

# CERVICAL CANCER DETECTION AND CLASSIFICATION BY USING EFFECTUAL INTEGRATION OF DIRECTIONAL GABOR TEXTURE FEATURE EXTRACTION AND HYBRID KERNEL BASED SUPPORT VECTOR CLASSIFICATION

S. Athinarayanan<sup>1</sup>, K. Navaz<sup>2</sup>, R. Kavitha<sup>3</sup> and S. Sameena<sup>4</sup>

<sup>1,2</sup>Department of Computer Science and Engineering, Annamacharya Institute of Technology and Sciences, India

<sup>3</sup>Department of Computer Applications, The MDT Hindu College, India

<sup>4</sup>Department of Computer Science and Engineering, Universal College of Engineering and Technology, India

## Abstract

*Planning of invigorating representation is a troublesome and testing process because of the unpredictability of the images and absence of models of the life systems that thoroughly catches the reasonable expressions in each structure. Cervical malignant growth is one of the noteworthy reasons for death among different kinds of the diseases in women around the world. Genuine and auspicious determination can keep the life to some dimension. Therefore, we have proposed a computerized dependable framework for the analysis of the cervical malignancy utilizing surface highlights and machine learning calculation in Pap smear images, it is extremely advantageous to anticipate disease, likewise expands the dependability of the determination. Proposed framework is a multi-organize framework for cell nucleus extraction and disease finding. To begin with, clamor expulsion is performed in the preprocessing venture on the Pap smear images. Exterior highlights are separated from these demand free Pap smear images. Next period of the proposed framework is classification that depends on these separated highlights, SVM classification is utilized. Over 94% exactness is accomplished by the classification stage, demonstrated that the proposed calculation precision is great at recognizing the disease in the Pap smear images.*

## Keywords:

*Cervical Cancer, Feature Extraction, DGTF, Classification, Hybrid Kernel SVM*

## 1. INTRODUCTION

Cervical malignant growth has been ascending to be second most elevated factor causing passing among ladies all-inclusive and the event of cervical disease is overwhelming in creating nations like Malaysia, Philippines, Thailand and Indonesia. The quantity of female setback is high and common in the nation of Vietnam pursued by Malaysia and Philippines. Despite the fact that the passing rate had fallen after the inception of Human Papilloma Virus (HPV) immunization, different grounds of the ailment has not yet been wiped out. Pathologist assesses the slides utilizing the magnifying instrument to watch the spread of irregular cells. Malignancy is dealt with dependent on the examination. Determination got from pathologists gets shifted since the subjectivity has command over the slide understanding and furthermore investigation of malignant growth depends on the aptitude and ability of the pathologist. Issues will emerge when legitimate direction isn't accessible for unpracticed pathologists, so it is fundamental to propose a framework that could help pathologist amid demonstrative procedure and consequently subjectivity connected to finding be decreased. The Papanicolaou (Pap smear) test is viewed as most money saving advantage malignant growth counteractive action and location program

designed up until this point. It was begat by George N. Papanicolaou and in his discovering; he detailed that influenced cells of the cervix can be recognized by the little example of cells assembled from the cervix. They had an exploration and demonstrated the power of vaginal smear for the simple discovery of cervical carcinogenic and precancerous changes. The conventional Pap test is a straightforward technique comprising of steps referenced underneath.

- To see the cervix, vagina is broadened by embedding's the speculum.
- Sample cells are gathered in and around the cervix by brush or swab.
- Cells are protected in a glass by applying some additive.
- To feature the auxiliary examples for examination utilizing magnifying lens in the lab, tests are set apart to improve the differentiation in the example.

Another strategy for test cell readiness for cytological test is the Liquid Based Cytology and it is to some degree like pap smear test however the thing that matters is that as opposed to spreading the cells in the glass slide, the cell is all of a sudden wash with additive fluid arrangement then it will be sent to the lab for expelling non-demonstrative materials such like bodily fluid, discharge and platelets. After the protection, a slight layer of cell is shaped onto a glass slide, with this cell, it is inspected under the magnifying lens in a similar way of conventional smear test.

