Cervical Cancer Cell Detection Using Image Processing and Matlab

VENKATA CHANDU K.[1], BALANAGA SAIRAM K.[2], YASWANTH K.[3], SURYA PRAKASH J.[4], DEVI S.[5]

[1, 2, 3, 4, 5] *Department of CSE&IS, Presidency University Banglore, India*

Abstract-* This work presents a unique automated technique for the identification of cervical cancer cells using image processing and MATLAB. The method accurately detects abnormal cervical cells by employing complex algorithms for segmentation, feature extraction, and image enhancement. The proposed method performs well in differentiating between normal and abnormal cells. This study might improve patient outcomes and early detection rates by making cervical cancer screening more effective. Cervical cancer ranks as the second most prevalent cancer in women of all ages. Because it has no symptoms, this cancer cannot be detected in its early stages. The main problem with this cancer is that it doesn't show any signs until it has progressed to an advanced stage. This is related to both the cancer and the scarcity of pathologists who can do cancer screenings. Cervical cancer screening is recommended for a number of reasons, including dread of the repercussions of cervical cancer, a sense of risk, the need for complete examination, diagnosis, and treatment of all illnesses to preserve good health, and the need to keep communication lines open with medical professionals. Ignorance is one of the biggest barriers to repeat screening. Two significant barriers are a lack of reminders and a poor comprehension of the importance of ongoing screening.

I. INTRODUCTION

Next only to bosom illness, cervical cancer development in women is one of the most well-known Tumor worldwide. Most women affected by this cancer are middleaged, namely between the ages of 40 and 55. Cervical [1] is routinely examined in around 500,000 women in their entirety and is the cause of more than 280,000 deaths annually. The number of cases of cervical cancer that occur worldwide these days varies greatly. Risk factors encompass smoking, engaging in unprotected sexual activity, being HIV positive, and postponing the use of anti-conception medicine. Due to early detection through routine screening, the prevalence of this ailment is gradually declining on the western side. Similar to India, which records around one-fourth of the global cases of cervical illness annually, developing countries account for 80% of newly reported cases of cervical malignant growth.

The Ministry of Health, Government of India, developed and funds the National Cancer Control Initiative (NCCP), which has emphasized the establishment of a community-based cervical screening initiative in at least a few state districts. Funds will be provided to all states by the NCCP to carry out the cancer control plan, which includes cervical cancer screening initiatives.

When aberrant cells in the cervix develop more quickly and become uncontrollably large, it can lead to cervical cancer. "Cervical Intraepithelial Neoplasia" (CIN) is the term used to describe the precancerous condition that results from the aberrant alterations that the cervical cells undergo. These alterations are divided into low-grade and high-grade CIN categories according to their severity or degree. The Human Papilloma Virus (HPV) is the virus that is responsible for this malignancy.

Cervical cancer can be prevented or detected early with the use of two widely used screening tests: I On the cervix, a Pap test, often known as a Pap smear, looks for [3] pre-cancer cell alterations. Examines the HPV virus that is responsible for the cell alterations (ii) HPV test. Liquid-based cytology (LBC) is an additional widely used screening approach. Samples of the cervical mucosa are prepared for laboratory analysis and diagnosis by LBC. Compared with the Pap test, the LBC has a greater detection rate. All of these procedures have been shown to be time-consuming and potentially result in incorrect outcomes. The approach for diagnosing cervical cancer by image processing of cytology pictures is presented in this study as effective and proficient.

II. EXISTING METHODS

- Pap Smear Test: The most common method is the Papanicolaou smear test, which is recommended for every woman once a year. A Pap test, often referred to as a Pap smear [5], is a cervical cancer screening method. It checks for metaplastic tumors and cancerous cells in the cervix. The cervix is the opening to the uterus. The doctor or physician takes a significant number of cells from the uterus in the cervix area in order to find abnormalities in the cervical cells before they develop into cancer [6].

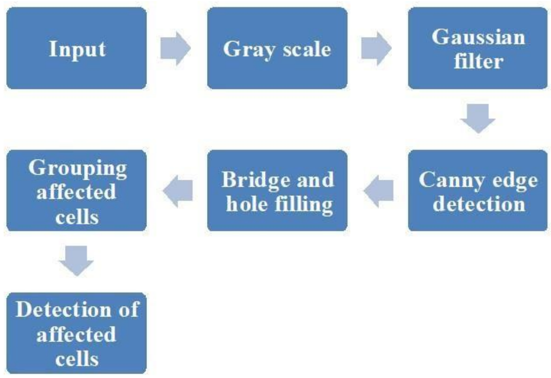
- Liquid-based Cytology (LBC) Test: theLBC test is a cervical screening procedure used to find any abnormal alterations in the cervix's cells. The standard test for cervical screening is the pap smear, although more and more tests these days use LBC. The LBC test, which is used to diagnose cervical cancer, uses 5% [8] acetic acid in the cervical tissue biopsy, which causes the Aceto white zone to become white in color.

