International Journal of u- and e- Service, Science and Technology
Vol.9, No. 11 (2016), pp.195-216
http://dx.doi.org/10.14257/ijunesst.2016.9.11.18

Density based Multiclass Support Vector Machine using IoT driven Service Oriented Architecture for Predicting Cervical Cancer

Sakthi A[1*] and Rajaram M[2]

[1]*Department of Electronics and Communication Engineering,*
Bharathidasan Engineering College,
Natrampalli, Vellore District, Tamil Nadu, 635852–INDIA.
[2]*Department of Electrical Engineering Anna University, Chennai*
Tamil Nadu, 600025–INDIA.
[1]*sakthijesma@gmail.com*,[2]*rajaramgct@rediffmail.com*

Abstract

Cervical Cancer stands out among those deadliest diseases, which threatens women in an alarming rate causing approximately 2, 66,000 mortalities per annum worldwide. This cancer can be diagnosed early enough through Pap smear test; a cervical cancer screening program. Finding out the true positive rates of the Cervical Cancer cells with precision is more complex when identifying the same categories of the cancer disease. Various researchers have proposed many approaches over the past four decades and the solutions are pertinent to cervical cancer; however, the challenge remains partially unresolved. The significant contribution of this paper is in two folds, firstly discuss a cloud ready Adapter Driven Service Oriented Architecture (RESPRO 3.0), developed by us for automated screening of Pap Smear can be extended to any International Classification of Diseases (ICD). Secondly, present an Internet of Things (IoT) driven Cervical Cancer prediction adapter built for RESPRO 3.0 based on Density based Multiclass Support Vector Machines (MCSVM) in combination with Polynomial Kernel Trick. The density parameters provide unique space in identifying cervical cancer cell categories compared to exising researches. This cloud solution's results are bench marked and verified against cyto technician's ground truth results, found to be highly satisfactory with respect to 93% Sensitivity and 99% Specificity while minimizing test repeatability ratio for the supervised training set of images.

Keywords*: Cervical Cancer, Density based Support Vector Machine (DSVM), Service Oriented Architecture (SOA), Kernel Functions, Polynomial Kernel Trick*

1. Introduction

Ever since the cell theory was understood as cell being the fundamental building block of human body, biologists have been keen in investigating the cell for human betterment in finding reasons behind disease patterns [1]. Cervical cancer is prevalent among women due to many reasons but in scientist's opinion the reasons are such as smoking, early marital sex, and weak immune systems. however concrete reasons are still debated. An early and effective treatment is the most important means to prevent pre-cancerous cells to develop in the human body. Using the pap-smear techniques developed by Georgios Papanicolaou, a specimen of cell is smeared onto a glass slide and colored, making it easier to examine the cells under a microscope for any abnormalities indicating a pre-cancerous stage [2], cyto-technicians thus detect pre-cancerous cells in the uterine cervix using microscopes; these cells that turn out to be Cervical Cancer in the near future when

*Corresponsing Author

ISSN: 2005-4246 IJUNESST

untreated. The term Pap-smear refers to samples of human cells stained by Papanicolaou method.

The purpose of the Papanicolaou method is to diagnose pre-cancerous cell changes before they progress to invasive carcinoma (Meisels & Morin, 1997) [3] as advancements in preventive treatment approaches does create a great impact in the disease life cycle. For the cervical cancer the preventive methods and treatments are held nationwide in the developing countries. Developed countries have special programmes that associates women after 30 years old get tested with the Pap smear screening programme once in definitive period of years.Many diagnostic centers, healthcare service providers are widely spread across the globe but the quality and accuracy of diagnostic results are prone to be erroneous due to different factors.The act of false results sometime deviate the treatment procedures to a worse situation. [4] Manual Microscopic diagnosis is one such area where due to human errors the actual results are either not found or missed out. As stated in the introduction this research work attempts to address the solution in two folds, firstly the solution approaches to processing pap smear images using the well architected clinical information processing systems devloped (RESPRO 3.0) by us with enriched Service Oriented Architecture platform that can take advantage of IoT scenarios, Secondly with Density based MCSVM to provide better results as compared to existing methods and techniques.This can be handled by heterogeneous systems to submit clinical data to the RESPRO 3.0 System that processes from single point with reduced cost and with minimized human errors [5].

This paper is organized as, Section 2 explores the related surveys and motivation factors for cervical cancer analysis. Section 3 details the Pap Smear Image processing methods and Bethesda system for interpreting pap smear results. Section 4 outlines the need for heterogeneous service oriented architecture that integrates with IoT and the platform developed to cater to healthcare needs. Section 5 details the Multi class Support Vector Machine, Proposed Density based MCSVM and the kernel tricks applied in our solution. Section 6 provides details of cervical cancer prediction automation using RESPRO 3.0 Adapter with tools and performance assessment, finally Section 7 concludes with merits of this research work, extensibility of the research, and future focus.

2. Related Work

Automation of cervical cancer screening gained a special attention over the the past 50 years [Meyer 1986 [6], Pycock & Taylor 1980 [7], Taneka *et. al*, 1977 [8], Tucker et al 1987 [9], Zahniser *et. al.*, 1979 [10]. Automated image segmentation for cell analysis is generally a difficult problem due to the large variability such as various microscopes, various stain patterns, re-agents colour formation, time lapse and so on neverthless vast majority of cell segmentation techniques have been discussed since 1960s on cell segmentation.

Extensive pre-analysis over pap smear images have been performed in the past. An expert system CYTOPATH is designed through meta model and based on the priori information squamous lesions are decided and severity is concluded [11], Squamous Intraepithelial Lesions (SIL) is identified to screen the pap smear true positives. [12] Preliminary results for the classification of Pap Smear cell nuclei with textural features, using Gray Level Co-occurrence Matrix (GLCM) is performed [13]. A composite classification scheme implemented by combining several classifiers with distinctly different design methodologies. The classifiers are selected from a number of state of the art pattern classification schemes with a view to obtain superior performance. In this scheme, no priori information except a set of pre-classified data is assumed to be available and pap smear images being tested for biological cell classification [14]. The use of connectionist methods such as multilayered perceptrons, radial basis function (RBF) networks and ensembles of such networks were investigated. RBF ensemble algorithms based on fluorescence spectra potentially provide automated and near real-time

implementation of pre-cancer detection in the hands of non-experts [15]. In the pre-processing stage, first a set of ten features are extracted from a Pap smear image and is used to form the feature space, then the standard 'The Bethesda System" (TBS) rules are translated into fuzzy rules, which are used to classify the Pap smear test into 'normal" or 'abnormal classes based on the extracted features [16]. Using two feature screening measures, the initial feature set is effectively reduced to a computationally manageable size. Based on pixel-level screening results, cancerous regions can thus be detected through a relatively simple procedure [17]. Fast Fuzzy-based Automatic Pap Screening Systems algorithm, that can be used for future Automated Cell screening systems. The algorithm detects areas of the smear where the cells are located. Identification of the best areas for screening provides the important degree for evaluation of each fields [18].

A method, which outlines the area of metaplastic changes, known as the 'Transformation Zone' (TZ). Via Gabor Filter analysis a new frequency domain design scheme for Gabor wavelet with varying tuning frequencies and orientations are proposed. These are specifically designed to reduce the redundancy in the wavelet-based representation of the cervical image. The scheme calculates the Gabor wavelets either at dynamic or at continuous scales and identifies the region of interests [19]. The application of the machine learning methodology of Support Vector Machines (SVM) using FTIR data to enhance and improve upon the standard Pap test results [20] which happens in the stage where malign cells are visible, our work in this paper uniquely differs from this paper [20] by contributing at the initial stage of pre cancerous cell identification that helps to avoid repeated tests. While Pap test undoubtedly facilitates diagnosis, it suffers from a number of weaknesses such as blurriness as well as the effects of unwanted noises, which leads to false diagnosis. In order to address some of these problems, this study [21] discusses a novel image processing approach involving segmentation and contrast enhancement techniques. Firstly, images of cervical cells on the Pap smears are selected using the segmentation technique. Then, four algorithms of contrast enhancement technique are applied on the selected cell to increase their contrast. A contrast enhancement technique, which is only applied on the cervical cell of interest. The proposed technique is divided into two stages. Firstly, the cervical cells of interest will be selected using the modified seed based region growing algorithm. Then, the contrast of cervical cell of interest in the ThinPrep® will be enhanced by using three contrast enhancement algorithms. The cervical cell of interest will be applied with linear contrast algorithm and the proposed nonlinear algorithms namely non-linear bright and nonlinear dark contrast to enhance the contrast of the ThinPrep® images [22].