## 2. LITERATURE REVIEW

Cervical malignancy compromises the lives of numerous ladies in our present reality. In 2014, the quantity of ladies contaminated with this ailment in the United States was 12,578, of which 4,115 kicked the bucket, with a passing rate of almost 32%. Malignancy information, including cervical malignancy datasets, speak to a huge test information mining systems since nonattendance of various expenses for blunder cases [1]. Cervical harm is one of the indisputable reasons of illness going in females around the globe. The Pap smear is the mind blowing dynamic screening test used to see the cervical pre-hazardous and destructive assortments in an exploratory of cervical cells subject to the shape assortments of the centers and cytoplasm [2]. Pap test has to a great degree changed the figure of women with cervical harmful development as it has revealed its ability to perceive 95% of the tumors of the vaginal neck. Cervical harmful development can be turned away if it is seen and treated early. Built up a novel strategy to consolidate a few models utilizing a Bayesian methodology. The technique chooses the most pertinent

properties from a few models, and delivers a Bayesian classifier which has a higher arrangement rate than any of them, and in the meantime is extremely proficient. In light of contingent data measures, the strategy disposes of unimportant factors, and joins or dispenses with ward factors; until the point when an ideal Bayesian classifier is acquired. [3]. Pap smear test is a physical screening process used to distinguish cervical malignant growth or precancerous changes in an uterine cervix by evaluating cervical cells dependent on shading, shape and surface properties of their cores and cytoplasm. A PC helped screening structure for Pap smear tests will be correct useful to anticipate cervical disease in the event that it builds the unwavering quality of the diagnosis. “The machine learning method is a good substitute for manual diagnosis in the analysis of Pap smear cervical cell images, reflecting its effective and accurate classification” [4]. Papanicolaon (PAP) Smear Screening Test is the most widely recognized type of conclusion for identifying cervical malignant growth in its beginning times. Cervical malignancy is an infection that happens when cells in the cervix territory prompt to deliver out of controller and assault close-by tissues or devour all through the body. Malignant growth or tumor can be separated into two gatherings i.e. amiable and harmful. Generous, that does not assault and abrogate the tissue in which it begins or spread to reserved destinations in the body (non-destructive tumor) while threatening, that assaults and cancels the tissue in which it concocts and can devour to different locales in the body by means of circulation system and lymphatic framework (carcinogenic tumor) [5]. Medicinal image preparing framework has prompt an expanding imperative and developing job for image handling and PC supported determination (CAD) frameworks in various clinical applications. Cervical malignant growth is the second most basic disease influencing ladies worldwide and the main source of malignancy mortality in creating nations. It very well may be restored in all patients, whenever identified early and treated satisfactorily. A robotized image examination arrangement of uterine cervical images breaks down and extricates symptomatic highlights in cervical images and can help the doctor with a recommended clinical analysis. Such a framework could be incorporated with a therapeutic screening gadget to permit screening for cervical disease by non-restorative personnel. Cervical malignant growth majorly affects lady’s lives worldwide and one in each five ladies experiencing cervical disease has a place with India. Henceforth the goals of this examination were to discover the learning of ladies in regards to cervical malignant growth, to decide screening practices and determinants, and to recognize factors for non-screening [6]. An epic cell and atypical cell highlights determined from vertical allotments of the epithelium in the histology pictures. Epithelium characterization of pictures is inspected with casting a ballot combination of vertical portions by methods for help vector machine (SVM) and Linear Discriminant Analysis (LDA) strategies. Preparing and testing depends on forget one technique for acquiring precise evaluation naming exactness. First request auxiliary measures and second request insights estimates dependent on GLCM were utilized for removing textural highlights. Test cell got from magnifying instrument is separated into three sections: cytoplasm regions, core regions and a cell territory. Characterization results are assessed utilizing three scoring plans and its disparities are analyzed by methods for classifiers [7]. To employee Artificial Neural system for

arrangement of cervical malignant growth cells in the cervix district of the uterus. Here, multi-layer (MLP) has been joined to find the cervical disease cells which have two phases: pre-processing and feed forward MPL neural systems. The unsegmented cell highlights are arranged into typical, LSIL, HSIL cells by Levenberg-Marquarat feed forward MLP neural system. The fell multi-layered preceptor (C-MLP) and extraordinary learning machine (ELM) and its execution likewise were broke down. Fluffy C implies calculation was likewise used in the neural system to make centre layer [8].