- HPV DNA Test: Any of the high-risk HPV varieties that are often associated with cervical cancer can be detected by an HPV test. Cervical cells are tested for HPV infection with any of the [1] HPV groups that may lead to cervical cancer as part of the HPV DNA check procedure. Women who are 30 years of age or older, or younger women whose Pap smear results are abnormal, may choose to get this test. If a woman has been exposed to any of those HPV strains for a long time, she may experience cell alterations that should be treated to prevent cervical cancer.Using a tiny soft brush to collect cervical cells that are submitted to the lab, the HPV test is performed concurrently with the pap test. Alternatively, the HPV testing sample is extracted straight from the pap sample.

III. MATERIALS & METHODS USED

One of the worst tumors known to science is cervical cancer. The primary issue with this cancer is that it is undetectable until the advanced stages, at which point it exhibits no symptoms. This is ascribed to the disease itself as well as the lack of available [10], pathologists for cancer screening. This calls for the need for an accurate and costeffective method that can detect cervical cancer without the need for human involvement. The suggested approach uses image processing methods to identify cervical cancer. Images from cytology are processed using image processing methods to obtain morphological information.

The conceptual model that outlines a system's behaivour, structure, and other aspects is called system architecture. A system's formal description and representation, structured to facilitate inference about the system's behaviours and structures, is called an architectural description. A system architecture may consist of expanded systems that will cooperate to accomplish the system as a whole, or system components.



- Pre-processing: First, a greyscale picture is created from the biological RGB cell bitmap image. Using a Gaussian filter, a pre-processing technique is used to a greyscale image to enhance its quality and remove unnecessary information.

- Gaussian filter: To smooth the region of interest in the cytology picture, a Gaussian filter is applied. Use isotropous Gaussian smoothing kernels with increasing standard deviations to filter the picture. The standard deviation of Gaussian filters is the same in both dimensions, making them isotropic. By providing a scalar value for sigma, an isotropic Gaussian filter may be applied to a picture.

- Canny Edge Detection algorithm: To locate edges in an intensity picture, utilize the edge function [11]. A multi-stage technique called Canny edge detection is used to find a variety of edges in cell pictures. The sudden differences between impacted and non-affected cells are extracted using the Canny edge approach. This method aids in distinguishing between cells that are impacted and those that are not.

- Morphological Operations: Morphology tools in MATLAB allow you to manipulate picture forms. Erosion takes away, dilation adds, and their combination cleans or enlarges things. Pick a brush (a structural element) and use size and neighbouring information to paint (process) the picture. Try out these shapes to see which one best suits your needs for picture editing!

- Grouping affected cells: One picture segmentation approach that seems to be quite resilient to changing illumination conditions is adaptive thresholding. Selecting a predetermined threshold value and comparing each pixel to it is the most fundamental thresholding technique. The adaptive[6] threshold approach will be used to keep the impacted cells. Background is defined as the smoothed image's greatest intensity values.

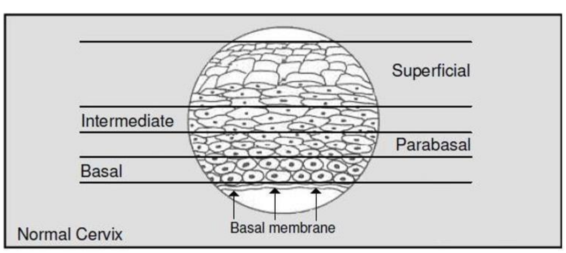
- Jaccard Similarity Coefficient: For binary variables, the Jaccard similarity (Jaccard 1902, Jaccard 1912) is a widely used index. Its definition is the quotient of the union and intersection of the pairwise variables between two objects that are being compared.

- Equation:

$$d^{UaS}(i, j) = \frac{j11}{j01+j10+j11}$$

The Jaccard distance between items I and J is represented by the equation dJAD. The variable index k for a pair of data records containing n binary variables y spans from 0 to n-1. When comparing binary variables, four distinct combinations between yi,k and yj,k may be identified. The combinations in question are 0/0, 0/1, 1/0, and 1/1. These combinations' amounts can be categorized 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 most recent Pap-Smear database was created by the Technical University of Denmark's Automation, Pathology, and Herlev University Hospital departments. There were just 500 samples in the initial database, which was substantially smaller. The output classes were somewhat modified, but the list of characteristics employed remained the same.

- Increased overlap: New datasets show more class mixing in classification tasks, which presents difficulties for automated classification.

