKSVM & ERSTCM+HKSVM classifiers.

………………………………………………………………………………………………….....

Corresponding Author:- S.Athinarayanan.
Address:-Research Scholar, Manonmaniam Sundaranar University, Tirunelveli, Tamilnadu, India.

Introduction:-

Cancer is incurable disease all around the world [1-3]. So many researchers, pathologists have found many number of methodologies to cure the cells which can be affected by cancer. In earlier days so many methodologies are available to detect the symptoms of cancer [4-6]. All around the world women were affected by two types of cancer viz., breast and cervical cancer [7,8]. Un controlled growth of cells in the part of breast is breast cancer [9].

The cells in part of the Cancer that forms in tissues of the cervix is cervical cancer [10]. Cervical Cancer cells originates from pre-cancerous, benign lesions in the uncontrolled cells which are present inside the cervix. The results interpreted by World Health Organization (WHO), the initial stage for the development of cervical cancer is mild dysplasia, thereafter its moderate dysplasia, severe dysplasia, finally leads to carcinoma in situ (CIS) as well as invasive cervical cancer [11,12].

Pap Smear is the most popular screening method to detect cervical cancer from cervicted portion (leisons). To detect the portion which is affected by cervical cancer, it requires a recognized laboratory, trained cytologist and their repeatable visited reports used for evaluate the results[13]. In order to detect cervical cancer automatically with the help of screening tests. Among them one of the most important method is the segmentation of cell nuclei from the stained specimens [14]. However, isolated nuclei of the cells in high-quality acquisitions provide difficult tricks in the segmentation is more number of nuclei with various characteristics under occurring different acquisition conditions in good-resolution scans of the complete microscope slide[15]. Thresholding is the most important method for segmentation process[16] and also it is an easiest method to convert gray scale image into binary image based on the global or local threshold value[17,[18]. Bi-level thresholding is a method which must be classifies the pixels into two groups. First one represents in which the pixels with gray levels are lie above a certain threshold value, whereas the other in which the pixels are lie below a certain threshold value [19,20].

Binaryclassification is one of the method which classifies the elements of a given cell into two groups based on the classification rule. Binary classification is one of the characteristic method used to determine if the patient which has been affected by the cancer[21]. Principally classification process focussed the detection of cancer with the help of pap smear screening test results. But, it is difficult to detect the critical stage of the cancer from the pap smear screening test images.

The Pap smear reporting classification has evolvedand been refined over time. In this present work, according to the WHO descriptive histological classification of the abnormal squamous cell of the image as mild dysplasia, moderate dysplasia, severe dysplasia and Carcinoma in Situ(CIS) [22] by using effective texture based feature extraction and classification methods. As a results produced from the proposed system we have determine whether the cell is become which stage of the cancer.

Related Works:-

First order statistical features and Second order grey level co-occurrence matrix (GLCM) with the SVM classifier was predict the treatment outcome of the patient is cured or relapsed from cervical cancer disease [23]. Higher-order statistical approaches have been successfully used in textural recognition. The author applied both low and high order statistical features to extract thecell texture, including maximum, minimum, range, edian, mean, standard deviation,energy, skewness, kurtosis and entropy for detecting cervical cancer [24]. A combination of features are calculated for a single stage of cancer and depending upon the number of positives we can predict for certain whether the cervical cancer has progressed to a particular stage or not [25].

Srinath et al. proposed or Computer aided detection for classifying the cervical cell as normal or abnormal based on the extracted selective feature of the nucleus like area value. Then the results were represented a given cell was normal means we come to a conclusion, that cell having the particular patient was not affected by the cervical cancer. But the results represented a given cell was abnormal means the particular patient was affected by the cervical cancer. Not only to end my proposed system whether the patient is affected by this cervical cancer, but also it is used for finding the severity of the cancer based on the above said same feature (area) up, down tuned values of the nucleus. According to the WHO classification of the pap smear image, abnormal cell was further classified into mild, moderate, Severe or CIS dysplasia respectively [26].Srinath et al. proposed test results shows the mean and area features based cell classification done as normal or abnormal cell and also the result would be abnormal means further the cell was used for identifying the stage of the cervical cancer. Hence the results would be more useful for the pathologists to reduce their work load and minimize the human error while improving the accuracy of diagnosis

[7]. The rest of the paper is organized as follows: Our proposed multi class cancer classification system is presented in section 2. The detailed experimental results and discussions are given in section 3, while the conclusion is summarized in section 4.

Cervical Cancer Severity Classification System:-

Current manual screening methods are costly and sometimes result in inaccurate diagnosis caused by human error. The introduction of machine assisted screening will bring significant benefits   to   the community, which can   reduce financial costs and increase screening accuracy**.** In this research article, we have developed, Severity class of cervical cancer classification system based on individual and combining individual texture features and hybrid kernel based support vector machine using pap smear images. Two major contribution of the proposed system is feature extraction and feature classification.

Feature extraction:-

The purpose of feature extraction is to reduce the original data set by measuring certain properties, or features, that distinguish one input pattern from another pattern. The extracted feature is expected to provide the characteristics of the input type to the classifier by considering the description of the relevant properties of the image into a feature space. In this paper, we proposed three novel feature extraction methods. From that three, two were individual feature extraction methods, they are Effective Extending Enriched Texton Co-Occurrence Matrix (EEETCM) and Enriched Rough Set Texton Co-Occurrence Matrix (ERSTCM) and the remained one was combining individual feature extraction features method named as Concatenated Feature Extraction (CFE). The CFE method represents all the individual feature extraction methods of EEETCM &ERSTCM features are combining together to one feature to assess their joint performance. The types of individual and combining individual feature extraction features method is given below.

Individual Feature Extraction Methods:
- Computation of Feature Vector F(V1) using EEETCM.
- Computation of Feature Vector F(V2) using ERSTCM.

Combining Individual Feature Extraction Features Method:
- Computation of Feature Vector F(V3) using CFE.

The detailed Process of the above feature extraction method is given below.

Individual Feature Extraction Methods:
Computation of Feature Vector F(V1) using EEETCM:
In this method will be contributed to improve EETCMby adding new textons to detect information that lost. sothereare6textons.Thenew textons are horizontal bottom and vertical right. The aim of this method is addition to prevent the loss of information when pixels are co-occurred on horizontal bottom and vertical right.