Extracting those cell nuclei which are abnormally large size, bizarre shape as well as hyper density to the different kinds of abnormal cells. The global information of the image and the local image condition were meantime considered. This method searched whole picture by scanning on different axes and determined the locations of abnormal cell nuclei with high contrast and find squamous intraepithelial neoplasia [23]. Edge Detection process and Pseudo-Color technique, with Color Space Extraction employed at preprocessing stage. First, the color space concept is applied to extract the original image of Pap smear into red plane, green plane and blue plane. Then the Seed Based Region Growing technique is applied to find boundaries of the cells. Pseudo-color technique is then embedded to the demarcated region to determine each part of the cell; nucleus, cytoplasm and background [24]. A fully automated method for the accurate detection of cell nuclei boundaries in conventional Pap smear images, based on the watershed transform for the extraction of nuclei and cytoplasm markers, which are used as starting points for the flooding process. A morphological reconstruction step is initially performed in each image. The watershed transform is then applied in the color morphological gradient image, which shows the boundaries of the more pronounced nuclei [25].These efforts in the cervical cancer pregnosis realm gave a real inspiration to think of a centralized processing platform that can be used for diverse needs of medical image

anlaysis. Hence a Service Oriented Architecture based system being developed indigenously with the future forecast in mind. As devices proliferating in day today's needs IoT scenarios going to play a key role in providing input and this can be harnessed in a great level especially in healthcare; The benefits and implementation strategy is harnessed from here [26-27] however we have uniquely scaled up this model with adpater driven solution pattern to identify possible auotmation for health challenges classified by ICD using our own implementation of RESPRO 3.0 solution.

3. Pap Smear System

In Papanicolaou smear method, to find pre-cancerous cells in the uterine cervix, a small cytological specimen from the uterine cervix is collected with a special cyto-brush and smeared onto a glass slide. Then the slide is stained using the Papanicolaou method. The different components of the cells are emphasized with specific colors. This glass slide is then viewed under a microscope, so cyto-technicians can diagnose the cells on the glass slide to identify the intensity of the pre-cancerous cells shown in Figure 1.

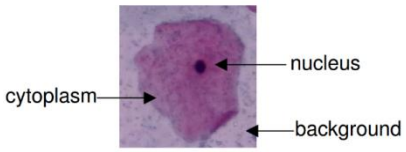


Figure 1. Single Pap Smear Image

Cyto-technicians use several different features to get a cell diagnosis. The size, color, shape and the texture of the nucleus and cytoplasm is used. The density of cells in a certain area, can influence the diagnosis. It takes a skilled cyto-technician to be able to differentiate between the different cells. [28] Every glass slide, can contain up to 3,00,000 cells. Therefore it is a time consuming job viewing the slides manually, this provides an opportunity for researchers to automate this cumbersome process.

3.1. The Papanicolaou Cells

Ideally specimens are taken from several areas of the cervix as shown at Figure 2. Depending on the area, the cyto-brush, cotton stick or the wooden stick is used. [2] The specimens most often contain cells from the columnar epithelium and the squamous epithelium. The columnar epithelium is located in the upper part of the cervix, and the squamous epithelium in the lower part. In between these two areas, the metaplastic epithelium is found, also called the transformation zone. [29-38] A single Pap smear slide may contain hundreds of thousands of cells and cytotechnician examine these cells under the microscope to determine premalignant cell changes based on the cell characteristics like size, color, shape and texture of nucleus and cytoplasm.

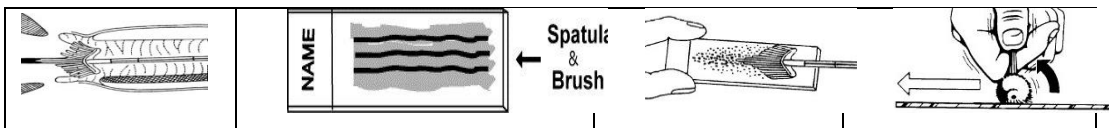


Figure 2. Extraction of Vaginal Cells and Staining with Reagents for Microscopic Observation

In the squamous epithelium there are 4 layers of cells. The cells start out being formed at the basal layer while maturing they move out through the parabasal layer, the

intermediate layer and at last out through the superficial layer. The cells in the basal layer divide and deliver cells to the layers above it. As illustrated at Figure 3, and Figure 4. while the cells mature and move through the layers they change shape, color and other characteristics [30]. When the cells reach through the superficial layer they are rejected and replaced by the cells coming from below. The basal layer has small round cells with a relative big nucleus and a little cytoplasma. When maturing, the nucleus becomes smaller and the cytoplasma becomes bigger. The shape of the cells become less round the more mature they are. The columnar epithelium only contains a single layer of cells, the basal layer here contains columnar cells and reserve cells. The reserve cells divides into new reserve cells and into columnar cells. The metaplastic epithelium consists of reserve cells from the columnar epithelium. When the cells have matured fully in the metaplastic epithelium, they look like; the cells found in the squamous epithelium. When the genetic information in a cell somehow has changed, the cell will not divide as it should. This is a pre-cancerous cell [31].

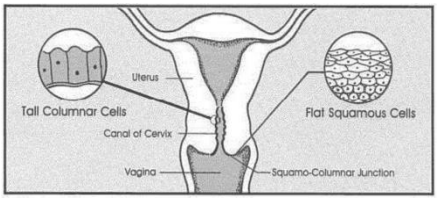


Figure 3. The Uterus in Details and the Location of a) Columnar Cells and b) Squamous Cells

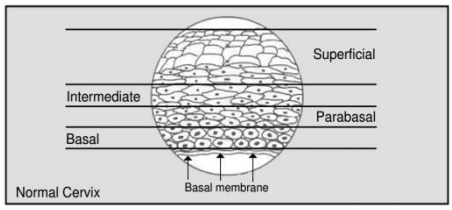


Figure 4. Development of the squamous Cells through the Four Layers (Image credit [37])

Depending on which kind of cell that is dividing incorrectly, it diagnoses like dysplasia and 'carcinoma in situ'. The dysplastic cells are divided up into mild, moderate and severe dysplastic. The graduation into different degrees of dysplasia are determined from the probability of the cells later on turning into malignant cancer cells. A high amount of the mild dysplastic cells will disappear without becoming malignant, whereas severe dysplastic cells quite likely will turn into malignant cells. In medical terms these are divided into 2 different main diagnoses such as '*Dysplasia*' and '*Carcinoma in Situ'*

3.1.1. Dysplasia: The Term "*plasia"* Means growth, and dysplasia means disordered growth. The cervical dysplasia are normally divided into 3 types: *mild, moderate and severe*, describing the risk that the cells turn into malignant cancer cells. [32] *Mild* means of course lowest risk. The characteristics of cells in dysplasia depends on the kind. In the mild dysplasia they have enlarged and light nucleus. For the *moderate dysplasia* the nucleus is larger and darker. In *severe dysplasia* the nucleus is large, dark and often deformed. The cytoplasm is dark and small when compared to the nuclei.

3.1.2. Carcinoma-in-situ:'Carcinoma-in-situ' means 'cancer in place' and is characterized by very large nucleus. In the past, there was a tendency to treat '*carcinoma-in-situ*' as a much more serious issue than *severe dysplasia*, in fact they are essentially the same. The pre-malignant cells are characterized by a larger Nucleus (N) with Cytoplasm (C) and a bigger N/C (ratio given by [33-34]

$$\frac{N}{C}Ratio = \frac{NucleusArea}{(NucleusArea+CytoplasmArea)} \quad (1)$$

3.2. Bethesda System

To describe the pap test results The Bethesda System (TBS) is used. This was developed in 1988 and had been revised in 1991 and 2001[35-36]. There are 6 BCC

(Basel Cell Carcinoma) codes as illustrated at Figure 5, with in the Bethesda System such as Negative for intraepithelial lesion or malignancy, Atypical squamous cells of undetermined signficance (ASC-US), Low grade squamous intraepithelial lesion (LSIL). This includes HPV and mild dysplasia, Atypical squamous cells cannot exclude high grade SIL (ASC-H), High grade squamous intraepithelial lesion (HSIL) encompassing moderate dysplasia (CIN2), as well as CIN 3, which includes severe dysplasia Carcinoma In Situ (CIS), Squamous cell carcinoma, Abnormal glandular cells including atypical glandular cells of undetermined significance (AGUS): endocervical adenocarcinoma; endocervical adenocarcinoma in situ; endometrical adenocarcinoma; extrauterine adenocarcinoma; adenocarcinoma.[37]. Clinical experts use the Bethesda System to describe the Pap test results and usually send the image processing centre where advanced microscopes used to identify pre-cancerous stage.

