Classification of Cervical Cancer from MRI Images using Multiclass SVM Classifier

Bethanney Janney.J *, Umashankar.G, Sindu Divakaran, Shelcy Mary Jo and Nancy Basilica.S

Department of Biomedical Engineering, School of Bio and Chemical Engineering,

Sathyabama Institute of Science and Technology, Chennai, Tamilnadu, India.

*Email: jannydoll@gmail.com

Abstract

Cervical Cancer is the abnormal growth of tissues in the lower, narrow part of the uterus (womb) called the Cervix which connects the main body of the uterus, to the vagina or birth canal. Cervical cancer is one of the most common types of cancer that can be seen in women worldwide. Early detection and proper diagnosis can prevent the severity level and reduce the death rates .In this paper, we have proposed an automated diagnosis system of cervical cancer using texture features and Multiclass SVM (Support Vector Machine) Classifier in MRI images. Initially the MRI images are pre-processed to remove undesirable noises and other effects. After pre-processing, the image is segmented by Region growing method to obtain the region of interest. Texture features are extracted from the segmented region. Almost 22 features are extracted at the region of a segmented area and then passed on to Multiclass SVM Classifier to detect if the given image is cancerous or not. The results of the proposed techniques provide effective results for classifying cancerous and the non-cancerous image.

Keywords: *Cervical Cancer; MRI Images; Multiclass SVM Classifier; Region growing segmentation; Texture features.*

1. Introduction

Cancer is the abnormal development of cells that have the ability to invade or spread to other parts of the body. Women worldwide are affected by Cervical Cancer which is the most common cancer seen especially in developing countries. Timely and precise diagnosis of cervical cancer is an important real world medical problem. Early detection is useful for cervical and colorectal cancer. Cancer treatment includes some combination of radiation therapy, surgery, chemotherapy, and targeted therapy. Important part of care are pain and symptom management. Palliative care is particularly important in people with emerging illness. The survival rate depends on the type of cancer and extent of disease at the start of treatment. During diagnosis in children under 15, the five-year survival rate in the developed world is on average 80%. Cervical cancer is due to the abnormal development of cells in the cervix that have the ability to invade or spread to other parts of the body. Early on, typically no symptoms are seen. Later some of the symptoms like pelvic pain, or pain during sexual intercourse, abnormal vaginal bleeding may be seen. While bleeding after sex may also indicate the presence of cervical cancer [1].

The main cause for cervical cancer is Human papillomavirus (HPV) infection which appears to be involved in the development of more than 90% of cases; however, most people who have had HPV infections do not develop cervical cancer. Other risk factors include smoking, a vulnerable immune system, birth control pills and having many sexual companions, but these are less important. Diagnosis is typically by cervical screening followed by a biopsy. Medical imaging is then done to determine whether or not the cancer has spread. The main aim of this paper is to propose a method for the detection of cervical cancer using contextual clustering along with region growing algorithm. The proposed work uses clustering along with the combination of region growing algorithm for segmenting MRI cervical images. The image is then processed using morphological image processing tools to detect the accurate dimensions of the cancer and classify by using Multi-class Support Vector Machine (SVM). Region growing combined with contextual clustering in this work reduces the number of steps of segmentation. Further use of classifier ensures reliability of detection [2].

2. Related Works

Wen wu et al. [3] stated that two improved SVM methods, Support vector machine-recursive feature elimination (SVM-RFE) and support vector-machine component analysis (SVM-PCA) can be used to diagnose cervical cancer and it is shown that SVM-PCA is superior to others.

Athinarayanan.S et al. [4] combined individual feature extraction method with classification techniques. They used four feature extraction methods namely Gray level co-occurrence matrix (GLCM) and text on co-occurrence matrix (ETCM) and other was combining individual feature extraction methods of Concatenated Feature Extraction (CFE). It is shown that the proposed Concatenated Feature Extraction (CFE) method with SVM Classifier had given the better results than all other remained CFE method with

classifier combinations and all other individual feature extraction and classification combinations.