Texton Detection:-

The Texton template defined in EEETCM are different from those in EETCM. In this method, six special texton types are used. defined on a 2 x 2 grid, as shown in figure.1(a-e). Denote the four pixels as V1, V2, V3 and V4. If the two pixels highlighted in the gray color have the same value, the grid will form a texton. Those 6 texton types are denoted as T1, T2, T3, T4, T5 & T6, respectively. From this 6 textons, four textons (T1,T2,T3 & T4) are used for ETCM [ ] method. T5 & T6 additionally added new textons.

The Working Mechanism of Texton detection is illustrated in figure.2. In the final segmented nucleus image, we move 2 x 2 block from left to right and top to bottom throughout the image to detect textons with 2 pixel as step length. If a texton is detected, the detected pixel pair values in the 2 x 2 grids are kept unchanged. Otherwise it will have the zero value. Finally we will obtain a Texton image, denoted by T(x,y). This process results was shown in figure.2(a-e). After the formation of final texton image, the feature vector F(V1) (Five features such as, ASM, entropy, IDM, contrast and Maximumprobability) is extracted from it.
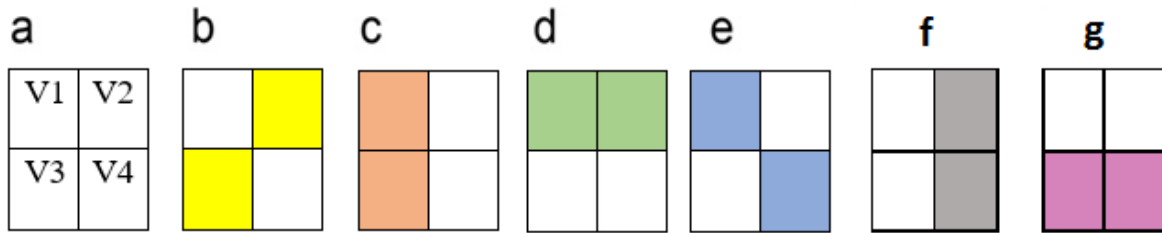
Figure.1. Special texton types of EEETCM (a) original 2x2 matrix (b)Texton T1 (c) Texton T2 (d) Texton T3 (e) Texton T4 (f) Texton T5 (g) Texton T6.
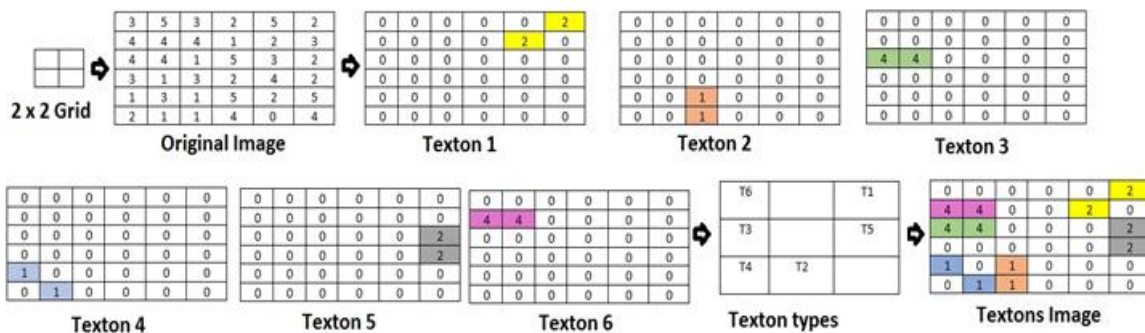

Figure.2: Texton image formation process using EEETCM (a) 2x2 Matrix (b) Original image intensity value (c) Texton location of the original image (d) Four texton types (e) Final texton image of EEETCM.

The six texton types used in EEETCM contain rich information than those in EETCM because the co-occurrence probability of two same-valued pixels is bigger than that of three or four same-valued pixels in a 2 x 2 grid. As for the texton detection procedure, the computational complexity of proposed EEETCM is slightly slower than EETCM because of additionally added 2 textons, but in terms of the effective texture information extraction the proposed EEETCM will better than EETCM.

