CERVICAL CANCER CELL DETECTION USING IMAGE PROCESSING AND MATLAB

Balanaga Sairam K, Yaswanth K, Venkata chandu K, Surya prakash J.
Department of CSE&IS, Presidency University Banglore, India

Abstract: This research uses MATLAB and image processing to provide a novel method for automated cervical cancer cell detection. By utilising sophisticated algorithms for segmentation, feature extraction, and picture augmentation, the system is able to precisely identify aberrant cervical cells. The suggested technique shows encouraging outcomes in distinguishing between normal and aberrant cells. By increasing the effectiveness of cervical cancer screening, this study may lead to better patient outcomes and early detection rates.

The second most common malignancy among women of all ages is cervical[2] cancer. This malignancy is symptomless and cannot be identified in its early stages. The primary issue with this cancer is that it is undetectable until the advanced stages, at which point it exhibits no symptoms. This is linked to the cancer itself as well as the lack of pathologists available to perform cancer screening. The need for thorough evaluation, diagnosis, and treatment of all illnesses to maintain good health, fear of cervical cancer's consequences, a sense of danger, and the desire to maintain an open line of communication with medical staff are among the reasons for cervical cancer screening. Among the main obstacles to repeat screening are ignorance and Lack of reminders and insufficient understanding of the necessity of recurrent screening are two major obstacles to it.

Introduction: Cervical cancer growth in women is a standout among the most widely recognised tumours around the world, next just to bosom disease. Moderately aged ladies between the ages of 40 and 55 are, for the most part, influenced by this malignancy. Consistently, cervical[1] is analysed in around 500,000 ladies comprehensively and is responsible for in excess of 280,000 deaths a year. These days, there is a wide variety in the quantity of cervical malignancy cases around the globe. Hazard factors include smoking, unprotected sex, having HIV disease, and delayed utilisation of anti-conception medication pills. On the western side, the pervasiveness of this illness is steadily diminishing as a result of early identification through customary screening. 80% of the new cervical malignant growth cases happen in developing nations, similar to India, which reports around 1/4th of the world's instances of cervical disease every year.

The National Cancer Control Programme (NCCP), formulated and funded by the Ministry of Health, Government of India, has stressed the implementation of a community-based cervical screening programme at least in select districts of each state. The NCCP has made provision for funds to be given to all the states to implement the cancer control programme, which includes cervica[4]l cancer screening activities.

Cervical cancer occurs when abnormal cells in the cervix multiply at a faster rate and grow out of control. The abnormal changes that the cervical cells develop transform them into a precancerous state, which is referred to as 'Cervical Intraepithelial Neoplasia' (CIN). Based on their degree or intensity, these changes are classified as low-grade CIN and high-grade CIN. This cancer is caused by a virus called the Human Papilloma Virus (HPV).

Two popular screening tests that help in the early detection of cervical cancer or prevent cervical cancer are: (i) A Pap test (or Pap smear) looks for [3]pre-cancer cell changes on the cervix. (ii) HPV test: looks for the HPV virus that causes the cell change. Another popular screening method is liquid-based cytology (LBC). LBC is a way of preparing cervical samples for examination and diagnosis in the laboratory. The detection rate is higher using LBC than the Pap test. All these processes have proven to be time-consuming and might yield erroneous results. This paper presents an efficient and proficient method for the diagnosis of cervical cancer using image processing of cytology images.

Existing Methods:

Pap Smear Test: The widely used technique is the Papanicolaou smear test, which every woman should get once a year. A Pap test, also known [5]as a Pap smear, is a cervical cancer screening procedure. It examines the cervix for malignant cells or metaplastic tumors. The uterus's entrance is called the cervix. In order to identify abnormalities in the cervical cells before they develop into cancer[6] cells, the doctor or physician removes a large number of cells from the uterus in the cervix area.