## 2.1 MAJOR CONTRIBUTION OF THE PROPOSED SYSTEM

In our proposed system, a cervical disease order framework mainly depends upon the feature extraction and classification of segmented cervical cell nucleus. In highlight extraction, numerous highlights are utilized to decide the significance of typical and unusual images. In this situation, it is fundamental to limit every one of the highlights separates that are resolved between the malignant growth image and the non-dangerous images. To play out the order arrange proficiently, a viable grouping calculation is required. In this exploration work, we abuse the proposed half and half calculation in the grouping stage to guarantee the arrangement execution. Our proposed technique comprises of three stages to be specific, cervical cell nucleus detection, highlight extraction and arrangement. In this paper, cervical cell nucleus detection is finished by utilizing pre-preparing and division process. In pre-preparing, anisotropic channel is connected to expel the commotion and upgrade the image for the phone nucleus discovery process. Four normal learning calculations, for example, Naïve Bayes, J48 Decision Tree, Sequential Minimal Optimization, and Random Forest were looked at as far as their exactness, hamming misfortune, correct match (subset precision) and positioning misfortune execution assessment measurements.

## 2.2 CERVICAL CANCER NUCLEUS DETECTION PROCESS

The original thinprep cervical cell shading Image, regardless of whether it is a typical, splendid or dim image, first it is changed over into their separate Red, Green and Blue plane Images. At that point every  $R$ ,  $G$  and  $B$  plane of a unique image was first locally edge by utilizing given conditions in Eq.(1)-Eq.(3). That is the mean estimation of the first info image was increased with the components  $\alpha$ ,  $\beta$  and  $\gamma$ , by altering the estimation of this elements we get fragmented division of the core. So Cytoplasm locales were expelled by playing out a Morphological shutting task utilizing a structure component littler than the littlest core and atomic in homogeneity was rectified by a Morphological opening of comparative size. The subsequent divided image for the particular R, G and B plane were utilized for recognizing the core with a larger number of subtle elements than the first image and it is given in Eq.(4)-Eq.(6)

$$MI_R = Mean(I_R) * \alpha \quad (1)$$

$$MI_G = Mean(I_G) * \beta \quad (2)$$

$$MI_B = Mean(I_B) * \gamma \quad (3)$$

$$SI_R = ((TI_R \bullet MC_R) \circ MO_R), \quad (4)$$

$$SI_G = ((TI_G \bullet MC_G) \circ MO_G), \quad (5)$$

$$SI_B = ((TI_B \bullet MC_B) \circ MO_B), \tag{6}$$

where  $SI_R, SI_G$  and  $SI_B$  are the subsequent portioned image for the particular  $R, G$  and  $B$  plane.  $TI_R, TI_G$  and  $TI_B$  is the edge image for the separate  $R, G$  and  $B$  plane,  $MC_R, MC_G$  and  $MC_B$   $MO_R, MO_G$  and  $MO_B$  are organizing components for the particular  $R, G$  and  $B$  plane. The images  $\bullet$  and  $\circ$  mean morphological shutting and opening, individually. At that point by consolidating  $R, G$  and  $B$  channel segmented image  $SI_R, SI_G, SI_B$ , we get the portioned cervical cell image as in condition Eq.(7).

$$SI_{RGB} = ((TI_{RGB} \bullet MC_{RGB}) \circ MO_{RGB}) \tag{7}$$

where,  $SI_{RGB}$  is the resulting segmented color image,  $TI_{RGB}$  is the threshold image;  $MC_{RGB}$  and  $MO_{RGB}$  are structure elements of both the closing and opening operation [9].

### 2.3 DIRECTIONAL GABOR TEXTURE FEATURES (DGTF) FEATURE EXTRACTION:

Directional Gabor's are utilized as they measure the heterogeneity in the SROI. Gabor channel is a Gaussian part work tweaked by a sinusoidal plane wave. There-fore, it gives directional surface highlights at a predefined Gaussian scale. Gabor part is characterized as given in Eq.(8).

$$g(x,y,\lambda,\theta,\psi,\sigma,\gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) x \cos\left(2\pi \frac{x'}{\lambda} + \Psi\right) \tag{8}$$

where  $x' = x\cos\theta + y\sin\theta, y' = -x\sin\theta + y\cos\theta'$

In this condition,  $\lambda$  speaks to the wavelength of the sinusoidal factor,  $\theta$  speaks to the introduction of the typical to the parallel stripes of a Gabor work,  $\Psi$  is the stage counterbalanced,  $\sigma$  is the width of the Gaussian, and  $\gamma$  is the spatial perspective proportion, and indicates the elasticity of the help of the Gabor function [10]. DGTF based Extracting no of highlights are appeared in the accompanying Table.1.