Abid Sarwar et al. [5] created a rich machine learning database of quantitatively profiled and calibrated cervical cancer cells obtained from pap-smear test slides. In addition to preprocessing and classification, they have also applied a novel hybrid ensemble system to the database in order to evaluate the effectiveness of both novel database and novel hybrid ensemble technique to screen cervical cancer by categorization of Pap smear data. The paper also presents a comparative analysis for prognosis of cervical cancer using multiple artificial intelligence based classification algorithms.

Pradeep Chandran et al. [6] proposed a paper to find-out the treatment output of cervical cancer patients that can be obtained from the parameters of pre-chemo radiotherapy magnetic resonance images. In this paper, segmentation method, spatial fuzzy clustering algorithm is used for segmenting Magnetic Resonance images to detect the Cancer in its early stages anatomical structures. The neural network is used to classify the stage of Tumor that is normal or abnormal. DTCWT texture features, fuzzy clustering segmentation is used as a explanatory variables for PNN classification. The segmentation results will be used as a base for a Computer Aided Diagnosis (CAD) system for early detection of Tumor. Decision making was performed in two stages: feature extraction using GLCM and the classification using PNN-RBF network. It is shown that classifier and segmentation algorithm provides better accuracy than other method.

Soumya.M.K et al. [7] proposed a paper on classification technique using Magnetic Resonance Images (MRI) to obtain the staging of cervical cancer patients. The construction of Nonlinear SVM classification models were based on both second-order texture features and transform features of the tumors. Energy based Transform features such as contour let and Gabor features are used for the prediction of output. However, contrast, correlation, energy and homogeneity which are based on second-order statistical features are significantly used for prediction of outcome from pretreatment MR images of cervical cancer tumors.

3. Proposed Methodology

The image dataset used in this work is a collection of cervical MRI scan images obtained from the cervical Cancer Dataset, hospitals and from different web sources. The images both healthy and malignant are acquired and stored as database. The different components involved in the proposed system for detection of cervical cancer from MRI images are depicted in Fig.1. Supervised contextual clustering is used in the proposed method and along with this the combination of region growing algorithm is used for segmentation of cervical MRI images. Then the images are processed using morphological image processing tools to detect the accurate dimensions of the cancer image provided and it is further classified by using Multi-class Support Vector Machine (SVM).A region growing approach along with the clustering is used to fix the threshold in order to segment the region of interest present in the MRI cervical images. The initial seed point is a voxel belonging to the cervix region, and the fuzzy rule fixes a value by selecting the voxels with intensity values lower than the given threshold. The initial seed point is automatically chosen by selecting the pixels which is present in the central slice of the MRI image and grows towards the entire cervix region present in the image. Finally morphological operations are applied and given to multiclass SVM classifier.
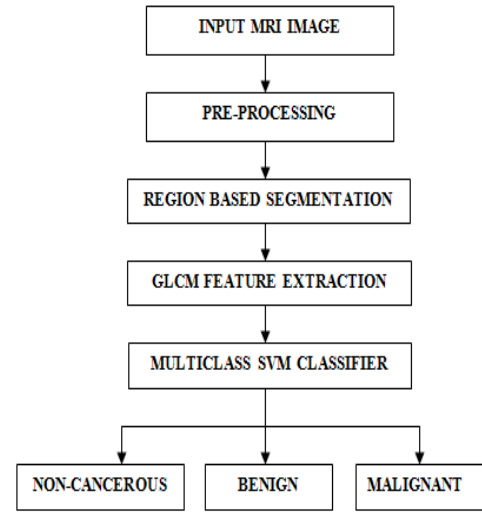

Fig. 1: Block diagram of Proposed Architecture

3.1. Pre-processing

It is a preliminary step that is used for noise removal, contrast enhancement and illumination equalization. Non-local means is an algorithm for image denoising. Unlike "local mean" filters, which take the mean value of a set of pixels surrounding a goal pixel to smooth the image, non-local means filtering takes an average of all pixels in the image, weighted by using how similar those pixels are to the goal pixel. This result in plenty extra post-filtering readability, and less loss of element in the image in comparison with local mean algorithms [8], [9].

