Journal of Advanced Computer Applications *ISSN No.: XXXX-XXXX*
vol-01, issue-01, December, 2024 *Pages 95-100*

JOURNAL OF ADVANCED COMPUTER APPLICATIONS

ISSN: XXXX-XXXX (Online)
Vol: 01 **Issue:** 01
Contents available at: https://www.swamivivekanandauniversity.ac.in/jaca/

Optical Imaging for Cancer Detection through Machine LearningApproaches

Shreya Adhikary[1] *

[1] *Department of Electronics and Communication Engineering, Swami Vivekananda University, Barrackpore-700121, WB, INDIA, Email ID: shreyaa@svu.ac.in**

ABSTRACT

Aspects of microscope imaging have already been investigated for the auto-matic identification of cervical cancer, while brightfield techniques have emerged for the detection of cervical cancer in fresh tissue. Using epithelial cells taken from the endo-cervical region, it might be a successful strategy. This study avoids the issue of nucleus segmentation and instead presents a technique for the automatic classification of cervical cells. This technique used a collection of spectral texture features (SPTF) that have been taken from single cell images that are transformed into two dimensions (2-D). Two 1-D functions, namely the radial function and the angular function accumulated from the frequency spectrum, are used to evaluate SPTFs in order to distinguish between normal and pathological cells. Support vector machines (SVMs) are used to evaluate the classification. The Gaussian SVM reaches to the greatest accuracy is 0.94, with sensitivity 0.89 and specificity 0.83. A promising performance more than 0.75 is obtained to achieve the total accuracy.

Keyword: Cervical cancer, Human papillomavirus, Spectral texture features, Linear Support Vector Machine, Non-linear Support Vector Machine.

I. INTRODUCTION

In 2023, cervical cancer ranked fourth globally among malignant tumors in women, accounting for 6.6% of all cancer diagnoses in females [1]. The cancer is developed by a sequence of alterations from precancer. However, it is generally found that, in 90% case, the early detection and treatment of precancerous lesions may prevent the individual cervical cancer . Human papillomavirus is responsible for this type of cancer.

Most of the proposed techniques employed the methods that are subject to do segmentation of the cytoplasm and nucleus [2-3]. To circumvent the problem of cell and nuclei segmentation, several attempts have focused on the categorization of the Papanicolaou (Pap) test using sets of properties derived from the frequency domain. The virtue of this

---

* Authors for Correspondence

approach is that rigorous segmentation of the cell and its nucleus is not elementary. However, it should be noted that this approach was only applied to the single-cell image captured at high resolution.

This work recommends a new procedure to classify traditional cervical cells and abnormal cells (precancerous cells) by utilizing a set of basic properties extracted from Fourier transform of low- resolution cell images. This allows us to bypass the cell segmentation process and a high-resolution image purchasing system. This section represents the method of this research work. The performance of this work is evaluated by using a publicly available dataset of 800 cervical Pap smear cell images, collected from HErlev Pap Smear Dataset [4]. The dataset are visualized in Fig. 1.

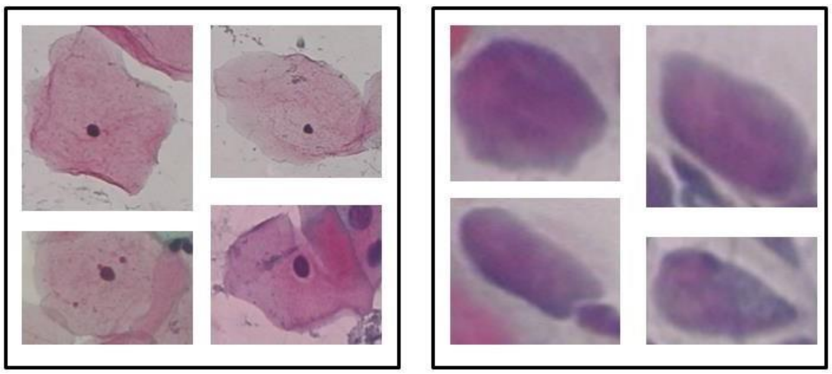

Fig. 1 HErlev dataset: normal (left column) and cancer (right column)

Firstly, the DFT transforms M×N gray scale image into 2-D Fourier transform, transforming components into polar coordinates and separated into radial and an-gular functions. All elements of the Fourier spectrum are transformed into polar coordinates, F(r, θ), which are divided into two 1-D functionalities, that is, the ra-dial functionality, F(r) and the angular functionality, F(θ) [5] .

II. METHODOLOGICAL ASPECT

A) Radial functionality and Angular functionality

The spectrum consists of frequency components along a circle of radius r and a radial line with angle θ, represented by mean, variance, and entropy through statistical evaluations. The features of normal and abnormal cells have been calculated using annular ring subsections and radian wedge subsections.