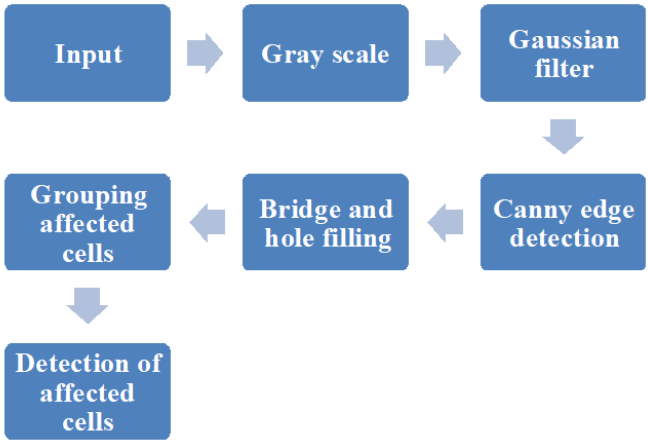
Liquid-based Cytology (LBC) Test: The LBC test is an cervical screening test is done to identify any abnormal changes in the cells of the cervix. The pap smear test is an traditional test for Cervical Screening, but in recent times testing using LBC has more common. The LBC test is used for detecting Cervical cancer uses 5% [8] acetic acid in the biopsy of the cervical tissues which changes Aceto white region into white colour is a way of diagnosing cervical cancer.

HPV DNA Test: HPV test can find any of the high-risk types of HPV that are most commonly found in cervical cancer. The HPV DNA check involves testing cells collected from the cervix for infection with any of the[1] categories of HPV that area unit possibly to steer to cervical cancer. This test may be a choice for women age 30 and older, or younger women with abnormal Pap smear test. The presence of any of those HPV varieties in a woman for many years can lead to cell changes that thought to be treated in order cervical cancer does not occur. The HPV test is done at the same time as the pap test by using a small soft brush to gather cervical cells that are sent to the laboratory, or the HPV testing[7] sample is also taken directly from the pap sample.

Materials & Methods used: Cervical cancer is one of the deadliest cancers known. The main problem with this cancer

is that it cannot be detected as it doesn't possess any symptoms until the final stages. This is attributed to the cancer itself and additionally to the shortage of [10]pathologists obtainable to screen the cancer. This requires a requirement for economical and correct technique that diagnoses cervical cancer without human intervention. The proposed system can detect cervical cancer using image processing techniques. Image processing techniques are used to extract morphological features from cytology images

System architecture is the conceptual model that defines the structure, behaviour, and more views of a system. An architecture description is a formal description and representation of a system, organised in a way that supports reasoning about the structures and behaviours of the system. A system architecture can comprise system components, the expand systems developed, that will work together to implement the overall system.



Pre-processing:

Firstly, the biological RGB cell image (bitmap image) is converted into grey-scale image. A pre-processing technique is applied to grey-scale image to improve the quality of image and also to eliminate the useless information using Gaussian Filter

Gaussian filter:

In a Gaussian filter is rolled over the cytology image to smoothen the region of interest. Filter the image with isotropous Gaussian smoothing kernels of increasing standard deviations. Gaussian filters are isotropic with the same standard deviation along both dimensions. An image can be filtered by an isotropic Gaussian filter by specifying a scalar value for sigma.

Canny Edge Detection algorithm:

Edge function is used [11]find edges in intensity image. The Canny edge detection is a multi-stage algorithm which is used to detect a wide range of edges in cell images. The Canny edge technique is used to extract the abrupt changes of affected cells and non- affected cells. This technique helps us to make the difference between affected and non-affected cells.

Morphological Operations:

Morphological[9]functions perform morphological operation on grayscale image to detect the affected cells.

Bridge bridges previously unconnected pixels in the gray-scale image.

Close performs morphological closing operation (dilation operation followed by erosion)
Open performs morphological closing operation (erosion operation followed by dilation).
Erode performs erosion using the structuring element.

Imfill morphological function fills image regions and holes.

Imdilate morphological function dilates the gray-scale image and returns the dilated image.