3.2. Region Based Contextual Clustering Segmentation

In region growing based contextual clustering technique, a region growing technique along with the clustering is used to restore the threshold as a way to segment the region of interest present in the MRI cervical images. The initial seed point is a voxel (3x3 pixels) belonging to the cervix region, and the fuzzy rule fixes a value with the aid of choosing the voxels with depth values lower than the given threshold. The preliminary seed point is automatically selected via deciding on the 3x3 pixel that is present inside the central slice of the MRI image and grows toward the complete cervix vicinity present in the image [10]. In a clustering technique along with the region growing, each pixel is related to one of the finite wide variety of threshold is grown to shape disjoint areas. The contextual clustering technique is based on supervised algorithm. It utilizes 3 X 3 covering windows of pixels to form a segmented image. The nature of segmented image depends upon a defined threshold value (T=120) by the client which is utilized to pick the closest areas for segmentation and a controlling parameter β which is in the scope of 0< β <1.The median estimation of the 9 pixels in the window, the aggregate number of intensity esteems (u) >1 inside the window, barring the officially distinguished median esteem. The relevant bunching portions an information into classification 1 (ω0) and classification 2 (ω1) in view of the grown region. The contextual value is calculated using equation (1).

$$Vcc = \text{Median value} + \frac{\beta}{\text{threshold}} \times \left\{ u - \frac{\text{window size}}{2} \right\} \qquad (1)$$

The value estimated is contrasted and a set threshold limit which is settled so as to pick the pixels for region growing. In the event that the estimated contextual value is not exactly or equivalent to the set edge, at that point 0 is relegated to the focal point of window else, 1 is appointed to the focal point of the window. The presence of 1 in a region in the image demonstrates the segmented region of interest**.** Consequently the desired feature is extracted.

3.3. GLCM Feature Extraction

The basic role of feature extraction is to separate beneficial characteristics from the image. A feature or descriptor is depicted as a work of 1 or extra estimations determining some measuring properties like color, texture or shape etc of the whole image or sub-image or one object is outlined as a characteristic or descriptor. During this work, we attention on extraction of features which might be one of a kind and standardize them. This is carried out using Gray level co-occurrence matrix (GLCM).it's miles a well-set-up statistical tool that extracts second order texture data from images. In a GLCM matrix, the quantity of rows and columns is adequate to the quantity of distinct grey levels or pel values within the image of that surface and it describes the frequency of 1 grey level showing in an exceedingly mere spatial linear relationship with another grey level among the realm of investigation. In a picture of mere intensity levels, the chance of various mixtures of grey levels co-occurring in a picture or image section is tabulated by GLCM [11]. A measure of the variation in intensity at the pel of interest is given by Texture feature calculations victimization the contents of the GLCM. regularly parameters i.e. relative distance between the pixel pair d this is measured in pixel range and their relative orientation θ. commonly is quantized in 4 guidelines 0º, 45 °, 90 ° and 135 °, diverse different combinations are also feasible[12][13]. The features extracted from the image are listed in Table1.