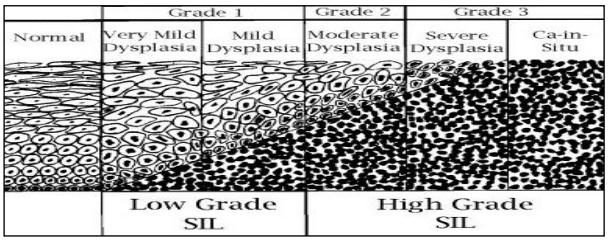


Figure 5. CIN Grades (Image Credit [37])

4. IoT Driven Clinical Analysis Architecture (IDCAA)

IDCAA highlights on heterogeneous healthcare needs with cost effective enterprise approach by making use of the available Internet of Things and Cloud Platform to automate the process of clinical diagnosis. This paper highlights the automation of Pap smear image analysis as a use case in the implementation. With over flourished mobile devices, self-communicating sensors and data transmitting devices data can be easily transmitted through centralized processing platform and depending on the need the solution can be leveraged. In healthcare scenario one could enable patient monitoring sensors, image capturing devices, X-Ray machines, microscopes, etc… to be tied to the services of IDCAA implemented as RESPRO 3.0, which is running on the Cloud Platform to process any healthcare data analysis and processing needs.Currently RESPRO 3.0 deals with Pap smear image analysis and it can be extended to any clinical needs in future into Software as a Service (SaaS) model.

4.1. Service Oriented Architecture

To implement this model, we propose [9] SaaS driven client-server architecture, where Server acts as service provider and client requests for services. The idea behind the proposal is to make the cell pattern image analysis method to be publicly avaialble for cheaper cost with greater quality. SOA processes the malign cells and provide results as per the diagnosis expected by clients. Experts can define and update the image pattern rules, recent trends, comments based on diagnosis. Clinical experts can meet through this solution and colloborate among themselves for an ideal diagnosis.

4.2. Architecture

The RESPRO 3.0 services has multiple facets; RESPRO Services exposed to get inputs from IoT devices using IDCAA model, Individual Clinics, Agencies, Laboratories and patients or anyone who is willing to get the intelligence diagnosis. RESPRO Bridge which receives the input messages from the RESPRO Services and process the initial message to the relevant RESPRO adapters which can pick up the message from the source and start

processing for the specific results. The processing includes image processing, signal analysis, results ground truth results comparison of specific input range for ICDs. Associated automation steps pertaining to image processing methods and techniques specific to each ICD is enabled by specific adapters. RESPRO Knowledge Base (KB) houses all the processing logics specific to each ICD and automations is constructed using meta-model approach. The meta-model which combines the relational structure of each ICD with lab test. Clinical details, bench mark figures for the true positive and true negatives, Sample inputs,

sample outputs, exceptional criteria, categories of test results, test result appearance, Colours of reagents, chromatic representations of the results sets. As illustrated at Figure 6. this KB is updated as an ongoing live feedback mechanism, thus new agents can enhance and develop new ICD adapters in the future. This solution can be an incompetent forum for various interested stakeholders such as healthcare agencies, doctors, patients, to take part and mature the systems and develop a colloborated, well matured automation systems.

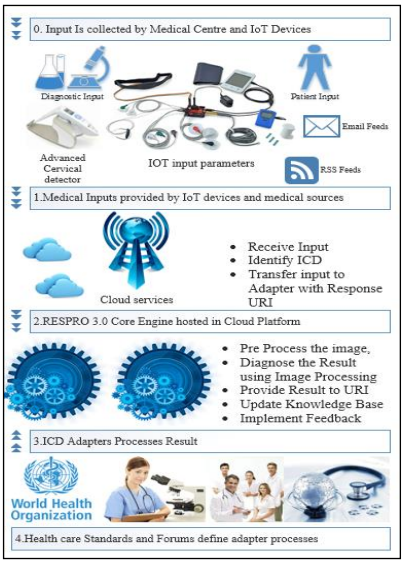


Figure 6. RESPRO 3.0 Architecture

a) Process flow at IoT Devices for Pap Smear Processing

- Blood or Tissue sample is collected from patient and stained based on Pap stain method
- A wireless camera (in the case of low end microscopes) is fixed to the microscope and connected to the web based software to transmit the images seen through microsoft to computer.
- The slide is exposed to the microscope, Cytologist observes and mark specific region of the slide for area of interest captures images with integrated camera.
- Image processing filters removes noise and enhances the image for processing.
- Once the image is filtered from noise, using RESPRO 3.0 services client pushes the image to the SOA architecture, where services are exposed for any device to consume the services and POST the image for processing.

b) Process flow at the Server

- SOA in the form of Windows Communication Foundation (WCF) services hosted in the cloud for anybody to access it free of cost and it acts like a core engine.
- SOA Web Service Function receives images from IoT client, and identifies the following parameter such as Test Name, Images and specific test conditions.

- After verification and validation, the received image is transferred to specific adapter that can perform image processing operations on the image and provide results.
- RESPRO engines has an inbuilt adapter for Pap Smear Processing adapter

c) Process flow at the Specific Adapter (Cervical Cancer Prediction Adapter)

- Looking at the image, Pap smear adapter process the images
- The details of the specific test details are stored in the form of meta models for decision making and provides the results as illustrated below at Figure 7.
- Details are captured and updated from KB that contains repository of image patterns with respect to slide tests, tissue patterns.
- KB hosts all the updated image processing tools for pap smear processing, in our case the image processing will apply SVM Filters for analyzing pap smear images
- The algorithms stored in the KB based on the expert's advice, community practices, forums, cytologists experience and specific adapter is tuned to perform with any input criteria for Pap smear test.
- Based on the test conditions the results are provided to the IoT device which initiated the requests
- Industry experts, Doctors, web sites, Expert systems, Healthcare software agents can update IDCAA functionality as per the recent standards channelized by experts group

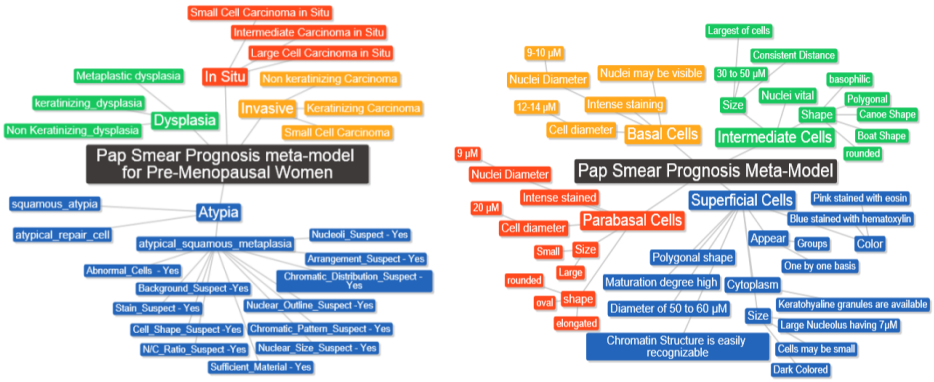


Figure 7. Architecture Meta Model Construction for Pre-Cancerous Detection used by Cervical Cancer Predicting Adapter

4.3. Automation Stages and Ranking Features

Features extracted from images using image-processing techniques such as Sober Edge detection, Prewitt Filters, Canny Edge detection, Red Filter, Green Filter, Blue Filter, Grey Scale Filter, Invert Filter, Brightness and contrast adjustments over the images provide means to operate on the images to extract the feature listed in Figure 8. Image Processing techniques involve starting from image registration, archival into RESPRO 3.0 native database, every feature once processed stored against the images so that it can be retrieved anytime for further processing or comparison during the ongoing image processing stages.

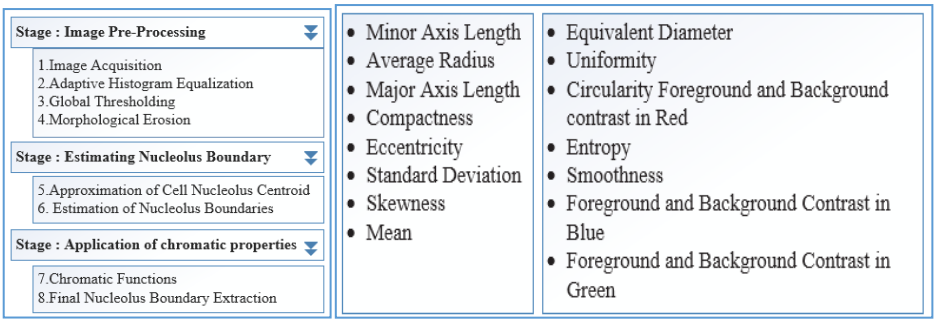


Figure 8. Automation Stages& Ranking Features

5. Support Vector Machines (SVMs)

Support Vector Machines are supervised models with associated learning algorithms that analyze data and recognize patterns, they are used for classification and regression analysis [39]. SVM is a design of classification of both linear and nonlinear data. It converts the original data into higher dimension. SVM is based on supervised learning where it is popularly used at various fields of science inclusive of object recognition and speech recognition. SVM performs classification by finding the hyperplane that maximizes the margin between the two classes. The vectors (cases) that define the hyperplane are the support vectors. SVM uses nonlinear mapping to convert the orignial data into higher dimension. Its objective is to construct a function which will correctly predict the class to which the new points belong and the old points belong. With appropriate nonlinear mapping, two data sets can always be divided by hyperplane. Hyperplane separates the tuples of one class from another and defines a decision boundary. [40] There are many hyperplanes that separates the data but only one will achieve maximum seperation. When data is non-linear and data set is inseperable then we use Kernels. SVM Becomes prominent when we use pixel maps as input. SVM has been performing extremly well in the field of pattern classification problems.