Computation of Feature Vector F(V2) using ERSTCM:-
As per Julesz [27] description a texton is a pattern which is shared by an image as a common property all over the image. Textures will be formed only if the adjacent elements lie within the neighbourhood. Texton Image has the discrimination power of color, texture and shape features. Based on the Julez's [27] textons theory the texton co-occurrence matrices (TCM) algorithm, can describe the spatial correlation of textons for image retrieval. With the use of limited number of selected pixels the algorithm computes different features. Extraction of effective texture information will be increased by calculating texton using Rough texture. This method is represented Enriched Rough Set Texton Co-occurrence Matrices. In this Method. As a texture descriptor, good retrieval accuracy can be achieved especially for directional textures. The texton can be inclined with the use of critical distances between texture elements. Texture element size determines the critical distance. Texture can be resolved into minute units, like orientation, Texton classes of colors, elongated blobs of specific width, aspect ratios and terminators of elongated blobs. If texture elements are expanded to a large extent in one orientation discrimination reduces. If the elongated elements are not jittered in orientation texture gradients increase at boundaries. Hence by using a sub image of size 3 x 3 a texton gradient can be obtained. Here 20 textons of 3 x 3 grids are proposed as shown in figure 3. Even the co-occurrence probability of same valued pixels in 3 x 3 grids is smaller than that of 2 x2 grid, but the textons developed using 2 x 2 grid may not give complete information regarding direction. The computational complexity for using the overlapped components of 20 textons is also less to obtain final texton image. The 20 textons of 3 X 3 grid can detect textons in all directions and also the corners of the textures. If three pixels are highlighted and have same value then, grid will form a texton as shown in Figure 4.
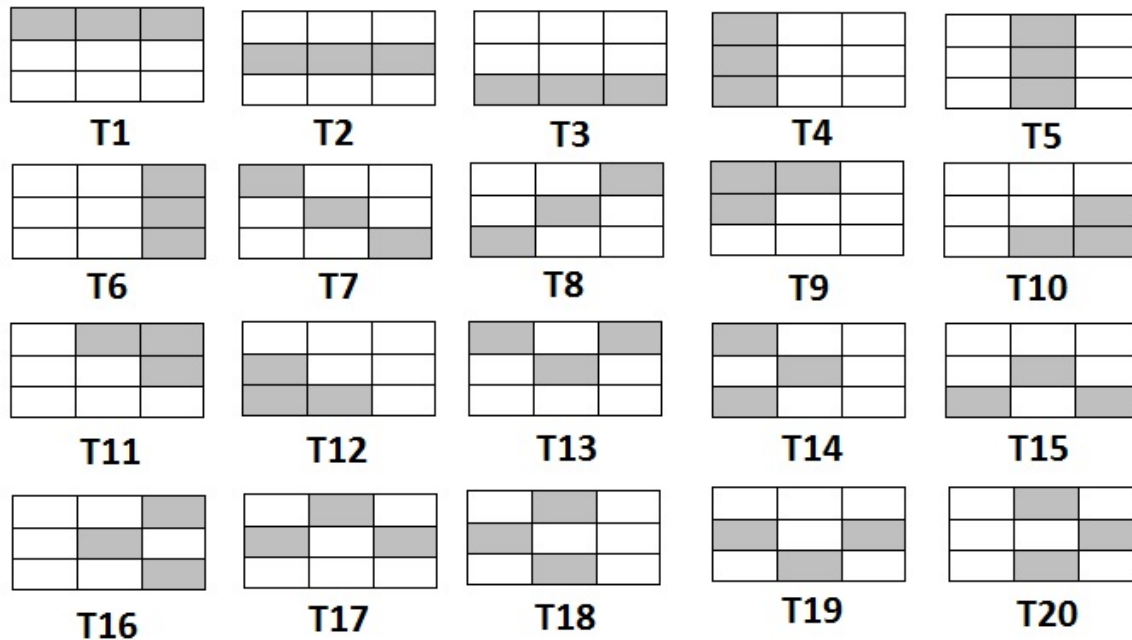
Figure.3 : 20 types of 3 x 3 Textons

Figure.4 : Texton Detection Process

Texton Detection:-

The Working Mechanism of Texton detection is illustrated in figure.4. In the final segmented nucleus image, we move 3 x 3 block from left to right and top to bottom throughout the image to detect textons with 3 pixel as step length. If a texton is detected, the same value of the three pixels in the 3 x 3 grids are kept unchanged and the remained pixels are zero value. Finally we will obtain a Texton image, denoted by $T(x,y)$. After the formation of final texton image, the feature vector $F(V2)$ (Seven features such as contrast, inverse difference moment, correlation, variance, cluster shade, cluster prominence and Homogeneity) is extracted from it.

Difference between ERSTCM and EEETCM:-
Texton Detection and Computational Complexity:-

Even thoughthe occurrence probability of same valued pixels in 3 x 3 grids in ERSTCM is smaller than the 2 x2 grid in EEETCM, but the 20 types of texton detection in ERSTCM contains rich and complete information than the six types of Texton detection in EEETCM method regarding all directions and corners of texture. This result was shown in the following figure.5.Then the computational complexity of the ERSTCMmethod is also less to obtain final texton image because the texton detected step length of this method is three pixels, when compared to the EEETCM method step length of two pixels.
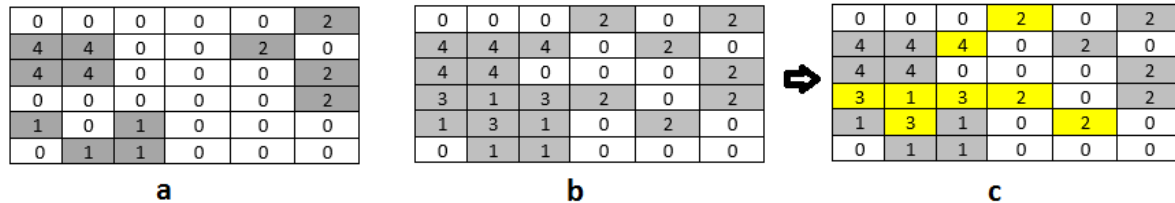
Figure.5. The Result of Texton detection a) EEETCM b) ERSTCM c) Difference between EEETCM & ERSTCM.

Combining Individual Feature Extraction Features Method:-
Computation of Feature Vector F(V3) using CFE:-
The name CFE Stands for Concatenated Feature Extraction. The CFE method represents all the individual feature extraction methods of EEETCM & ERSTCM features are combining together to one feature to assess their joint performance. Hence, total Feature vector uses F(V3) = F(V1) + F(V2) dimensional vector as the concluding image features in cervical cell classification.

Table.1. Summary of Texture Features:-
The three feature extraction methods and its features details are given in the following table.1.

Feature category	Features	Number of features
EEETCM	ASM, Entropy, IDM, Contrast &Maximum Probability.	5 features
ERSTCM	Contrast, Inverse difference moment, correlation, variance, cluster shade, cluster prominence and Homogeneity	7 features
CFE	The features of EEETCM & ERSTCM	12 features

Table.1. Features Summary Table.

Nucleus feature Classification:-

The individual feature extraction methods of EEETCM & ERSTCM &combining individual feature extraction features method (CFE) features are tested over Hybrid Kernel based Support Vector Machine (HK-SVM) classifier.The result of this classifier represent classifies the abnormal cell of the pap smear image into any one of the classes of mild dysplasia, moderate dysplasia, severe dysplasia and Carcinoma in Situ (CIS). This Examination was conducted over a set of single cervical cell based pap smear images. The dataset contains four classes of abnormal cell images, with a total of 512 images. From that 512 images, mild dysplasia -107, moderate dysplasia -124, severe dysplasia - 139, & Carcinoma in Situ (CIS) were 142 images.

Feature Classification using HK-SVM:-
After feature extraction process, In-order to detect the severity of the cancer from the pap smear images, we perform the classification step. In our Proposed method, we have developedHybrid kernel based SVM method for classifyingthe abnormal cell of the pap smear image intoany one of the severity classes of mild, moderate, severe or CIS dysplasia. In computer science, such type of the classification process explained as following two phases.
There are two phases in the support vector machine namely,
(1) Training phase and (2) Testing phase.