Table 1: Summary of image features extracted

| Feature	Feature Description	Feature	Feature Description
F1	Uniformity / Energy	F12	Sum of Squares
F2	Entropy	F13	Sum Average
F3	Dissimilarity	F14	Sum Variance
F4	Contrast / Inertia	F15	Sum Entropy
F5	Inverse difference	F16	Difference variance
F6	Correlation	F17	Difference entropy
F7	Homogeneity	F18	Information measures of correlation (1)
F8	Autocorrelation	F19	Information measures of correlation (2)
F9	Cluster Shade	F20	Maximal correlation coefficient
F10	Cluster Prominence	F21	Inverse difference normalized (INN)
F11	Maximum probability	F22	Inverse difference moment normalized

3.4. A Multiclass SVM Classifier

The images are further classified using Multi-class Support vector machines (Multi-class SVMs). Support vector networks are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. Given a set of training examples, each marked as belonging to one or the other of two categories, SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier. A Multiclass SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall. A multiclass SVM classi-

fier is utilized for better classification of cervical MRI images of type Non-Cancerous, Benign, and Malignant [14], [15].

4. Results and Discussion

The experimental results of the classification method using MRI cervical images with different types of cervical cancer are described in this section. The MRI image database consists of normal and abnormal cervical images. These images are collected from The Cancer Imaging Archive (TCIA). The dataset in this proposed work consists of three classes of normal, benign and malignant based 180 cervical MRI images. From these 90 images were used for training phase and the remaining 90 images used for the testing phase. It is observed that the outcome acquired because of the utilization of previously mentioned strategies is considerably more precise than alternate techniques.
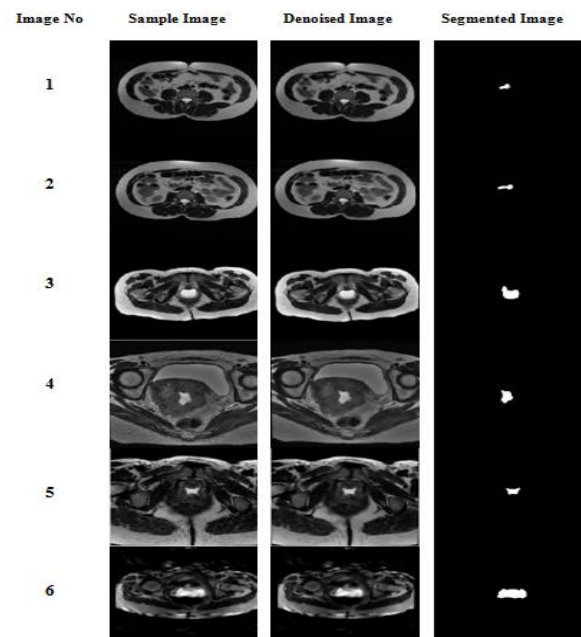
Fig. 2: Segmentation results of the images

The input MRI cervical cancer images are pre-processed using Non-local means algorithm which provides much greater post-filtering clarity, and less loss of detail in the image. After pre-processing, the region growing based contextual clustering algorithm is used for segmentation followed by 22 features are extracted from the segmented area and further classified by multiclass SVM classifier which accurately identified 85 images as malignant, 55 images as benign and 40 images as non-cancerous. The segmented result of 6 sample images is shown in Fig. 2.
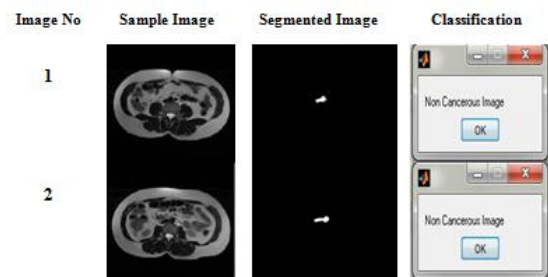
Fig. 3: Classification of Non-cancerous Images