5.1. Binary Classification Using SVM

SVM is primarily designed for binary classification the various properties of SVM are duality, robust, kernel, margin, convexity and sparseness. [41-42] SVM describes two classess primarily

1) Scenario where data are linear separable
2) Scenario where data are linear inseperable

From Figure 9, The SVM derivative can be described as, given some training data D, a set of n points of the form $D=\{(x_i,y_i \mid x_i \in \mathbb{R}^P, y_i \in \{-1,1\}\}^n{}_{i=1}$ where the y_i is either 1 or −1, indicating the class to which the point x_i belongs. Each x_i is a P-dimensional real vector. To find the maximum-margin hyperplane that divides the points having $y_i = 1$ from those having $y_i = -1$

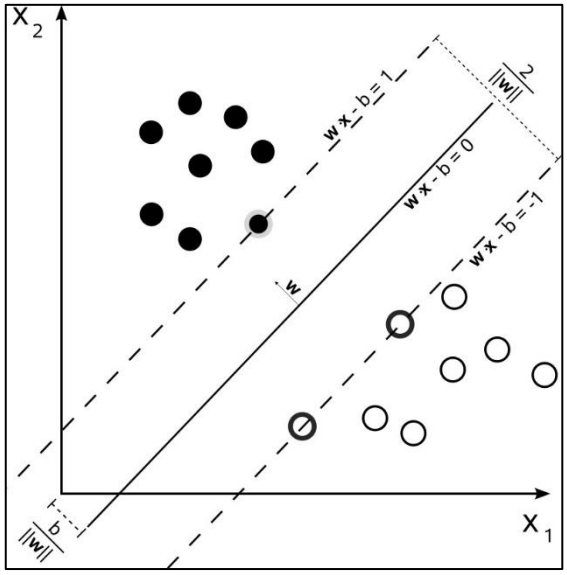


Figure 9. Depicting Support Vector Parameters Located at both Sides of Maximum-Margin Hyperplane

Maximum-margin hyperplane and margins for an SVM trained with samples from two classes. Samples on the margin are called the support vectors. $W.X - b = 0$ where . denotes the dot product and W the normal vector to the hyperplane. The parameter $\frac{b}{||W||}$ determines the offset of the hyperplane from the origin along the normal vector W. If the training data are linearly separable, we can select two hyperplanes in a way that they separate the data and there are no points between them, and then try to maximize their distance. The region bounded by them is called "the margin". These hyperplanes can be described by the equations$W.X - b = 1$ and $W.X - b = -1$ Geometrically, the distance between these two hyperplanes is $\frac{2}{||W||}$, so to maximize the distance between the planes we want to minimize ||W||. As we also have to prevent data points from falling into the margin, we add the following constraint: for each i either $W.Xi - b \geq 1$ for x_i of the first classor $W.Xi - b \leq -1$ for X_i of the second.This can be rewritten as$y_i(W.X_i - b) \geq 1 \, for \, all \, 1 \leq i \leq n$. We can put this together to get the optimization problem, Minimize (in w , b)||W|| subject to (for any i=1,…,n) $y_i(W.X_i - b) \geq 1$. Figure 10, provides the the description of how input space is received from the SVM algorithm and the classsification done using dynamically identified SVM parameters.

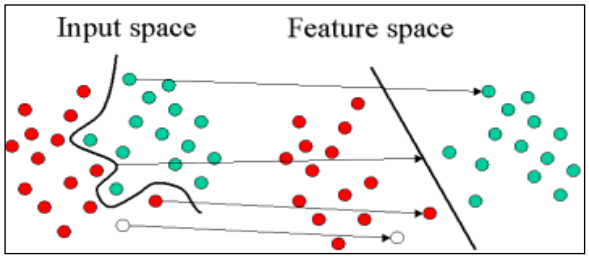


Figure 10. Input Features Converted into Features Space while Performing SVM Binary Classification

5.2. Multiclass Support Vector Machine (MSVM)

There are two most popular approaches described below when handling with Multiclass classification through binary data such as, One-Vs-One and One-Vs-All (One-

versus-rest classifiers). [43-45] Our experiments is based on one-versus-all classifier where each category is split out and all of the other categories are merged and to choose the class which classifies the test data with greatest margin. It divides an m class problem into m binary problems. The learning step of the classifiers is done by the whole training data considering the cervical cancer cell patterns from the labelled class as positives and all other cervical cancer cells as negatives. In validation phase, a pattern is presented to each one of the binary classifiers and then classifier which provides a positive output indicates the output class. In numerous cases, the positive outcome is not unique and some tie-breaking techniques are handled internally with threshold values. The most familiar approach uses the confidence on the classifiers to decide the last outcome, predicting the class from the classifier with the maximum confidence score.

5.2.1. Kernel Trick: The Kernel Trick is a very interesting and powerful tool. It is powerful because it provides a bridge from linearity to non-linearity to any algorithm that can be expressed solely on the terms of dot products between two vectors. [46-47]. It comes from the fact that, if we first map our input data into a higher-dimensional space, a linear algorithm operating in this space will behave non-linearly in the original input space. Now the Kernel Trick is really interesting because that mapping does not need to be ever computed. If our algorithm can be expressed only in terms of an inner product between two vectors, all we need is replace this inner product with the inner product from some other suitable space. That is where resides the "trick": wherever a dot product is used, it is replaced with a Kernel function. [48] The kernel function denotes an inner product in feature space and is usually denoted as,

$$K(x, y) = < \Psi(X), \Psi(Y) > \quad (2)$$

Using the Kernel function, the algorithm can then be carried out into a higher-dimension space without explicitly mapping the input points into this space. This is highly desirable, as sometimes our higher-dimensional feature space could even be infinite dimensional and thus unfeasible to compute the desired results.

5.2.2. Polynomial Kernel: For the classification of multi class cervical cancer cells, polynomial kernel is chosen due to the well-structured samples we had gathered. [49-50] The Polynomial kernel is a non-stationary kernel; Polynomial kernels are well suited for problems where all the training data is normalized. Adjustable parameters are the slope alpha, the constant term c and the polynomial degree d

$$K(x, y) = (\alpha X^T y + c)^d \quad (3)$$

It is highly believed that MCSVM performs superior to conventional statistical and neural network classifiers. MCSVMs deliver state-of-the-art performance in real-world applications such as text categorization, hand-written character recognition, image classification, bio sequences analysis, *etc.*, and are now established as one of the standard tools for machine learning and data mining. The goal of the SVM is to separate multiple clusters with the set of unique hyperplanes that have the greatest margins to the edge of each clusters. [51-52] Various kernels including Polynomial, Radial Basis Functions and Hyperbolic tangent can be used for mapping the original sample space into a new Euclidean space and then the linear classifier can be designed for classification. Since, Polynomial kernel function is a non-stationary kernel, Polynomial kernels are well suited for problems where all the training data is normalized, because the input features used in our classification approaches are well normalized and we preferred Polynomial Kernel over other kernel functions.

5.3. Proposed Density based Multi Class Support Vector Machine(DMCSVM)

SVMs were developed to participate in binary classification. Nevertheless, functions of bifold classification are very restrained exceptionally in cancer classification the place most of the classification problems involve more than two classes. A number of methods to generate multiclass SVMs from binary SVMs have been proposed by researchers and is still a continuing research topic. This section provides a brief description of some methods implemented to solve multi-class classification problem with SVM in present study.