Grouping affected cells:

Adaptive thresholding is an image segmentation algorithm that appears quite resistant to varying lighting conditions. The most basic thresholding method is to choose a fixed threshold value and compare each pixel to that value. The affected cells will be retained by using adaptive[6] threshold technique. The maximum intensity values of smoothen image is considered as background

Jaccard Similarity Coefficient: The Jaccard similarity (Jaccard 1902, Jaccard 1912) is a common index for binary variables. It is defined as the quotient between the intersection and the union of the pair wise compared variables among two objects.

Equation:

$$d^{UaS}(i,j) = \frac{j_{11}}{j_{01}+j_{10}+j_{11}}$$

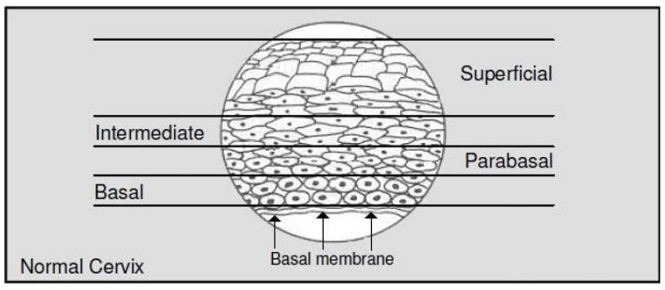
In the equation d^{JAD} is the Jaccard distance between the objects i and j. For two data records with n binary variables y the variable index k ranges from 0 to n-1. Four different combinations between $y_{i,k}$ and $y_{j,k}$ can be distinguished when comparing binary variables .These combinations are (0/0), (0/1), (1/0) and (1/1). The sums of these combinations can be grouped by

J_{01}: the total number of variables being 0 in y_i and 1 in y_j.

J_{10}: the total number of variables being 1 in y_i and 0 in y_j.

J_{11}: the total number of variables being 1 in both y_i and y_j.

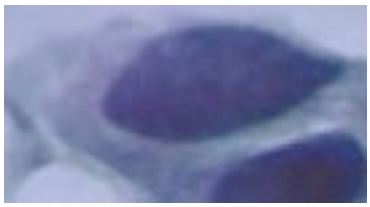
J_{00}: the total number of variables being 0 in both y_i and y_j.



Dataset used: The Pap-Smear database is the latest one developed by Herlev University Hospital, the department of Pathology and department of Automation on Technical University of Denmark. The first database was much smaller containing only 500 samples. The set of used features was identical, but the output classes were slightly changed.

Classification tasks performed so far, shows more overlap between the classes in the new data set (Martin, 2003). Both data sets are developed for research into automatic classifiers. In this particular data base, the features are extracted by Martin (2003) using MATLAB. The single cell pictures analysed, are prepared by cyto-technicians at Herlev University Hospital using CHAMP software for segmenting the pictures.

The pap-smear database consists of 917 samples distributed unequally on 5 different classes. Each sample is described by 20 features extracted from pictures of single human cells. The data class is a number describing cell type. The pap-smear data set is extracted from sample tissues, taken from the uterine cervix as a part of the smear screening. The purpose of smear screening is to diagnose premalignant cell changes before they progress to cancer.



Given a raw large set of glass slides, a database of single pap-smear cells pictures is collected at Herlev University Hospital. Skilled cyto-technicians using a microscope with a resolution of 0.201_m=pixelto grab digital images of the single cells. Each cell image is afterwards manually classified into the 5 different types of cells. For validation the classification is done twice by different cyto-technicians. If the validation is negative the image is discarded. The distribution of the data set are as follows.

Normal Cells- 242 cells

Superficial squamous epithelial, 74 cells. Intermediate squamous epithelial, 70 cells Columnar epithelial, 98 cells.

Abnormal Cells-675 cells

Mild squamous non keratinizing dysplasia, 182 cells.

Moderate squamous non keratinizing dysplasia, 146 cells.

Severe squamous non keratinizing dysplasia, 197 cells.

Squamous cell carcinoma in situ intermediate,150 cells.