The segmented images and classification results of Non-cancerous and Benign cervical images are displayed in Fig. 3, and Fig. 4.
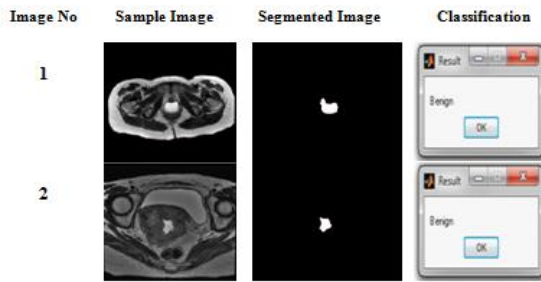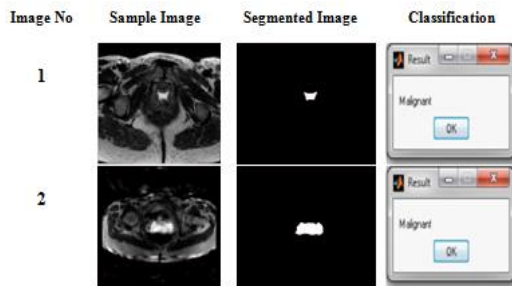
Fig. 4: Classification of Benign Images

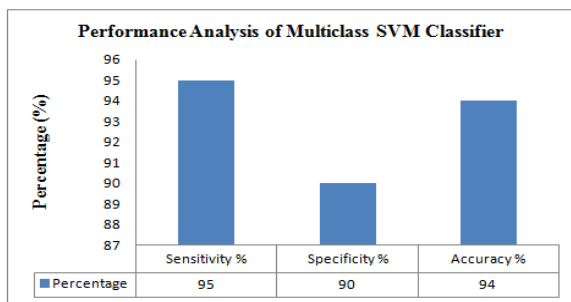
Fig. 5: Classification of Malignant Images

Fig. 6: Performane Analysis of Multiclass SVM Classifier

The segmented images and classification results of malignant cervical images are shown in Fig. 5. To validate the performance of classifier three different metrics have been selected, the measures are like accuracy, specificity and sensitivity [16]. The performance analysis has been made by the following equation.

$$Accuracy = (TP + TN) / (TP + TN + FP + FN)$$

$$Sensitivity = (TP) / (TP + FN)$$

$$Specificity = (TN) / (TN + FP)$$

Where TP = True positive, FP = false positive, FN = false negative and TN = True negative. The execution of a classifier could be assessed as far as the quantity of TP and FPs.

Table 2: Performance results for multiclass SVM classifier

Input Image Data set	Evaluation Metrics						
	TP	TN	FP	FN	Sensitivity %	Specificity %	Accuracy %
	76	18	2	4	95	90	94

The Table2. demonstrates the performance of classification technique. This multiclass SVM classification technique gives effective outcome and the evaluation of the table clearly indicates the superior performance of the proposed classifier through achieving 94% of accuracy, 95% of sensitivity and 90% of specificity, which might be a good deal higher than those of the prevailing methods. It's far very clean from the above discussions that the proposed

approach plays higher in view of classifying the cervical MRI images into Non-cancerous or malignant or benign. The Fig. 6 shows that proposed system classification technique of cervical MRI images compared with specificity, sensitivity and accuracy.

5. Conclusion

Over the past several years, the role of MRI in gynecology oncology has evolved. Owing to its superior soft tissue delination, MRI is now the most widely used imaging modality in the initial staging of primary cervical cancer. In this paper, it is proposed by using GLCM features and Multi-class SVM classifier, the MR Images can be differentiated in a more accurate manner. The proposed method extracts region of interest from MR Images using Region growing method. The texture features are extracted using Grey Level Co-Occurrence Matrix (GLCM). The experiment is performed on standard datasets to evaluate the performance of the proposed model and to reach the specific goal of classifying cervical cancer as Non-Cancerous, benign and malignant. Hence it is able to produce results in a more precise manner efficiently.

Acknowledgement

References

[1] Aswathy.S, M. A. Quereshi, B. Kurian, and K. Leelamoni, "Cervical cancer screening: current knowledge & practice among women in a rural population of Kerala, India", *Indian Journal of Medical Research*, Vol. 136, No. 2, (2012), pp. 205–210.

[2] Yessi Jusman, Siew Cheok Ng, and Noor Azuan Abu Osman," Intelligent Screening Systems for Cervical Cancer", *The Scientific World Journal*, Vol. 2014, (2014), pp. 1-15.