In this paper, [53] a binary determination tree structure that uses Density based SVM for making the binary selections within the nodes. The DMCSVM-BDT (Density based Multi Class Support Vector Machines utilizing Binary Decision Tree) classifier architecture takes advantage of both the efficient computation of the tree architecture and the high classification accuracy of SVMs. Utilizing this architecture, N-1 SVMs needed to be trained for an N SVMs are $\lceil \log 2N \rceil$ class problem, best at most required to be consulted to classify a pattern. This may result in a dramatic improvement in recognition speed when addressing problems with tremendous quantity of classes to competently replicate the traits of the goal data set; we recommend a density weighted for Multi category support vector machine. Within the proposed DMCSVM, the relative density of each information factor situated on the density distribution of the goal category label making use of the ok-nearest neighbour (k-NN) procedure. The distance between xi and the k^{th} nearest neighbour of x_i is denoted as$\left(x_i, x_i^k\right)$; where $\left(x_i^k\right)$ is the kth nearest neighbour of data point xi. Using k-NN distance, the density weight of data point x_i is defined as

$$w_i = 1 - \frac{d\left(x_i, x_i^k\right)}{\max_{j \epsilon \, trainset} d(x_j, j)} \tag{4}$$

Density weight measures the relative density centered on the density distribution of the target data by comparing the k-NN distance of each and every data point with the maximum k-NN distance of the dataset. Density weight falls within the range $0 \leq w_i \leq 1$. To measure the density weight in feature space, can use the kernel function to map knowledge into high dimensional area. The distribution of the data in feature space may be exclusive from the original data distribution. So as to obtain a more appropriate description, we estimate the density weight in real space. With the density weight estimation method in the equation (4), an information point placed in a comparatively excessive-density area is close to its neighbours, so the distance between that data point and its k^{th} nearest neighbour decreases, and eventually the density weight will become larger. In relatively low-density areas, data points are far from each other, so the density weight value will be low. To apply the density weight, the objective function is defined as:

$$\min_{R,C,\xi_i} R^2 + C \sum_{i=1}^{N} w_i \xi_i$$

$$s.t. \|\phi(x_i) - c\|^2 \leq R^2 + \xi_i, \quad \forall i = 1,,,, N. \tag{5}$$

We impose the weight w_i on each data pointx_i . The data point in excessive-density regions acquires a larger weight, so the outcome of the slack variable is compounded. Therefore, to minimize the objective function, the spherical description will shift toward the high-density regions. On the other hand, with decreasing weight in relatively sparse areas, the influence of each data point will be reduced and there is no pressure to keep data lying outside the spherical description. By introducing the Lagrangian function for

(5), and let partial differentiation of the Lagrangian [59] function is equal to 0, we have the Wolf dual form

$$\max_{\beta} \sum_{i=1}^{N} \beta_i k(x_i, x_i) - \sum_{i,j=1}^{N} \beta_i \beta_j k(x_i, x_i) \quad s.t. 0 \le \beta_i \le w_i c_i$$

$$\sum_{i=1}^{N} \beta_i - 1, \forall_i - 1,2, \dots N \tag{6}$$

Notice that, the upper bounds for Lagrange multipliers = i N β, 1, i are no longer the same instead, each of them is respectively controlled by the corresponding weight. The primal variables can be recovered from the optimal β as 1

$$C = \sum_{i=1}^{N} \beta_i \emptyset(x_i), \quad R = \sqrt{\beta^T \, diag(K) - \beta^T \, K\beta} \tag{7}$$

Therefore, by introducing density weight into the search for the optimal description of the dataset, can shift its description boundary to dense areas. In our proposed DMCSVM, update core set by using density weight and then train the core set by using SVM algorithm.

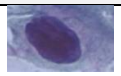
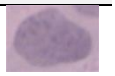
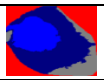
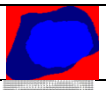
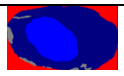
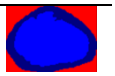
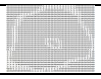
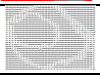
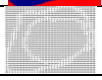
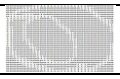
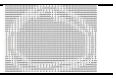
The DMSCVM algorithm is introduced as follows, Here, the core set, the ball's center, and radius at the t^{th} iteration are denoted by St, ct, and Rt respectively and input of a termination parameter ε.

Step 1): Initialize ε,t=0, by S_t ,c_t , and R_t
Step 2): Update the core set: Terminate if there is no training point
Step 3): Represend each data point as vector
For each vector d∈ D do
Using k-NN distance, the density weight of data point xi is defined in equation (4)
impose the weight w_i on each data pointx_i in equation (5)
Find = such that ϕ (z) is furthest away from c_t in the corresponding feature space and set S_{t+i}= S_t
Step 4): Find the new MEB: The C_{t+1} and R_{t+1} are computed by equation (7).
Step 5): Set t=t+1 and go back to step 2).
Step 6): Train the core set using SVM algorithm.

6. Experimental Classification Results and Analysis

The below table Table 1, illustrates a sample of Pap smear cells with its extracted features as per Figure 8, calculations invovled in extracting these feature during image processing are referred here [33]. These features used as input to DMCSVM algorithm that we developed for the test set and multi classes were identified.

Table 1. Pap Smear Cells with Expert Identified Class and their Calculated Ranking Features*

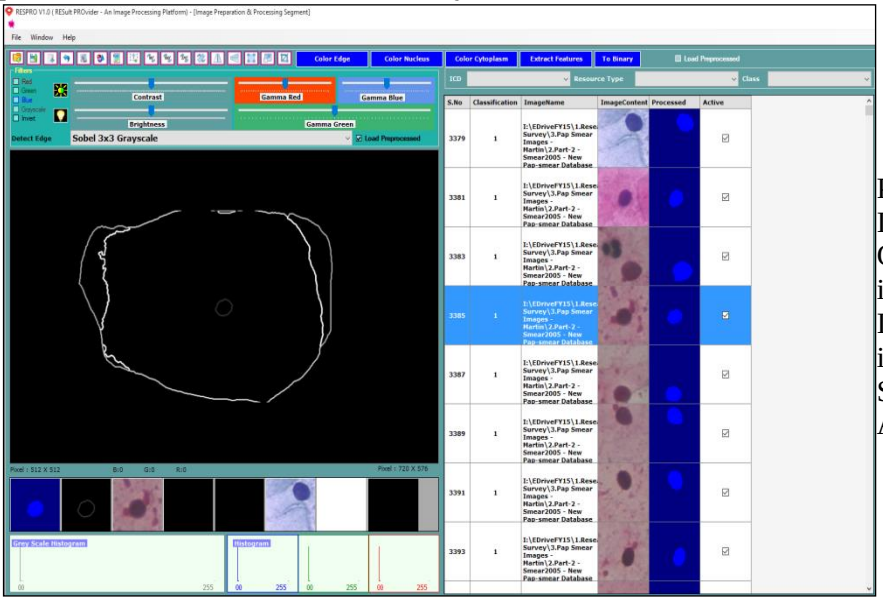
N= Nucleus C= Cytoplasm	Normal Columnar	Normal Superficial	Normal intermediate	Light dysplastic	moderate dysplastic	Severe dysplastic	Carcinomain Citu
Cell Image							
Cell Mask							
Binary image							
N_Area	804	1475	2396	3091	4527	5908	3271
C_Area	27804	65192	5390	6166	4328	14109	1816
N/C ratio	0	0	0	0	1	0	1
N_YCol	86	118	144	90	69	71	141
C_YCol	193	217	204	146	95	130	170
N_Short	30	33	43	60	72	72	57
N_Long	35	55	71	77	80	107	79
N_Elong	1	1	1	1	1	1	1
N_Round	1	1	1	1	1	1	1
C_Short	182	302	93	120	103	136	70
C_Long	242	340	127	123	132	195	96
C_Elonged	1	1	1	1	1	1	1
C_Rounded	1	1	0	1	0	0	0
N_Perimeter	101	141	188	257	245	285	228
C_Perimeter	674	1062	332	433	359	666	257
N_Position	0	0	0	0	0	0	0
N_Maximum	44	65	86	95	213	215	108
N_Minimum	37	49	55	60	213	210	79
C_Maximum	649	1728	114	172	126	357	56
C_Minimum	655	1674	147	158	189	343	91

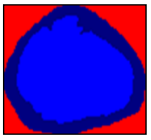
*numbers to be measured in Nano Meter^2

6.1. Image Pre Processing and Preparation for Training Set

We followed 80:20 distribution model to split our data for training and testing on the experts pre-classified 917 normal pap smear images obtained from here [33] a nd collected from Three hospitals from Bangalore and Vellore of south India. Table displays features of each sample of well known cervical cancer's seven class classification. Images are stored in the Sql Server 2012, using SQL procedures images are retrieved, using byte array data type images are archived. An extensive image processing library natively developed using dot net framework 4.5 using visual studio 2013 as part of RESPRO 3.0 Framework. The basic representation of the image is bitmap type and handles 512 X 512 to 720X 720 pixel for image processing. Figure 11, details our tool's help in pre processing each image using various image processing capabilites such as Red Pass Filter, Green Pass Filter, Blue Pass Filter, Invert and Grayscale capabilities, managing contrast,

brightness, Gamma Red, Gamma Blue, Gamma Green, and Edge Detections such Laplasian, Sobel, Prewit, Kirsch and many others.