Training Phase:-
The features of the three methods of EEETCM, ERSTCM and CFE for the corresponding number of the training images are given as input to the training phase. The input function gives the set of values which are non-separable. All the possible separations of the point set can be achieved by a hyperplane. For that, a set of data drawn from an unknown distribution, $((x1, y1), . . . , (x1, yl), xi) \in Rn, yi \in \{-1, 1\}$ and also a set of decision functions, or hypothesis space $f\lambda : \lambda \in \Lambda$ are given, where Λ (an index set) is a set of abstract parameters, not necessarily vectors. $f\lambda : Rn \to \{-1, +1\}$ is also called a hypothesis.

The set of functions $f\lambda$ could be a set of Radial Basis Functions or a multilayer neural network. All the possible separations of the point set can be achieved by a hyperplane. In the Lagrange optimization formulation, we can find the optimal separating hyperplane normal vector,

A kernel is any function K : Rn × Rn → R. This corresponds to a dot product for some feature mapping K(X1, X2) = φ(X1) · φ(X2) for some φ The kernel function can directly compute the dot product in the higher dimensional space. Introduce kernel-based Lagrange multipliers αi ≥ 0 ∀i

$$L_p = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{i=1}^{n} \alpha_i \alpha_j y_i y_j K(x_i . x_j)$$

---------- (1)

Minimize L p with respect to w, b and maximize with respect to αi.
In a convex quadratic programming problem, the plane is a nonlinear combination of the training vectors

$$w = \sum_{i=1}^{n} \alpha_i y_i K(x_i)$$

------------- (2)

Thus, the hyperplane is separated into two clusters. The sample representation of this process is shown in figure. 5.

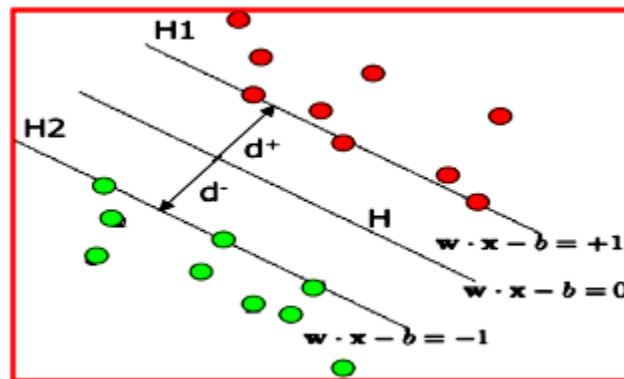

Fig. 5. Sample representation of separating optimal hyperplane

Testing Phase:-
The features of the three methods of EEETCM, ERSTCM and CFE for the corresponding number of the training images are given as input to the testing phase.
The class of an input data x is then determined in Eq.(3)

$$\text{class}(x) = \text{sign}(\varphi(x) \cdot w - b) = \text{sign}\left(\sum y_i \lambda_i \varphi(x_i) \cdot \varphi(x) - b\right)$$

------ (3)

We have analyzed the kernel equation from the existing work [28] and used them in the proposed work namely, RBF and polynomial function. Radial Basis Function The support vector will be the center of the RBF and σ will determine the area of influence. This support vector has the data space.

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$$

Polynomial Kernel Function Polynomial kernels are of the form K(x,z) = (1 + xTz)d .
Where d = 1, a linear kernel and d = 2, a quadratic kernel are commonly used.
Let k1(RBF) and k2(Polynomial) be kernels over Ξ ×Ξ,Ξ ⊆ Rp, and k3 be a kernel over Rp × Rp. Let function ϕ : Ξ → Rp. The four kernel-based formulations are represented by

- $k(x, y) = k_1(x, y) + k_2(x, y)$ is a kernel   (i)
- $k(x, y) = \alpha k_1(x, y)$ is a kernel, when $\alpha > 0$ (ii)
- $k(x, y) = k_1(x, y) k_2(x, y)$ is a kernel   (iii)
- $k(x, y) = k_3(\varphi(x)\varphi(y))$ is a kernel   (iv)

Substitute the four equations (i) – (iv) in Lagrange multiplier Eq.(1) and get the proposed hybrid kernel. It is exposed in Eq.(4).

$$L_p = \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{i=1}^{n}\alpha_i\alpha_j y_i y_j (k_1(x_i.x_j) + k_2(x_i x_j))$$

$$L_p = \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{i=1}^{n}\alpha_i\alpha_j y_i y_j a k_1(x_i x_j)$$

$$L_p = \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{i=1}^{n}\alpha_i\alpha_j y_i y_j k_1(x_i.x_j) k_2(x_i x_j)$$

$$L_p = \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{i=1}^{n}\alpha_i\alpha_j y_i y_j k_3\varphi((x_i.x_j)\varphi(x_i x_j))$$

------- (4)

Substitute the four theorems in Quadratic function equation (2) and get,

$$w = \sum_{i=1}^{n} \alpha_i y_i (K_1(x_i) + k_2(y_i))$$

$$w = \sum_{i=1}^{n} \alpha_i y_i \alpha K_1(x_i y_i)$$

$$w = \sum_{i=1}^{n} \alpha_i y_i \alpha K_1(x_i) k_2(y_i)$$

$$w = \sum_{i=1}^{n} \alpha_i y_i K_3\varphi(x_i)\varphi(y_i)$$

---------------------- (5)

Hence, finally HKSVM Process used to test the features to identify the abnormal cell of the pap smear image into any one of severity classes of cervical cancer.

Experimental Results and Discussion:-

Experimental image Data set. :-

For any machine learning algorithm, the database with which it is trained plays an important role. It is said that a machine can be made to learn and reproduce any human behaviour, provided it is trained with suitably precise database. The database prepared in this work consists of four classes of abnormal cell based 512 single cervical cell presented in pap smear images. These Images were collected from Muthamil Hospital, Tirunelveli District, Tamilnadu State, India in 2013 and these images were taken with 100X lens magnification using Olympus ch20i Microscope. Each image was examined and diagnosed by pathologists of that hospital before being used as reference for this study. The Sample Image data sets are shown in the table.2. The sample Pap smear experimental images of various classes is shown in Figure 6. The data set contains 512 images and these images were histologically diagnosed and graded based on World Health Organization (WHO) criteria as the following abnormal class distribution:

- Mild dysplasia, 107 images.
- Moderate dysplasia, 124 images.
- Severe dysplasia, 139 images.
- Carcinoma in Situ. 142 images.