Table.1. DGTF features

DGTF	DGTFs are calculated at $\lambda$ for $2\sqrt{2}, 4, 4\sqrt{2}, 8, 8\sqrt{2}$ and $\theta$ for $0^\circ, 22.5^\circ, 45^\circ, 67.5^\circ, \text{ and } 90^\circ$ are varied. Four statistical parameters are calculated for each filter output in the marked SROI and are taken as 100 features in the feature bank. These parameters are: (1) mean intensity, (2) standard deviation, (3) Skewness, (4) Kurtosis	25*4=100 features
------	--	-------------------

### 2.4 HYBRID KERNEL-SVM CLASSIFIER

The symptomatic models, half and half portion based SVM has been produced for enhancing the characterization procedure. The extricated surfaces highlights are utilized for the division of two classes, for example, disease and non-carcinogenic. Since the surface component pursues the non-linearity, non-straight SVM is expected to do the partition of hyper plane. To do non-direct undertaking, bit capacities are presented in SVM arrangement. Various parts are consolidated to build up another cross breed portion that will enhance the characterization errand of isolating

the preparation information. By presenting the crossover bit, SVMs gain adaptability in the decision of the type of the limit, which require not be direct and even not to have the equivalent useful frame for all information, since its capacity is non-parametric and works locally.

- *Radial basis function Kernel:* The support vector will be the centre of the RBF and  $\sigma$  will determine the area of influence. This support vector has the data space in Eq.(9).

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \tag{9}$$

- *Quadratic Kernel Function:* Polynomial kernels are of the form  $K(\vec{x}, \vec{z}) = (1 + \vec{x}^T \vec{z})^d$ , where  $d = 1$ , a linear kernel and  $d = 2$ , a quadratic kernel are commonly used.

Let  $k_1(RBF)$  and  $k_2(Quadratic)$  be kernels over  $\Xi \times \Xi, \Xi \subseteq R^p$  and  $k_3$  be a kernel over  $R^p \times R^p$ . Let function  $\varphi: \Xi \rightarrow R^p$ . The four kernel based formulations are represented by in Eq.(10) and Eq.(11).

$$k(x, y) = k_1(x, y) + k_2(x, y) \text{ is a kernel} \tag{10}$$

$$k(x, y) = k_1(x, y)k_2(x, y) \text{ is a kernel} \tag{11}$$

- *Hybrid Kernel:* By substituting Eq.(10 and Eq.(11) in Lagrange multiplier Eq.(9), we get the proposed hybrid kernel.