Pap Smear Image Cell Converted into binary Image which is fed into SVM Algorithm

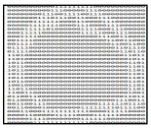


Figure 11. RESPRO 3.0 Architecture's Image Processing User Interface

6.2. Pap Smear Image Processing through Multi Class Support Vector Machine

For SVM training each images is feature extracted and processed, Sobel 3X3 Grayscale filter is applied on top of that Invert filter is applied in order to differentiate the nucleous and cytoplasm. In order to process the image into DMCSVM training, the images are converted into a binary image, once converted to binary image all the images are stored in a flat file with its predefined class. So every image is now represented as 32X32 matrices. They are also available in a pre-processed form in which digits have been divided into non-overlapping blocks of 4X4 and the number of on pixels have been counted in each block. This generates 8X8 input matrices where each element is an integer in the range of 0..16. Figure 12, details how we train the DMCSVM with experts pre-classified pap smear images into our tool.

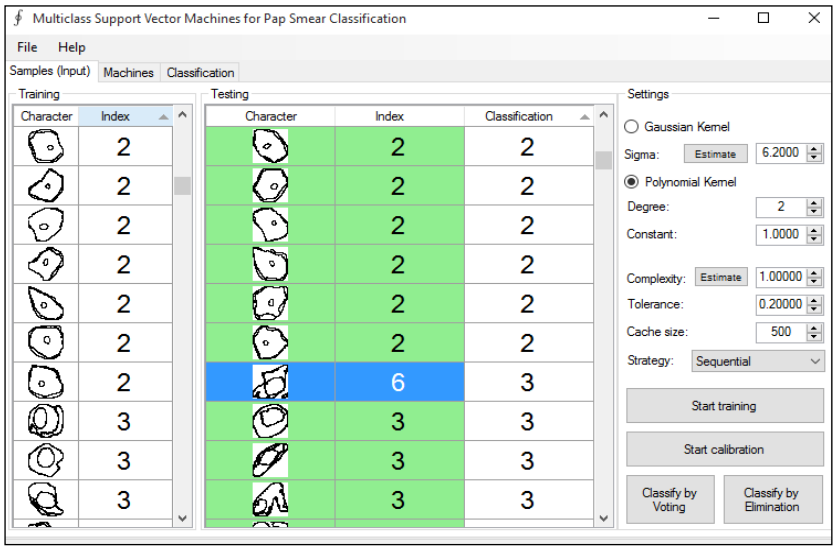


Figure 12. Training Images Classified Automatically using Polynomial Kernel Trick with Multiclass SVM

All the pap smear images are labelled and imported into the Density based Mutli class support vector machine processing tool. [54] Polynomial Kernel function is chosen for the processing with the degree as 2 and constant value as 1,. Complexity is chosen as 1, tolerance is given as 0.2 and sequential strategy is chosen for training the data, these parameters are derived based on the data visualization and best converging points in differntiating each classes. As indiciated above the total of 917 images are seprated into 80:20 ratio for training set (734) and test set (183). [55-56] Customized software has successfully got trained with the 80% of images, Figure 13, to Figure 20, provides the snapshot of rightly classified pap smear cells for the seven class categories using our tool. There are some images found as true negatives and did not respond successful and some other as false positivies . The performance of the system is explained in section 6.3 of this paper that details each classes and its performance.

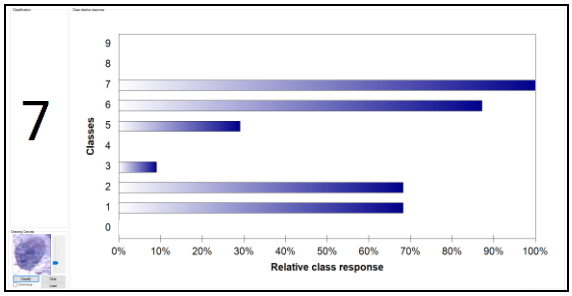


Figure 13. 'Carcinoma in Situ' Cell Classified Accurately as 7th Class

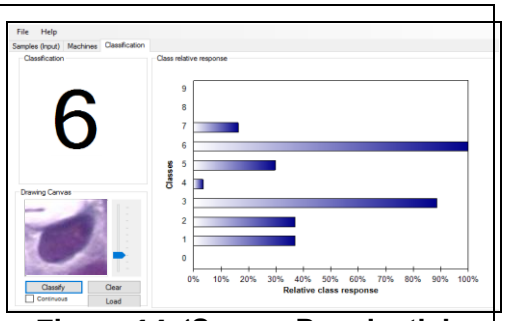


Figure 14. 'Severe Dysplastic' Cell Classified Accurately as 6th Class

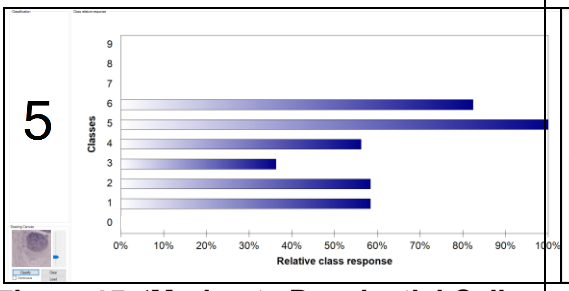


Figure 15. 'Moderate Dysplastic' Cell Classified accurately as 5th Class

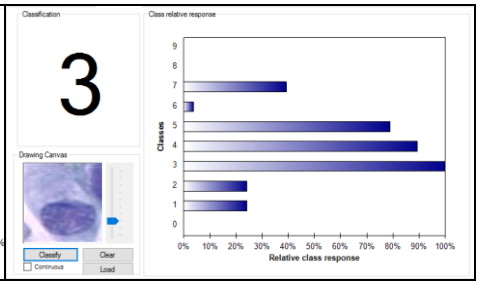


Figure 17. 'Normal Intermediate' Cell Classified Accurately as 3rd Class

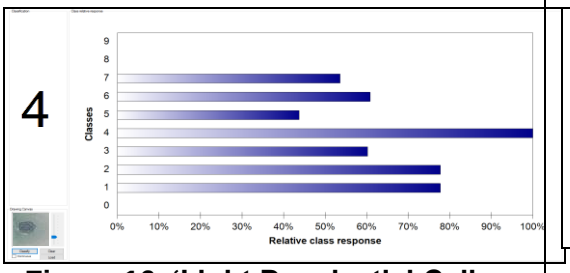


Figure 16. 'Light Dysplastic' Cell Classified Accurately as 4th Class

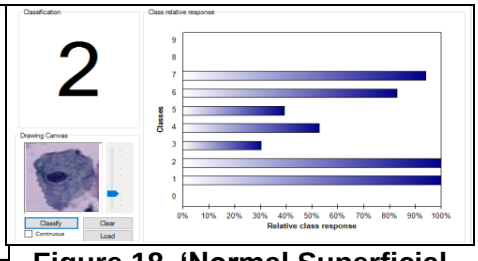


Figure 18. 'Normal Superficial Cell Classified Accurately as 2nd Class

Samples (Input) | Machines | Classification

Machine	Subproblem	Support Vectors	Threshold
1	1-vs-0	0	-1
2	2-vs-0	0	-1
3	2-vs-1	101	0.445876373719
4	3-vs-0	0	-1
5	3-vs-1	100	0.802884675195
6	3-vs-2	98	0.567720201530
7	4-vs-0	0	-1
8	4-vs-1	156	0.058765523208
9	4-vs-2	160	-0.43699037807
10	4-vs-3	186	-1.32690662175
11	5-vs-0	0	-1
12	5-vs-1	137	0.499371770195
13	5-vs-2	137	0.158071211780
14	5-vs-3	173	-0.78201951912
15	5-vs-4	233	0.808928861283
16	6-vs-0	0	-1
17	6-vs-1	144	0.861590931503
18	6-vs-2	144	0.449902392733

Figure 20. Multi class SVM Augments Scoring Based on One-Versus-All Method

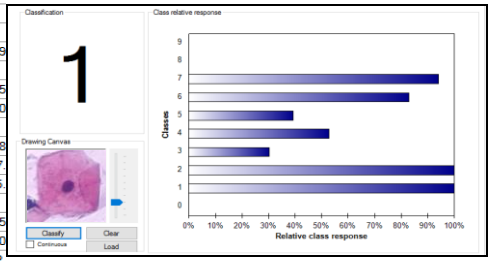


Figure 19. 'Normal Columnar Cell Classified Accurately as 2nd Class

6.3. Performance Assessment

Following are some of the existing results highlighted for comparative study in relation with the our findings of this paper

Prior Study	Method	Groups	Train	Test	Features	Accuracy	Error Rate
1 [57]	SVM - Polynomial Kernel	7	39	1	15	~98%	~1%
2 [58]	SVM - Polynomial Kernel	5	32	8	68	~90%	~10%
3 [59]	SVM - Linear Kernel	7	917	217	9	~93%	~7%
This Paper	MultiClass SVM Polynomial Kernel with Density Features	7	734	183	20	~96%	~4%

The cytopathological experts' classification of cells was assumed to the gold standard in this investigation. For the proposed DMCSVM classifier makes the decision of binary classification and classifies the seven class classfier based on the proposed approach, Table 2. depicts the confusion matrix of each pap smear cell being identified by our algorithm, accompanied with calculated accuracy and error rates were used to assess its classification performance. [60-61] The accuracy, senstivity and specificity of the model were compared with the clinical expert's judgement, the below confusion matrix provids better view of the performance of this algorithm; where TP = True Positive, TN= True Negative, FP = False Positive, FN = False Negative, Over all Accuracy $\frac{TP+TN}{(TP+FP+TN+FN)} =$ 96%, Sensitivity $\frac{TP}{(TP+FN)} = 92\%$, Specificity $\frac{TN}{(TN+FP)} = 99\%$, Table 3. provides the overall performance of the algorithm such as Positive Predictive Value PPV = 92%, Negative Predictive Value (NPV) = 99%, Informedness (Sensitivity + Specificity - 1) = 92%, Markedness (Precision + NPV -1) = 92%. The results are indicative that, Multiclass Support Vector Machines proven to be a good approach while solving multi class problems with density parameters taken into account.