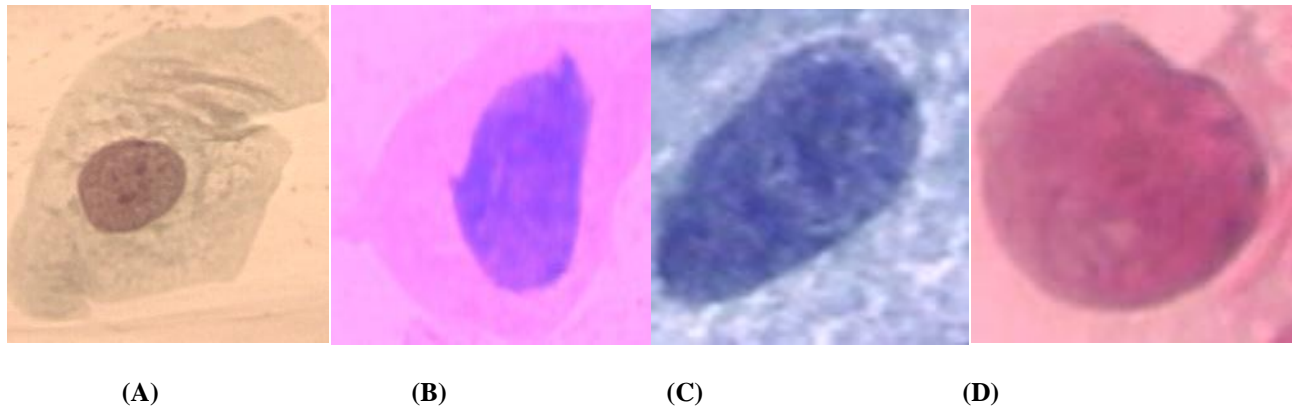|			
(A)	**(B)**	**(C)**	**(D)**

Figure.6. Some of the cells found in cervix: (A-C) mild, moderate and severe dysplasia,(D) Carcinoma in situ

The Implementation was done in the tool of matlab.

Experimental Results& Comparative Analysis:-
This section describes the experimental results of the proposed classification method using Pap smear images with different types of cervical cancer. In the proposed method, the experimental image data set is divided into two sets such as training set and testing set. The details of this set was shown in table.2. The classifiers are trained with the training images and the classification accuracy is calculated only with the testing images. In the testing phase, the testing dataset is given to the proposed technique to find the cancers type in smear images and the obtained results are evaluated through evaluation metrics namely, sensitivity, specificity and accuracy [29], it is given by (eqn.6-8)

$$Sensitivity = TP/(TP + FN) \text{ ----------(6)}$$

$$Specificity = TN/(TN + FP) \text{ ----------(7)}$$

$$Accuracy = (TN + TP)/(TN + TP + FN + FP) \text{ --------(8)}$$

Where TP corresponds to True Positive, TN corresponds to True Negative, FP corresponds to False Positive and FN corresponds to False Negative. These parameters for a specific category, say, Mild dysplasia are as follows: TP is True Positive (an image of 'Mild dysplasia' type is categorized correctly to the same type), TN = True Negative (an image of 'Non- Mild dysplasia' type is categorized correctly as 'Non- Mild dysplasia' type), FP =False Positive (an image of 'Non- Mild dysplasia' type is categorized wrongly as 'Mild dysplasia' type) and FN is False Negative (an image of 'Mild dysplasia' type is categorized wrongly as 'Non- Mild dysplasia' type).'Non- Mild dysplasia' actually corresponds to any of the three categories other than 'Mild dysplasia'. Thus, 'TP & TN' corresponds to the correctly classier images and 'FP & FN' corresponds to the misclassified images.

Table. 2. Experimental image dataset for classification

Cancer Type	Training data	Testing data	Total no of images
Mild dysplasia	50	57	107
Moderate dysplasia	50	74	124
Severe dysplasia	50	89	139
carcinoma in situ	50	92	142
Total	**200**	**312**	**512**

Individual feature Extraction and Classification Combination Results and Comparison: -
Thus, different parameter values are obtained for each class and also for the different classifiers. These parameters are estimated from the confusion matrix which provides the details about the false and successful classification of images from all categories for each classifiers. The confusion matrix of the EEETCM+HKSVM method is illustrated in Table.3.

Table 3 Confusion matrix of proposed method (EEETCM+HKSVM)

Class predicted	Ground Truth Class (Assigned by Pathologist)			
	Mild dysplasia	Moderate dysplasia	Severe dysplasia	Carcinoma in situ
Mild dysplasia	45	4	3	5
Moderate dysplasia	4	62	6	2
Severe dysplasia	3	4	78	4
Carcinoma in situ	5	4	3	80

In the Table 3, the row-wise elements correspond to the four categories and the column-wise elements correspond to the target class associated with that abnormal category. Hence, the number of images correctly classified (TP) under each category is determined by the diagonal elements of the matrix. The row-wise summation of elements for each category other than the diagonal elements corresponds to the 'FN' of that category. The column-wise summation of elements for each category other than the diagonal element corresponds to the 'FP' of that category. Similarly, 'TN' of the specific category is determined by summing the elements of the matrix other than the elements in the corresponding row and column of the specific category.

For example, among the 57 Mild dysplasia testing images, 45 images have been successfully classified (TP) and the remaining 12 images (first row-wise summation) have been misclassified to any of the non- Mild dysplasia categories (FN). Similarly, 12 images (first column-wise summation) from the other three categories (non-Mild dysplasia) have been misclassified as Mild dysplasia category (FP). In the Table 4, the classification accuracy of EEETCM with HKSVM in class 1(Mild dysplasia) type cancer is 92.31%, class 2(Moderate dysplasia) is 92.31%, class 3(Severe dysplasia) is 92.63% and class 4(carcinoma in situ) is 92.63%.