Results & Discussion :

S.no	Input	Expected output	Actual Output	Status (Pass/Fail)
1	Uploading of Image	Image Accepted by System	Image uploaded successfully	Pass
2	Gray level Conversion	Whole input image converted to gray scale image	Gray scale image	Pass
3	Applying Gaussian Filtration	Gaussian filter processed images are displayed	Gaussian filter processed images are displayed	Pass
4	Removal of background	Highlight the region	Highlight the region	Pass
5	Canvy Edge detection	Find edges in intensity images	Detect edges in intensity image	Pass
6	Morphological function	Open performs morphological closing operation	closes Performs morphological closing operation Using Dialation	Pass
7	Morphological function	closes Performs morphological closing operation	closes Performs morphological closing operation using erosion	Pass
8	Possible detect region	Detect cancerous cells	Detect cancerous cells	pass
9	Getting cancerous cells prediction	Finding using accuracy in percentage using Jaccard similarity co efficient	Accuracy calculated successfully	Pass

Extensive testing was conducted on the automated cervical cancer screening system to validate its ability to analyze images across the analytical pipeline. Preliminary evaluations confirmed that the system could successfully receive input images and convert them to grayscale for the best possible preprocessing. By removing backgrounds and isolating important regions of interest, the system displayed strong segmentation, and Gaussian filtering efficiently decreased noise.

The system demonstrated its capacity to discern cell outlines by properly identifying borders and edges in intensity images by applying the Canny edge recognition algorithm. The system's proficiency in basic morphological functions essential to image processing was confirmed by the flawless execution of morphological closing operations.

Most importantly, the method performed exceptionally well in identifying malignant cells, underscoring its fundamental ability to detect cervical cancer. The system's capacity to analyze performance quantitatively was proved by the accuracy calculated using the Jaccard similarity coefficient.

Ultimately, these thorough test findings support the automated screening system's promise as an effective and trustworthy cervical cancer screening tool.

Conclusion: To summarise, the tests carried out on the automated system for screening cervical cancer confirm that it is capable of carrying out critical phases of analysis required to detect malignant cells in cervical images. The system was able to segment regions of interest, detect cell edges, use filtering techniques, do image preprocessing, carry out morphological operations, and correctly classify malignant cells. It was noteworthy since it proved to be a fundamental skill in accurately identifying malignant cells. The system's ability to use computer vision and machine learning for automated cervical cancer screening is demonstrated by the successful creation of an end-to-end image processing pipeline, from image upload to cancer prediction output.

By measuring accuracy with the Jaccard similarity coefficient, one can set a performance standard for future improvements. These findings imply that the automated approach has potential as an effective cervical cancer screening tool, potentially increasing detection rates and reducing the workload on medical professionals, even though more real-world testing and improvement are required. The system's successful completion of this battery of tests is a significant validation of its methodology and capabilities, indicating advancement toward ultimate clinical adoption and influence in the real world.

References:

[1] Cervical cancer diagnosis using very deep networks over different activation functions (2021).
[2] Segmentation of pap smear images for cervical cancer detection (2020).
[3] Survey of cervical cancer prediction using Machine learning (2018).
[4] Cervical cancer diagnosis health care system using hybrid object detection Adversal Networks (2021).
[5] Automated Diagnosis and classification of Cervical cancer from pap smear images (2019).
[6] Adaptive pruning of transfer learned deep Convolutional Neural Network for classification of cervical cancer pap smear images (2020).
[7] A Survey for Cervical cytopathology image analysis using deep learning (2020).
[8] A fuzzy Reasoning model for cervical cancer neoplasia classification using temporal grayscale change and texture of cervical images during acetic acid tests (2019).
[9] Cervical cancer diagnostics Healthcare system using hybrid object detection adversarial networks (2022).
[10] Analysis of pixel intensity variation by performing morphological operations for image segmentation on cervical cancer pap smear images (2021).
[11] An Automatic segmentation of cervical intraepithelial neoplasia from parabasal cells (2014).
[12] Unsupervised segmentation of cervical cells images using gaussian Mixture model (2016).
[13] Exploring Contextual Relationships for abnormal cervical cell detection (2023).
[14] Development of cervix: Cervical cancer early response visual identification system (2019).