Table 2. Confusion Matrix for All Groups (Displayed in Grey) of Pap Smeared Cells along with Class Level Performance

Cell Type	Normal Columnar	Normal Superficial	Normal Intermediate	Light Dysplastic	Moderate Dysplastic	Severe Dysplastic	carcinoma in situ	True Positive	False Positive	False Negative	True Negative
Normal Columnar	13	1	1	0	0	0	0	13	2	1	165
Normal Superficial	1	12	0	0	0	1	0	12	2	0	166
Normal Intermediate	0	1	17	0	0	0	0	17	1	1	161
Light Dysplastic	0	0	1	35	0	0	0	35	1	0	143
Moderate Dysplastic	0	0	0	0	27	0	0	27	0	0	151
Severe Dysplastic	0	1	0	1	1	37	0	37	3	0	141
Carcinoma in situ	0	0	0	0	1	2	25	25	3	2	153
								166	12	4	NA

Table 3. Snapshot of Overall Performance for Sensitivity, Specificity & Accuracy

Cell Type	TPR	TNR	PPV	NPV	FPR	FDR	FNR	Accuracy	Informedness	Markedness
normal columnar	.93	.99	.88	0.99	.01	0.13	0.07	0.20	0.92	0.87
normal superficial	.86	1.0	1.0	0.99	00	0.00	0.14	0.20	0.86	0.99
normal intermediate	.95	.99	.95	0.99	.01	0.05	0.05	0.20	0.94	0.94
light dysplastic	.97	.99	.97	0.99	.01	0.03	0.03	0.20	0.97	0.97
moderate dysplastic	1.0	1.0	1.0	1.0	00	0.00	0.00	0.20	1.00	1.00
severe dysplastic	.93	.99	.95	0.98	.01	0.05	0.08	0.19	0.91	0.93
carcinoma in situ	.90	.99	.93	0.98	.01	0.07	0.10	0.19	0.89	0.91
Average	.93	.99	.95	0.99	.01	0.05	0.07	0.20	0.93	0.94

7. Conclusion and Suggestions

This paper has emphasised that SVM outperforms multiclass problems especially in the realm of Cervical Cancer predication. It still performs better with larger dataset when each of the cervical cancer image features are attributed with a density parameters. This highlights the quickest convergence of predictions. This paper also has highlighted a cloud driven solution framework to the needs of the era; as Cervical cancer is posing a critical threat to the women's health condition, determining the pre cancerous stage affirmatively helps calibrate the treatment procedures. The service oriented architecture enabled with adapter driven solution can be scaled up to growing any ICD's requirement. The prominence of the web and the embracement of the open source community can contribute to the solution in their own way trough this Software as a Service Model.

Acknowledgements

Author would like to thank Dr.Devarajan, Prof/EEE., Govt College of Technology, Dr. P. Uma Maheshwari, Prof/CSC., Info Institute of Engineering, for their valuable guidance and support. We also extent our gratitude to the Cytologists, Nurses of Baptist Hospital Bengaluru for their help in collecting data at various stages for this worthwhile effort.

References

[1] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, "Molecular Biology of the Cell", 5th ed. New York: Garland Science, **(2007)**.
[2] J. Jantzen, J. Norup, G. Dounias and B. Bjerregaard, "Pap-smear Benchmark Data For Pattern Classification", **(2002)**.
[3] A. Meisels and C. Morin, "Cytopathology of the uterus", ASCP Press, 2. Edition, **(1997)**.
[4] P. Soda, "Early Experiences in the Staining Pattern Classification of HEp-2 Slides", Twentieth IEEE International Symposium on Computer-Based Medical Systems (CBMS'07).
[5] A. Botta, "University of Napoli Federico II, Naples, Italy ; de Donato, W. ; Persico, V. ; Pescape, A. On the Integration of Cloud Computing and Internet of Things, published at Future Internet of Things and Cloud (FiCloud), 2014 International Conference on, 10.1109/FiCloud.2014.14, **(2014)**, pp. 23-30.
[6] F. Meyer, "Automatic screening of cytological specimens", CVGIP, vol. 35, **(1986)**, pp. 356-369.
[7] D. Pycock and C. J. Taylor, "Use of the magiscan image analyzer in automated uterine cancer cytology". Analytical and Quantitative Cytology, vol. 2, **(1980)**, pp. 195-202.
[8] N. U. Tanaka, Ikeda, Ueno, Watanabe, Y. Imasato and R. Kashida, "Fundamental study on automaticcyto screening for uterine cancer 111: New system utilizing the pattern recognition method." ActaCytologica, vol. 21, **(1977)**, pp. 85-9.
[9] J. H. Tucker, Burger, Husain, Ploem, Schenck, Schwarzmann and Stenkvist, "Measuring the accuracy of automated cervical cytology prescreening systems based on image analysis."CAAC Com of the Europ Com, **(1987)**, pp.1-51.
[10] D. J. Zahniser, P. S. Oud, M. C. T. Raaijmakers, G. P. Vooys and R. T. Van de Walle, "BioPEPR: "Asystem for the automatic prescreening of cervical smears", Histochem Cytochem, vol. 27, **(1979)**, pp. 635-41.
[11] W. Mitlehner, E . Cronin, J. Bronzino nd R. Veranes, "Cytopath: "An Expert System For The Classification And Diagnosis Of Squamous Lesions In The Pap Smears Of Pre-Menopausal Women. Department of Engineering", Hartford, CT, **(1990)**.
[12] J. S. J. Lee, J. N. Hwang, D. T. Davis and A. C. Nelson, "Integration of Neural Networks and Decision Tree Classifiers ForAutomated Cytology Screening", **(1991)**.
[13] R. F. Walker, P.Jackway, B. Lovell and I. D. Longstaff, "Classification Of Cervical Cell Nuclei Using Morphological Segmentation And Textural Feature Extraction"Department of Electrical and Computer Engineering University of Queensland, **(1994)**.
[14] R. Balasubramanian, S.Rajan, R. Doraiswami and M. Stevenson, "A Reliable Composite Classification Strategy", The University of New Brunswick, Fredericton, Canada, **(1998)**.
[15] K. Tumer, N. Ramanujam, J. Ghosh and R. R. Kortum, "Ensembles of Radial Basis Function Networks for Spectroscopic Detection of Cervical Precancer", IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, vol. 45, no. 8, **(1998)** August.
[16] Z. Li and K. Najarian, "Automated Classification of Pap Smear Tests Using Neural Networks", Department of Computer Science, IEEE, **(2001)**.
[17] Y. Liu, T. Zhao and J. Zhang, "Learning Multispectral Texture Features For Cervical Cancer Detection", The Robotics Institute, Carnegie Mellon University,Pittsburgh, PA, I EEE, **(2002)**.
[18] P. Sobrevilla, E. Lerma and E. Montseny, "An Approach to a Fuzzy-based Automatic Pap Screening System –FAPSSaddressed to Cytology Cells Detection", IEEE, **(2003)**.
[19] V. V. Raad, "Design of Gabor Wavelets for Analysis of Texture Features in Cervical Images", Proceedings of the 25 Annual International Conference of the IEEE EMBS Cancun, Mexico, **(2003)** September 17-21.
[20] E. Njoroge, S. R. Alty, M. R. Gani and M. Alkatib, "MRCOG Classification of Cervical Cancer Cells using FTIR Data", Proceedings of the 28th IEEE EMBS Annual International Conference New York City, USA, **(2006)**, Aug 30-Sept 3.
[21] N. A. M. Isa, N. Mustafa, K. Z. Zamli, M. Y. Mashor, "Improvement of Contrast Enhancement Technique for Cervical Cell of Pap Smear Images by Reducing the Effect of Unwanted BackgroundInformation", (**2006)** IEEE.
[22] N. Mustafa, N. A. M. Isa, M. Y. Mashor and N. H. Othman, "Colour Contrast Enhancement on Preselected Cervical Cell for ThinPrep® Images", **(2007)**.
[23] C. W. Chang, M. Y. Lin, H. J. Harn, Y. C. Harn, C. H. Chen, K. H. Tsai and C. H. Hwang, "Automatic Segmentation of Abnormal Cell Nuclei from Microscopic Image Analysis for Cervical Cancer