## 3. EXPERIMENTAL RESULTS AND COMPARATIVE ANALYSIS

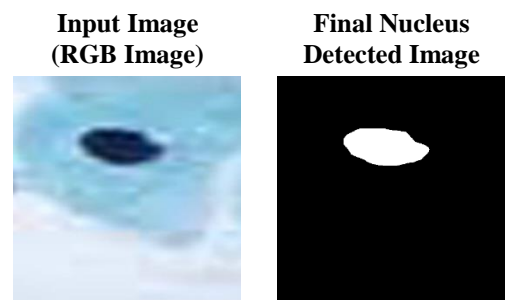
### 3.1 EXPERIMENTAL IMAGE DATASET

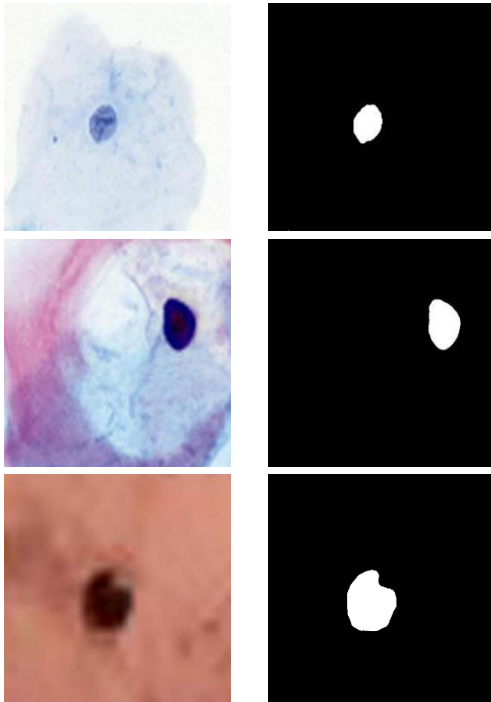
In this study 100 images (80 abnormal and 20 normal) have been used which were derived from Pap smear tests collected from Private Hospital situated Tirunelveli district of Tamil Nadu State in India. The images were taken with 100x lens magnification using Olympus ch20i Microscope. Each image was examined and diagnosed by pathologists of that hospital before being used as reference for this study.

### 3.2 EXPERIMENTAL RESULTS

In this experimental result section as per the three process collected images could be processed. The Result was given in following Table.2.

Table.2. Experimental Results





### 3.3 FEATURE EXTRACTED RESULT

After the segmented nucleus (ROI), we have to extract the statistical four features of the nucleus mean intensity, standard deviation, Skewness and Kurtosis as five different wavelengths ( $\lambda$  for  $2\sqrt{2}$ , 4,  $4\sqrt{2}$ , 8,  $8\sqrt{2}$ ) and five different directions ( $0^\circ$ ,  $22.5^\circ$ ,  $45^\circ$ ,  $67.5^\circ$ , and  $90^\circ$ ) and its around 100 features and based on this 100 features will be applied to the hybrid kernel SVM classification will give more efficient than the existing system.

### 3.4 PERFORMANCE EVALUATION OF PROPOSED SYSTEM

Classifier execution assessment of this work is directed with generally utilized factual measures, affectability, particularity, precision and rate [12] [13]. True Positive (TP) is characterized as the quantity of effectively distinguished positive pixels; True Negative (TN) is characterized as accurately recognized negative pixels. For instance, in an analytic test, assessment concentrating on the nearness of unusual tissues, tumor tests is considered in the positive class and typical tissues will be in the negative classification. False Positive (FP) speaks to the check of typical tissues inaccurately distinguished as a tumor, and False Negative (FN) gives the tally of anomalous examples mistakenly recognized as would be expected tissue [11]. Higher estimations of affectability, the extent of effectively arranged positives, demonstrate better execution of the strategy in anticipating positives. Explicitness estimates how well the framework can anticipate the negatives. Exactness estimates the general rightness of the classifier in foreseeing the two positives and negatives, and by and large blunder rate is ascertained according to the accompanying the equation Eq.(12)-Eq.(15).

$$Sensitivity = \frac{TP}{TP + FN} \quad (12)$$

$$Specificity = \frac{TN}{TN + FP} \quad (13)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (14)$$

$$Error\ rate = 1 - Accuracy \quad (15)$$

### 3.5 COMPARATIVE ANALYSIS

We have compared our proposed cervical cancer classification system (HKSVM), against the neural network techniques. The neural networks, we have utilized for comparative analysis are Radial Basis Function (RBF) neural network. The performance analysis has been made by plotting the graphs of evaluation metrics such as sensitivity, specificity and the accuracy are shown in Table.3.

Table.3. Experimental Results of Existing and Proposed Method

Evaluation metrics		Texture Features with HKSVM (Proposed)	Texture Features with RBF
Input Cervical Cell image data set	TP	38	37
	TN	9	8
	FP	1	2
	FN	2	3
	Sensitivity	95	92.5
	Specificity	90	80
	Accuracy	94	90
Total error (%)		6	10

By analyzing the results, the performance of the proposed technique has significantly improved the cancer detection compared with Radial Basis Function (RBF) neural network classifier. Based on the experimental results our proposed method produces better results compared to other neural network based classifiers.