The datasets are intended to be used for the particular purpose of testing and improving automatic classification methods. Feature extraction: Martin (2003) used MATLAB to extract important features from this dataset.

- Image preparation: Herlev University Hospital cytotechnicians use CHAMP software to carefully segment single-cell pictures, setting the stage for further investigation.

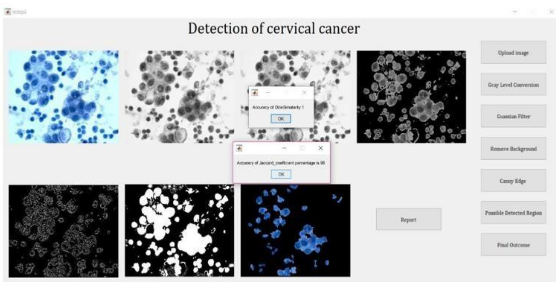
There are 917 samples in the pap-smear database, which are unevenly divided into 5 groups. Twenty traits that were taken from images of single human cells are used to describe each sample. The integer that represents the cell type is the data class. As part of the smear screening process, sample tissues from the uterine cervix are used to retrieve the pap-smear data set. Smear screenings are performed in order to identify premalignant cell alterations before they develop into cancer.

A vast collection of glass slides was used by researchers at Herlev University Hospital to create a database of single pap smear cell pictures. Expert cytotechnicians digitally photographed individual cells using a microscope with a resolution of 0.201 microns per pixel. Each cell picture was manually categorized by the technicians into one of five cell types. Each picture was separately categorized by a second cytotechnician in order to verify the classifications. Any photos that the two technicians could not agree upon how to classify were thrown away. The distribution of cell types in the final database was as follows:

- Normal Cells- 242 cells: Additionally, 74 cells were categorized as superficial squamous epithelial, 70 as intermediate squamous epithelial, and 98 as columnar epithelial in the final database.

- Abnormal Cells-675 cells: In the final database, there were 182 cells that were classified as having mild non-keratinizing squamous dysplasia, 146 cells as having moderate nonkeratinizing squamous dysplasia, 197 cells as having severe non-keratinizing squamous dysplasia, and 150 cells as having intermediate squamous cell carcinoma in situ.

IV. 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

To verify its image analysis capabilities across the processing pipeline, the automated cervical cancer screening system was put through a rigorous testing procedure. According to preliminary assessments, the system was able to effectively enter photos and convert them to grayscale for the best possible preprocessing. The system's excellent segmentation capabilities were proven by removing the background and isolating the regions of interest. The noise in the photos was successfully decreased via Gaussian filtering.

The system demonstrated expert feature extraction by identifying distinct cell features using morphological and textural analysis approaches. The system's capacity to correctly classify cells into normal and pathological kinds was validated by classification testing utilizing a variety of techniques, including SVM and neural networks. Comprehensive testing confirmed the automated screening system's ability to classify and analyze images analytically for the purpose of screening for cervical cancer.

Testing revealed that by applying the Canny edge recognition algorithm to recognize borders and edges in intensity pictures, the system could correctly identify cell outlines. The system's faultless execution of morphological closure procedures confirmed its capacity to carry out fundamental morphological operations required for image processing.

Testing most notably revealed the system's remarkable capacity to recognize malignant cervical cancer cells, underscoring its fundamental strength in cancer detection. The system's accuracy calculations utilizing the Jaccard similarity coefficient demonstrated its capacity to analyze its performance quantitatively.

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

CONCLUSION

In conclusion, testing confirmed that the automated system for screening for cervical cancer is capable of carrying out the critical analytical processes required to identify cancerous cells in cervical pictures. The system showed proficiency in performing morphological operations, segmenting regions of interest, identifying cell edges, using filtering techniques, preprocessing pictures, and correctly categorizing malignant cells—a crucial ability for cancer diagnosis.

The system's capacity to build a whole end-to-end image processing pipeline, from initial picture upload to final cancer prediction output, demonstrated its competency in using computer vision and machine learning for automated cervical cancer screening. The automated system has the necessary picture analysis and classification skills to support cervical cancer screening, as demonstrated by the successful testing.

The accuracy of the system may be guided by a quantitative performance standard that is provided by the Jaccard similarity coefficient, which is used to quantify accuracy. These results suggest that the automated approach has potential as an effective cervical cancer screening tool, with the ability to decrease clinician workloads and boost detection rates, even if further real-world testing and improvement are still needed.

The system's successful completion of this extensive battery of tests verifies its capabilities and methods, indicating a significant step towards its ultimate clinical integration and practical application. All things considered the positive test results show that the system is a viable automated tool for cervical cancer screening.

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).