The image information is divided into a frequency component with a small annular ring, divided into 8 equidistance annular rings with mean, variance, and entropy. The cervical dataset is utilized by two labels, i.e., normal and abnormal. Normal data represents the healthy volunteer and abnormal data specifies diseased person with cancer. This work presents machine learning approaches to predict and classify cervical cells. It includes stages of data collection, model training, testing, and comparing outcomes. The dataset is divided into two parts as training and testing part with 80% and 20% respectively. Vapnic invented the Sup-port Vector Machine (SVM) to solve data classification and regression problems by dividing original data into different groups and creating a hyperplane to predict new data labeling [6]. The SVM algorithm aims to construct the optimum decision boundary or line that can divide n-dimensional space into classes so that we can quickly classify the data points. The optimal decision boundary, called hyperplane, which is separated the two types of data. The hyperplane is created by SVM by selecting the extreme points and vectors. The SVM classifier is based on these extreme situations, which are referred to as support vectors. The SVM algorithm is often reported to achieve better results than other classifiers, although it has been indicated that the main reason to use an SVM instead is because the problem might not be linearly separable. In that case, an SVM with a non-linear kernel such as the Radial Basis Function (RBF) would be suitable [7].

B) Performance Measure

The classification is analyzed by some statistical methods are described below:

Sensitivity: The fraction of diseased cases predicted as diseased

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

Specificity: The fraction of normal cases predicted as normal one

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

Accuracy: The fraction of cases the model correctly identified

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Where, True Positive (TP): Predicted diseased, actual diseased, False Positive (FP): Predicted diseased, actual non-diseased False Negative (FN): Predicted non-diseased, actual diseased True Negative (TN): Predicted non-diseased, actual non-diseased.

III. RESULT AND DISCUSSION

This section presents two SVM-based approaches for classification of cervical cancer data, evaluated using 10-fold cross-validation. Two target variables are diagnosed, and the total accuracy of these methods is plotted.

The non-linear SVM can classify the abnormal dataset and achieve high classification accuracy of 0.94. The linear SVM and non-linear SVM reduce computation costs and can perform diagnosis with less factors or components. Though, the non-linear SVM method effectively detects faulty samples, with a total accuracy of 0.94 for cervical cancer dataset. Classification results of HErlev Pap test dataset using linear and non-linear SVM techniques are graphically shown in fig. 2 and also depicted in tabular form in Table 1.

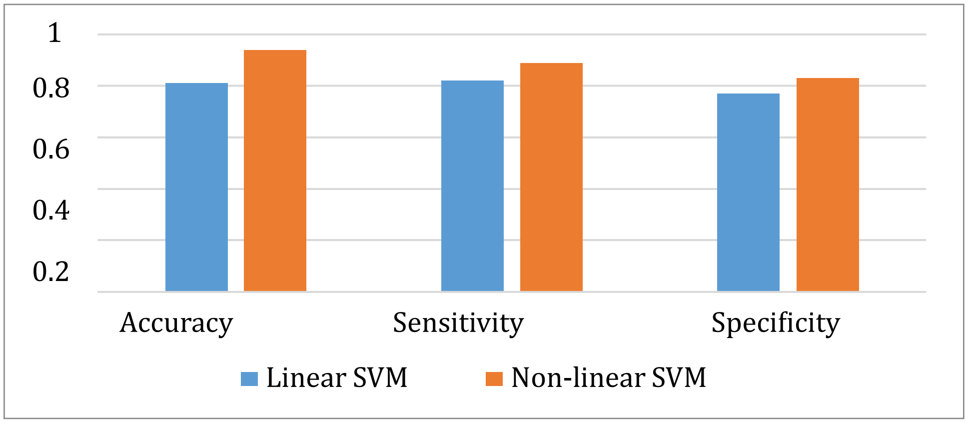

Fig. 2 The classification results of the present study with respect to accuracy, sensitivity and specificity.

Table 1: Classification performance of HErlev Pap test dataset using linear and non-linear dataset.

Classifiers	**Accuracy**	**Sensitivity**	**Specificity**
Linear SVM	0.81	0.82	0.77
Non-linear SVM	0.94	0.89	0.83

IV. CONCLUSION

This paper shows the cervical cancer risk factors and compares three SVM-based approaches for classification of cervical cancer datasets. Linear SVM and Non-linear SVM effectively classify malignant and benign cancers. With linear SVM shows improved classification results. However, linear SVM has better capability with the same number of features. This research study investigates the classification of normal and abnormal cervical cells using linear and non-linear SVM classifiers with simplified set of SPTFs. Linear SVM has low classification performance, while non-linear SVM achieves highest accuracy and specificity.

REFERENCE

1. https://www.who.int/health-topics/cervical-cancer#tab=tab_1
2. J. V. Lorenzo-Ginori, W. Curbelo-Jardines, J. D. L´opez-Cabrera, and S.B. Huergo-Su´arez, Cervical cell classification using features related to morphometry and texture of nuclei, Lecture Notes in Computer Science, vol. 8259 LNCS, no. PART 2, pp. 222–229 (2013)
3. M. E. Plissiti, C. Nikou, and A. Charchanti, Combining shape, texture and intensity features for cell nuclei extraction in Pap smear images, Pattern Recognition Letters, vol. 32, no. 6, pp. 838–853 (2011)

4. https://paperswithcode.com/dataset/herlev
5. T. Chankong, Automatic Classifying of Cervical Cells Using Fourier Spectral Features, Proc. 2018 4th Int. Conf. Green Technol. Sustain. Dev. GTSD,pp. 759–762 (2018)
6. V. Cortes, Corinna and Vapnik, Photonit neural networks and learning machines the role of electron-trappingmaterials, IEEE Expert, vol. 7, no. 5, pp. 63–72(1992) https://pubmed.ncbi.nlm.nih.gov/9691570/