The miss classification rate of class 1(Mild dysplasia) and class 2(Moderate dysplasia) type cancer is high compared to the other two classes.

Table 4 Performance measure of EEETCM with HKSVM

Class predicted	TP	TN	FP	FN	Sensitivity (%)	Specificity (%)	Accuracy (%)
Mild dysplasia	45	243	12	12	78.95	95.29	92.31
Moderate dysplasia	62	226	12	12	83.78	94.96	92.31
Severe dysplasia	78	211	12	11	87.64	94.62	92.63
carcinoma in situ	80	209	11	12	86.96	95.00	92.63
Overall Sensitivity, Specificity & Accuracy Results					84.93	94.97	92.46

The Performance of overall classifier sensitivity, specificity and accuracy was determined by using the following equations (eqn: 9 -11).

$$Overall classification Sensitivity = \left(\frac{\sum_i TP(Class(i))}{\sum_i TP(class(i)) + FN(class(i))} \right) * 100 \text{ ------- (9)}$$

$$Overall\ classification\ Specificity = \left(\frac{\sum_i TN(Class(i))}{\sum_i TN(class(i)) + FP(class(i))} \right) * 100 \text{ ----------(10)}$$

$$Overall\ classification\ accuracy = \left(\frac{\sum_i TP(Class(i) + TN(Class(i))}{\sum_i TP(Class(i)) + TN(Class(i)) + FP(Class(i)) + FN(Class(i))} \right) * 100 \text{----(11)}$$

Based on the similarly approach of the Table.3, the confusion matrix of the ERSTCM+HKSVM method is illustrated in Table 5.

Table 5 Confusion matrix of proposed method (ERSTCM+HKSVM)

Class predicted	Ground Truth Class (Assigned by Pathologist)			
	Mild dysplasia	Moderate dysplasia	Severe dysplasia	Carcinoma in situ
Mild dysplasia	47	2	3	5
Moderate dysplasia	3	64	3	4
Severe dysplasia	2	2	81	4
Carcinoma in situ	3	3	3	83

In the Table 5, the row-wise elements correspond to the four categories and the column-wise elements correspond to the target class associated with that abnormal category. Hence, the number of images correctly classified (TP) under each category is determined by the diagonal elements of the matrix. The row-wise summation of elements for each category other than the diagonal elements corresponds to the 'FN' of that category. The column-wise summation of elements for each category other than the diagonal element corresponds to the 'FP' of that category. Similarly, 'TN' of the specific category is determined by summing the elements of the matrix other than the elements in the corresponding row and column of the specific category.

For example, among the 57 Mild dysplasia testing images, 47 images have been successfully classified (TP) and the remaining 10 images (first row-wise summation) have been misclassified to any of the non- Mild dysplasia categories (FN). Similarly, 8 images (first column-wise summation) from the other three categories (non-Mild dysplasia) have been misclassified as Mild dysplasia category (FP). In the Table 6, the classification accuracy of ERSTCM with HKSVM in class 1(Mild dysplasia) type cancer is 94.23%, class 2(Moderate dysplasia) is 94.55%, class 3(Severe dysplasia) is 94.55% and class 4(carcinoma in situ) is 92.95%. The miss classification rate of class 4 (Carcinoma in situ) type cancer is high compared to the other three classes.

Table 6 Performance measure of ERSTCM with HKSVM

Class predicted	TP	TN	FP	FN	Sensitivity (%)	Specificity (%)	Accuracy (%)
Mild dysplasia	47	247	8	10	82.46	96.86	94.23
Moderate dysplasia	64	231	7	10	86.49	97.06	94.55
Severe dysplasia	81	214	9	8	91.01	95.96	94.55
carcinoma in situ	83	207	13	9	90.22	94.09	92.95
Overall Sensitivity, Specificity & Accuracy Results					88.14	96.04	94.22

For comparative analysis, the proposed cervical cancer classification system (ERSTCM + HKSVM) is compared to other classification system (EEETCM + HKSVM). The overall classification performance measured parameters of sensitivity, specificity & accuracy of the proposed method is 88.14%, 96.04% & 94.22% and this method will be provide the better results than the existing method EEETCM + HKSVM results in terms of 84.93%, 94.97% & 92.46%. The overall classification results of sensitivity, specificity and accuracy of existing and proposed method are shown in **Figure 7**.
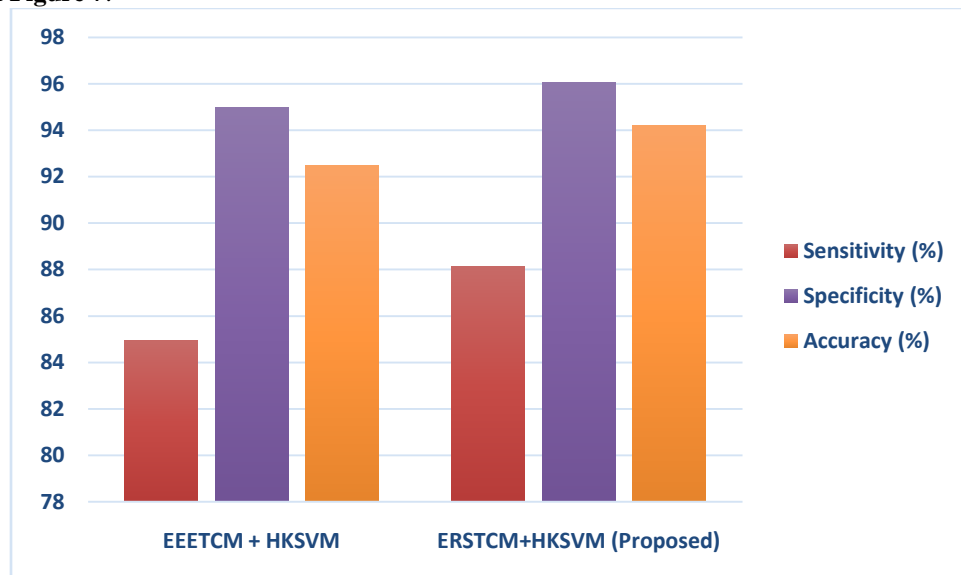

Figure.7. Comparison Results of the Proposed & Existing System in individual feature Extraction methods with Classification Technique.