### 4. CONCLUSIONS

In this paper, we have built up a mechanized cervical malignant growth indicative framework with ordinary and unusual classes. The restorative basic leadership framework was structured with the Texture highlights and piece based Support Vector Machine. The proposed methodology includes highlight extraction and arrangement. The advantage of the framework is to help the doctor to settle on an ultimate conclusion decisively. As indicated by the exploratory outcomes, the proposed strategy is proficient for the order of image into typical and unusual. For similar investigation, our proposed methodology is contrasted and other neural systems RBF. The precision level (94%) for our proposed strategy demonstrated that the proposed calculation chart is great at recognizing the malignant growth in the exploratory images.

### REFERENCES

- [1] Hayder K. Fatlawi, "Enhanced Classification Model for Cervical Cancer Dataset based on Cost Sensitive Classifier", *International Journal of Computer Techniques*, Vol. 4, No. 4, pp. 1-8, 2017.

- [2] S. Athinarayanan, M.V. Srinath and R. Kavitha, "Computer Aided Diagnosis for Detection and Stage Identification of Cervical Cancer by using Pap Smear Screening Test Images", *ICTACT Journal on Image and Video Processing*, Vol. 6, No. 4, pp. 1244-1251, 2016.
- [3] M. Martinez, L.E. Sucar, H.G. Acosta and N. Cruz, "Bayesian Model Combination and Its Application to Cervical Cancer Detection", *Proceedings of Ibero-American Conference on Artificial Intelligence*, pp. 622-631, 2006.
- [4] G. Sun, S. Li, Y. Cao and F. Lang, "Cervical Cancer Diagnosis based on Random Forest", *International Journal of Performability Engineering*, Vol. 13, No. 4, pp. 446-457, 2017.
- [5] S. Athinarayanan and M.V. Srinath, "Classification of Cervical Cancer Cells in PAP SMEA Screening Test", *ICTACT Journal on Image and Video Processing*, Vol. 6, No. 4, pp. 1234-1238, 2016.
- [6] S. Aswathy, M.A. Quereshi, B. Kurian and K. Leelamoni, "Cervical Cancer Screening: Current Knowledge and Practice among Women in a Rural Population of Kerala, India", *Indian Journal of Medical Research*, Vol. 136, No. 2, pp. 205-210, 2012.
- [7] M. Anousouya Devi, S. Ravi and J.V.S. Punitha, "Detection of Cervical Cancer using the Image Classification Algorithms", *International Journal of Circuit Theory and Applications*, Vol. 9, No. 3, pp. 153-166, 2016.
- [8] G. Peng, B. Koyel and R.S. Joe, "Nuclei-Based Features for Uterine Cervical Cancer Histology Image Analysis with Fusion-Based Classification", *IEEE Journal of Biomedical and Health Informatics*, Vol. 20, No. 6, pp. 1595-1607, 2016.
- [9] K. Thangavel, P.P. Jaganathan and P.O. Easmi, "Data Mining Approach to Cervical Cancer Patients Analysis using Clustering Technique", *Asian Journal of Information Technology*, Vol. 5, No. 4, pp. 413-417, 2006.
- [10] S. Athinarayanan and M.V. Srinath, "Robust and Efficient Diagnosis of Cervical Cancer in Pap Smear Images using Texture Features with RBF and Kernel SVM Classification", *ARPJ Journal of Engineering and Applied Sciences*, Vol. 11, No. 7, pp. 4504-4515, 2016.
- [11] S. Athinarayanan and M.V. Srinath, "Severity Analysis of Cervical Cancer in PAP SMEAR Images by using EEETCM, ERSTCM and CFE Method based Texture Features and Hybrid Kernel Based Support Vector Machine Classifier", *International Journal of Advanced Research*, Vol. 4, No. 11, pp. 2751-2764, 2016.
- [12] S. Athinarayanan and M.V. Srinath, "Effective Detection of Cervical Cancer During Pap Smear Screening Test", *Elysium Journal of Engineering Research and Management*, Vol. 1, No. 1, pp. 10-14, 2014.
- [13] S. Athinarayanan and M.V. Srinath, "Multi Class Cervical Cancer Classification by using ERSTCM, EMSD and CFE Methods Based Texture Features and Fuzzy Logic Based Hybrid Kernel Support Vector Machine Classifier", *IOSR Journal of Computer Engineering*, Vol. 19, No. 1, pp. 23-34, 2017.