Screening", Proceedings of the 2009 IEEE 3rd International Conference on Nano/Molecular Medicine and Engineering, Tainan, Taiwan, October 18-21, **(2009)**.
[24] S. N. Sulaiman, N. A. Mat Isa and I. A. Yusoff, "Overlapping Cells Separation Method for Cervical Cell Images", IEEE, **(2010)**.
[25] M. E. Plissiti, C. Nikou and A. Charchanti, "Watershed-based segmentation of cell nuclei boundaries in Pap smear images", **(2010)** IEEE.
[26] P. P. Pereira, J. Eliasson, R. Kyusakov, J. Delsing, A Raayatinezhad and M. Johnsson, "enabling Cpnnectivity for Mobile Internet of Things AZpplications Service Oriented System Engineering (SOSE)", **(2013)**, IEEE 7th International Symposium on year: 2013, pp. 518-526., DOI: 10.1109/SOSE.2013.33.
[27] S. Karnouskos, A. W. Colombo, T. Bangemann, K. Manninen, R. Camp, M. Tilly, P. Stluka, F. Jammes, J. Delsing and J. Eliasson, "A SOA-based architecture for empowering future collaborative cloud based industrial automation," in 38th Annual Conference of the IEEE Industrial Electronics Society(IECON2012), Montr´eal, Canada, **(2012)**.
[28] L. Koss P. Lisboa, E. Ifeachor and P. Szczepaniak, "The Application of PAPNET to Diagnostic Cytology".Artificial Neural Networks in Biomedicine, Springer, **(2000)**, pp. 51–68.
[29] A. Kale, "Segmentation And Classification Of Cervical Cell Images", Post Doctoral Thesis, **(2010)**.
[30] B. W. Matthews, "Comparison of the predicted and observed secondary structure of T4 phage lysozyme". Biochemica et Biophysica Acta, vol. 405, **(1975)**, pp. 442–551.
[31] J. Byriel, "Neuro-Fuzzy Classification of cells in Cervical Smears", **(1999)** Febuary.
[32] D. Landwehr, "Web based pap-smear classification", Dept. of Automation, Bldg 326, 2800 Lyngby, Denmark, **(2006)**.
[33] E. Martin, "Pap Smear Classification" a detailed study **(2003)** September, pp 3-17.
[34] R. C. Gonzalez and R. E. Woods, "Digital image processing", Pearson Education Inc Publishing Company, **(2009)**.
[35] E. S. Cibas and S. Z. Ali, "The Bethesda System for Reporting Thyroid Cytopathology", American Journal of Clinical Pathology, vol. 132, **(2009)**, pp. 658-665.
[36] Sited at http://ladiesfirstproviders.vermont.gov/cervical-care/pap-smear-classification-system- bethesda-2001.
[37] D. Solomon, D. Davey and R. Kurman, "The 2001 Bethesda System: Terminology for Reporting Results of Cervical Cytology", 002; vol. 287, no. 16, pp. 2114-2119.
[38] Sited at http://www.sh.lsuhsc.edu/fammed/OutpatientManual/PapSmear.htm.
[39] C. Cortes and V. Vapnik, "Support-vector networks", Machine Learning, vol. 20, no. 3, **(1995)**, pp. 273.
[40] L.Auria and R. A. Moro, "Support Vector Machines as a techinique for Solvency Analysis".
[41] A. Kaushik, G. Singh and A. Bhatia, "SVM Classification in Multiclass Letter Recognition System", Global Journal of Computer Science and Technology Software & Data Engineering, Suresh Gyan Vihar University, India, Global Journals Inc. (USA) Online ISSN: 0975-4172 & Print ISSN: 0975-4350.
[42] J. P. Lewis, A Short SVM (Support Vector Machine) Tutorial. CGIT Lab / IMSC, University of Southern California.
[43] A. J. Smola and B. Scholkopf, "A Tutorial on Support Vector Regression. NeuroCOLT2 Technical Report Series", **(1998)**.
[44] S. K. Shevade, "Improvements to SMO Algorithm for SVM Regression", **(1999)**.
[45] G. W. Flake and S. Lawrence, "Efficient SVM Regression Training with SMO".
[46] V. Vapnik, "The Nature of Statistical Learning Theory", Springer, New York, **(1995)**.
[47] A. Tareef, Y. Song, W. Cai and D. D. Feng, "Automated Three-Stage Nucleus and Cytoplasm Segmentation of Overlapping Cells", 13th International Conference on Control, Automation, Robotics and Vision (ICARCV 2014), (**2014)**.
[48] C. R. Souza, "Kernel Functions for Machine Learning Applications", Web., **(2010)** March 17.
[49] Wikipedia contributors, "Sequential Minimal Optimization", Wikipedia, The Free Encyclopedia, http://en.wikipedia.org/wiki/Sequential_Minimal_Optimization (accessed April 24, **2015**).
[50] Wikipedia contributors, "Support Vector Machine", Wikipedia, The Free Encyclopedia, http://en.wikipedia.org/wiki/Support_vector_machine (accessed April 24, **2010**).
[51] S. Theodoridis and K. Koutroumbas, "Pattern Recognition 2nd ed.", San Diego, C.A, USA;Academic, **(2003)**.
[52] N. Christianni and J. S. Taylor, "An Introduction to Support Vector Machines and Other kernel-based Methods". Cambridge, U.K.:Cambridge University Press, **(2000)**.
[53] S. Lu, C. Zhu and C. Jiao, "Density Weighted Core Support Vector Machine", Advances in Computer Science: an International Journal, **(2015)**, vol. 4, no. 6, pp. 150-155.
[54] J. C. Platt, "Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines", Microsoft Research, **(1998)**.
[55] A. Asuncion and D. J. Newman, "UCI Machine Learning Repository". Irvine, CA: University of California, School of Information and Computer Science **(2007)**.
[56] A. Kirillov, "The AForge.NET Framework". The AForge.NET Computer Vision, Artificial Intelligence and Robotics Website, (**2010)**.

[57] J. Zhang, Y. Liu and T. Zhao, "SVM Based Feature Screening Applied To Hierarchical Cervical Cancer Detection", The Robotics Institute, Carnegie Mellon University Pittsburgh, PA 15213
[58] M. O. Monteagudo, C. Mihai, H. Sahli and A. T. Crispi, "Combined Hierarchical Watershed Segmentation and SVM Classification for Pap Smear Cell Nucleus Extraction", Computación y Sistemas, ISSN 1405-5546, vol. 16, no.2, **(2012)**, pp. 133-145.
[59] J. Zhang and Y. Liu, "Cervical Cancer Detection Using SVM Based Feature Screening" Springer-Verlag Berlin Heidelberg MICCAI (**2004**), LNCS 3217, pp. 873–880
[60] T. Fawcett, , "An Introduction to ROC Analysis", Pattern Recognition Letters, vol. 27, no. 8, **(2006)**, pp. 861 – 874.
[61] D. M. W. Powers, "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation", (PDF). Journal of Machine Learning Technologies, no. 1, **(2011)**, pp. 37–63.

Authors

Sakthi Asokan, received B.E., M. Tech., degrees in Electronics and Communication Engineering from Anna University, and S.R.M. Institute of Research and Technology, Chennai. She has been working as a Lecturer and Associate Professor with more than 7 Years of experience. Her special interest's areas are medical image analysis, Robotics, Dot Net Programming and automation. Her current interest includes in creating impactful contribution in the field of medical image analysis specifically on cervical cancer detection and also spans around Machine Learning, Robotics, and Biomedical, Internet of Things, and Cloud Computing.

M. Rajaram, received his B.E. degree in Electrical and Electronics Engineering in 1981 and M.E. degree in Power Systems in 1988. Besides having a strong technical expertise and analytical skills, he received his Ph.D. degree in 1994. He has contributed to the areas of Computer Networks, High Voltage Engineering, Measurement and Instrumentation, Adaptive Controller, Electro Magnetic Theory and Distributed Computing. He has 157 publications in renowned research journals, 111 research publications in International Conferences, 73 research publications in National Conferences, more than 100 technical reports and six technical books some of which he has co-authored. Currently, he is the Vice-Chancellor of Anna University, Chennai. South India.