[3] Wen Wu and Hao Zhou,"Data-Driven Diagnosis of Cervical Cancer With Support Vector Machine-Based Approaches", *IEEE Access*-Data-Driven Monitoring, Fault Diagnosis and Control of Cyber-Physical Systems ,Vol. 5, (2017), pp. 25189 - 25195.

[4] Athinarayanan.S, M.V.Srinath,"Automatic Detection and Classification of Cervical Cancer in Pap Smear Images using ETCM & CFE methods based Texture features and Various Classification Techniques", *International Journal of Engineering and Computer Science*, Vol.5, No.7 ,(2016), pp. 17368-17387.

[5] Abid Sarwar,"Artificial Intelligence Based Semi-automated Screening of Cervical Cancer using a Primary Training database", *Journal of HPV and Cervical Cancer*, Vol.1, No.1, (2016), pp.1-10, 2016.

[6] Pradeep Chandran.K, Smt.U.V.Ratna Kumari, "Improving Cervical Cancer Classification On MR images Using Texture Analysis And Probabilistic Neural Network", *International Journal of Science, Engineering and Technology Research (IJSETR)*,Vol.4, No.9, (2015), pp. 3141-3145.

[7] Soumya.M.K, K. Sneha and C. Arunvinodh,"Cervical Cancer Detection and Classification Using Texture Analysis", *Biomedical & Pharmacology Journal*, Vol. 9, No.2, (2016), pp. 663-671.

[8] Wang X, Wang H, Yang J, Zhang Y, "A New Method for Nonlocal Means Image Denoising Using Multiple Images", *PLoS ONE*, Vol. 11, No.7, (2016), http:// doi.org/10.1371/journal.pone.0158664.

[9] He N, Lu K,"Images denoising by improved Non-local means algorithm", *Communications in Computer and Information Science, Springer,* Vol.164, (2011), pp.33-39.

[10] Prabin, A. and Veerappan. J, "Automatic segmentation of lung ct images by cc based region growing", *Journal of Theoretical & Applied Information Technology*, Vol. 68, No.1,(2014), pp. 63-69.

[11] Bethanney Janney J and Emalda Roslin S, "Classification and Detection of Skin Cancer using Hybrid Texture Features, *Biomedicine*,Vol. 37, No.2, (2017), pp.214-22.

[12] Dhumal.S, Agrawal.S,"MRI Classification and Segmentation of Cervical Cancer to Find the Area of Tumor", *International Journal for Research in Applied Science & Engineering Technology (IJRASET),* Vol. 3, No.7 ,(2015), pp. 21-26.

[13] Edwin Jayasingh Mariarputham and Allwin Stephen, "Nominated Texture Based Cervical Cancer Classification", *Computational and Mathematical Methods in Medicine*, Vol. 2015,(2015) pp. 1-10.

[14] Sukumar.P, and R. K. Gnanamurthy, "Computer Aided Detection of Cervical Cancer Using Pap Smear Images Based on Adaptive Neuro Fuzzy Inference System Classifier", *Journal Of Medical Imaging And Health Informatics*, Vol. 6, No. 2, (2016), pp.312–319.

[15] Turid Torheim, Eirik Malinen, Knut Kvaal, Heidi Lyng, Ulf G. Indahl, Erlend K. F. Andersen, and Cecilia M. Futsæther,"Classification of Dynamic Contrast Enhanced MR Images of Cervical Cancers Using Texture Analysis and Support Vector Machines", *IEEE Transactions on Medical Imaging*, Vol. 33, No.8, (2014), pp.1648-1656.

[16] Ebtihal Almansour and Arfan Jaffar M, "Classification of Dermoscopic Skin Cancer Images Using Color and Hybrid Texture", *.International Journal Compuert Science Network Secuityr*, Vol.16, No.4, (2016), pp.135-139.