Individual and Combining individual features & Classification combination comparison:-
Based on the similarly approach of the Table.3, the confusion matrix of the ERSTCM+HKSVM method is illustrated in Table 7.

Table 7 Confusion matrix of proposed method (CFE+HKSVM)

Class predicted	Ground Truth Class (Assigned by Pathologist)			
	Mild dysplasia	Moderate dysplasia	Severe dysplasia	Carcinoma in situ
Mild dysplasia	53	1	2	1
Moderate dysplasia	1	69	2	2
Severe dysplasia	2	2	84	1
Carcinoma in situ	1	1	2	88

In the Table 7, the row-wise elements correspond to the four categories and the column-wise elements correspond to the target class associated with that abnormal category. Hence, the number of images correctly classified (TP) under each category is determined by the diagonal elements of the matrix. The row-wise summation of elements for each category other than the diagonal elements corresponds to the 'FN' of that category. The column-wise summation of elements for each category other than the diagonal element corresponds to the 'FP' of that category. Similarly, 'TN' of the specific category is determined by summing the elements of the matrix other than the elements in the corresponding row and column of the specific category.

For example, among the 57 Mild dysplasia testing images, 53 images have been successfully classified (TP) and the remaining 4 images (first row-wise summation) have been misclassified to any of the non- Mild dysplasia categories (FN). Similarly 4 images (first column-wise summation) from the other three categories (non-Mild dysplasia) have been misclassified as Mild dysplasia category (FP). In the Table 8, the classification accuracy of CFE with HKSVM in class 1(Mild dysplasia) type cancer is 97.44%, class 2(Moderate dysplasia) is 97.12%, class 3(Severe dysplasia) is 96.47% and class 4(carcinoma in situ) is 97.44%. The miss classification rate of class 2(Moderate dysplasia) and class 3(Severe dysplasia) type cancer is high compared to the other two classes. The individual class and overall class accuracy calculation procedure are given in Eqn(10) and (11).

Table 8 Performance measure of CFE with HKSVM

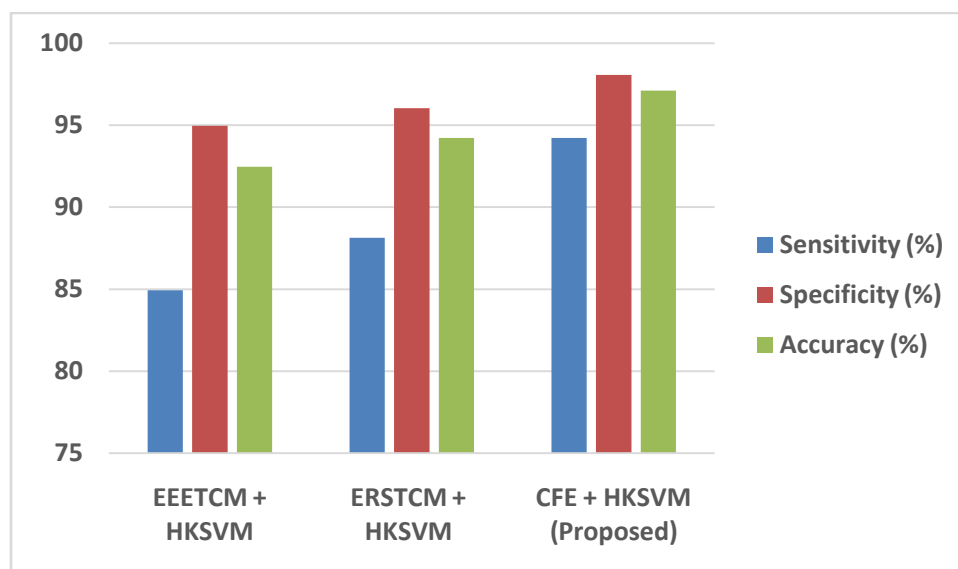| Class predicted	TP	TN	FP	FN	Sensitivity (%)	Specificity (%)	Accuracy (%)
Mild dysplasia	53	251	4	4	92.98	98.43	97.44
Moderate dysplasia	69	234	4	5	93.24	98.32	97.12
Severe dysplasia	84	217	6	5	94.38	97.31	96.47
carcinoma in situ	88	216	4	4	95.65	98.18	97.44
Overall Sensitivity, Specificity & Accuracy Results					94.23	98.07	97.11

Figure.8. Comparison Results of the Proposed & Existing System in Combining individual feature Extraction methods with Classification Technique.

For comparative analysis, the proposed cervical cancer classification system (CFE + HKSVM) is compared to the other methods (EEETCM + HKSVM & ERSTCM + HKSVM), the overall classification performance measured parameters of sensitivity, specificity &accuracy of the proposed method is 94.23%, 98.07%& 97.11% and this method produced better results than the existing methods of EEETCM + HKSVM & ERSTCM + HKSVM Results. The overall classification results of sensitivity, specificity and accuracy of existing and proposed method are shown in **Figure 8**.

Conclusion:-

In this paper, a novel approach of individual and combining individual feature extraction method with classification technique developed for identify the severity stage for the abnormal cell of the pap smear image. Two major contributions of this paper are feature extraction and classification. In feature extraction, we have taken the advantage of combining individual methods such as EEETCM & ERSTCM texture features as one feature to assess the joint performance. In Classification, multiple kernels are combined and developed a hybrid kernel based SVM classifier for improving the classification process. For comparative analysis, our proposed approach is compared with existing works in individual feature extraction with the classification method, the proposed ERSTCM+HKSVM produce better result than EEETCM+HKSVM and combining individual feature extraction with the classification method, the proposed CFE+HKSVM produce better result than EEETCM+HKSVM & ERSTCM+HKSVM methods in terms of the statistical parameters results sensitivity, specificity & accuracy. Hence, finally our proposed individual and combining individual method are proved good at detecting the severity class of the cervical cancer in the pap smear image.

References:-

1. William Donald Kelley, Cancer: Curing the Incurable without Surgery, Chemotherapy, or Radiation. Paperback; 2001.
2. Kiichi Sugimoto. Prognostic factors after palliative resection for colorectal cancer with incurable synchronous liver metastasis. OJGas; 2013: 3, 259-8.
3. Tami Boehmer. From Incurable to Incredible: Cancer Survivors Who Beat the Odds. Paperback, Import; 2010.
4. Eli.G.Jones. Cancer its Cases, Symptoms and Treatment. B.Jain Publisher; 2003.
5. Carol E. Desantis MPH. Chun Chie Lin, Angela B.Marieto, Rebecca L.Siegal, Kevin.D.Stein, Joan.L.Kramer, Rich Alteri MD. Cancer Treatment and Survivorship Statistics, CA- A Cancer Journal of Clinicians; 2014:64, 252-20.
6. Eleanor Chadza, Ellen Chirwa, Alfred Maluwa, Address Malata, Abigail Kazembe, Angela Chimwaza. Factors that contribute to delay in seeking cervical cancer diagnosis and treatment among women in Malawi; 2012:4, 1015- 8.
7. S.Athinarayanan, M.V.Srinath. Pap Smear Screening Test and it's Applications in Cervical Cancer Detection and Classification. International Journal of Advanced Research in Computer Science (IJARCS). 2013; 4: 164-5.
8. S.Athinarayanan, M.V.Srinath. Cervical Cell for ThinPrep Image Enhancement using Morphological Segmentation. International Journal of Management, IT and Engineering (IJMIE).2012; 2: 285-11.
9. Erin J. Aiello Bowles, Robert Wellman, Heather Spencer Feigelson, Adedayo A. Onitilo, Andrew N. Freedman, Thomas Delate,Larry A. Allen, Larissa Nekhlyudov, Katrina A. B. Goddard, Robert L. Davis, Laurel A. Habel, Marianne Ulcickas Yood,Catherine Mccarty, David J. Magid, Edward H. Wagner. Risk of Heart Failure in Breast Cancer Patients After Anthracycline and Trastuzumab Treatment: A Retrospective Cohort Study. JNCI oxford journal. 2012; doi: 10.1093/jnci/djs317
10. N.Sakthi Priya. Cervical Cancer Screening and Classification Using Acoustic Shadowing. International Journal of Innovative Research in Computer and Communication Engineering. 2013; 1:1676-4.
11. Diane Solomon; Diane Davey; Robert Kurman; et al. The 2001 Bethesda System Terminology for Reporting Results of Cervical Cytology. Jama. 2009; doi:10.1001/jama.287.16.2114.
12. Philippa Holowaty."Natural History of the Dysplasia of the Uterine Cervix". JNCI(Journal of the National Cancer Institute). 1998 ; 91 : 252-7.
13. Rafael C. Gonzalez and Richard Eugene Woods. Digital Image Processing.2 ed. Prentice Hall; 2002.
14. Masahiro Ishikawa,, Sercan Taha Ahi, Fumikazu Kimura, Masahiro Yamaguchi, Hiroshi Nagahashi, Akinori Hashiguchi, Michiie Sakamoto. Segmentation of Sinusoids in Hematoxylin and Eosin Stained Liver Specimens Using an Orientation-Selective Filter. Open Journal of Medical Imaging: Scientific Research. 2013; 3: 144-11.

15. Alden A. Dima, John T. Elliott, James J. Filliben, Michael Halter, Adele Peskin, Javier Bernal, Marcin Kociolek, Mary C. Brady, Hai C. Tang and Anne L. Plant. Comparison of segmentation algorithms for fluorescence microscopy images of cells. Wiley Online Library.2011;79: 545-15.

16. Swe Zin Oo1, Aung Soe Khaing. Brain tumor detection and segmentation using watershed segmentation and morphological operation. IJRET: International Journal of Research in Engineering and Technology. 2014; 3: 367-8.

17. P.W.Kitchin and A.Pugh. Chapter, "Processing of Binary Images",International Trends in Manufacturing Technology. 1983;doi: 10.1007/978-3-662-09771-7_3. pp:21-42.

18. Wilma a. Winnick, rhea l. Dornbush. classification and identification thresholds for three types of verbal materials. Springer.1996; 1:181-4.

19. Ali Mohammad Nickfarjam,"Novel Supervised Bi-Level Thresholding Technique Based on Particle Swarm Optimization". Springer.2014; 39: 753-14.

20. S. A. Mohammad, R. Akbar and S. H. Mohammad. "An Efficient Method based on ABC for optimal multilevel thresholding", IJST, Transactions of Electrical Engineering. The Islamic Republic of Iran 2012; 36: 37-13.

21. Daniel X. Le. "Classification of binary document images into textual or nontextual data blocks using neural network models", Springer Verlag publication. 1995; 8: 289-16.

22. Norup, Jonas, Master thesis "Classification of pap-smear data by transductive neuro-fuzzy methods" 2005.

23. Soumya M. K, Sneha K, Arunvinodh C. "Cervical Cancer Detection and Classification Using Texture Analysis". Biomed Pharmacol J 2016;9(2).

24. K. Jian. "Fundamentals of Digital Image Processing". Prentice Hall, Englewood Cliffs, 1989.

25. S. Pradeep Kumar Kenny and S. Allwin, "Combination Feature Set for the Detection of Cervical Cyto Images", Sylwan Journal, Vol. 158, No. 6, pp. 423-429, 2014.

26. S.Athinarayanan, M.V.Srinath. "Computer Aided Diagnosis for Detection and Stage Identification of Cervical Cancer by using Pap Smear Screening Test Images"in ICTACT Journal on Image and Video Processing (IJIVP), ISSN:0976-9102 (Online), Vol:06, Issue:04, May 2016.

27. B. Julesz. Texton gradients: The texton theory revisited. Biological Cybernetics, 54(1986) 245-251.

28. S.Athinarayanan, M.V.Srinath."Robust and efficient diagnosis of cervical cancer in pap Smear images using textures features with RBF and Kernel SVM classification" in ARPN Journal of Engineering and Applied Sciences (ARPN), ISSN: 1819-6608, Vol:11, Issue:7, April 2016.

29. Chen, L.; Chen, C.L.P.; Lu, M.: A multiple-kernel fuzzy C-means algorithm for image segmentation. IEEE Trans. Syst.Man Cybern. B Cybern. 41(5), 1263–1274 (2011).