Development of Deep Learning-Based Approach for Early Diagnosis of Cervical Cancer

Thesis Submitted for the Award of the Degree of

DOCTOR OF PHILOSOPHY

in

Computer Science and Engineering

By

Nahida Nazir

Registration Number: 42000478

Supervised By	**Co-Supervised By**
Dr. Baljit Singh Saini (22078)	**Dr. Abid Sarwar**
Computer Science and Engineering	**Computer Science and Applications**
Professor	**Senior Assistant Professor**
Lovely Professional University, Punjab	**University of Jammu, Jammu**

LOVELY PROFESSIONAL UNIVERSITY, PUNJAB
2024

DECLARATION

I, hereby declare that the work presented in the thesis entitled "**Development of Deep Learning-Based Approach for Early Diagnosis of Cervical Cancer**" in fulfilment of degree of **Doctor of Philosophy** is outcome of research work conducted by me under the supervision of Dr. Baljit Singh Saini and Dr. Abid Sarwar, working as Professor and Senior Assistant Professor, in the Department of Computer Science and Engineering of Lovely Professional University, Punjab, India, and University of Jammu, India . In accordance with the general practice of reporting scientific observations, due acknowledgments have been made whenever this work is based on the findings of other investigators. This work has not been submitted, in part or in full, to any other university or institute for the award of any degree.

Signature of Scholar

Name of the scholar: Nahida Nazir

Registration No.: 42000478

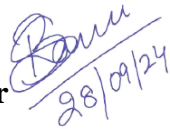Department/school: Department of Computer Science and Engineering

Lovely Professional University,

Punjab, India

CERTIFICATE

This is to certify that the work presented in the Ph. D. thesis entitled "**Development of Deep Learning-Based Approach for Early Diagnosis of Cervical Cancer**" submitted in fulfillment of the requirement for the award of degree of **Doctor of Philosophy (Ph.D.)** in the Department of Computer Science and Engineering, is a research work carried out by Nahida Nazir, Registration No. (42000478), is bonafide record of her original work carried out under my supervision and that no part of thesis has been submitted for any other degree, diploma or equivalent course.

Signature of Supervisor   **Signature of Co-Supervisor**

Name of supervisor: Dr. Baljit Singh Saini   Name of Co-Supervisor: Dr. Abid Sarwar

Designation: Professor   Designation: Senior Assistant Professor

Department: Computer Science and Engineering   Department: Computer Science and IT

University: Lovely Professional University, Punjab   University: University of Jammu

ABSTRACT

Cervical cancer, ranking as the fourth most prevalent cancer among women globally, remains a significant health concern, emphasizing the urgent need for accurate and timely detection methods. In response, the groundbreaking development of DeepCervix-Net emerges as an innovative AI-based approach, strategically designed to automate the comprehensive analysis of cervical cell images, from segmentation to classification, with the overarching goal of enhancing early detection strategies and subsequent medical interventions. The foundational pillar of DeepCervix-Net lies in its meticulously structured three-stage process: semantic segmentation, feature extraction, and classification. Each stage operates cohesively to meticulously dissect pap-smear images, identify intricate cell structures, and categorize them according to the latest Bethesda Classification System. The initial stage, semantic segmentation, is important in precisely segmenting the cytoplasm and nucleus within cervical cells. DeepCervix-Net utilizes a modified U-Net architecture, which integrates Attention modules and Residual blocks to highlight spatial information and elevate segmentation accuracy. This stage's precision in identifying and segmenting cell structures forms a crucial foundation for subsequent analysis. Following semantic segmentation, Cell Profiler, a powerful tool, is integrated to meticulously extract morphological and textural features from individual cells. This critical step enables the characterization of diverse cell structures, furnishing detailed insights that significantly enrich the subsequent classification task. The extracted features serve as a comprehensive blueprint for understanding cell characteristics, essential for identifying potential cancerous and precancerous cells. The concluding stage entails the utilization of the innovative Artificial Neural Network Architecture model, precisely classifying cervical cells into seven classes. This robust classification process facilitates early detection of abnormal cells, indicative of potential cancerous or precancerous conditions. The use of deep learning techniques at this stage improves accuracy and makes decision-making easier for medical practitioners. DeepCervix-Net shows great promise in automating the segmentation, feature extraction, and classification of cervical cells for cancer screening. With a demonstrated overall accuracy rate of 98%, surpassing comparable techniques utilizing VGG16 and XGboost, it stands as a testament to its efficacy and potential to revolutionize cervical cancer diagnosis. By using the capabilities of deep learning techniques and deploying a holistic multi-stage approach, DeepCervix-Net offers a promising solution to significantly enhance the efficiency of cervical cancer screening. This groundbreaking innovation holds the

potential to reshape the landscape of cervical cancer detection, potentially reducing mortality rates and enhancing the overall well-being of women worldwide.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANFIS	Adaptive Neuro Fuzzy Inference System
ANN	Artificial Neural Network
AGC	Atypical Glandular Cells
ASC-US	Atypical Squamous Cells of Undetermined Significance
AttFPN	Attention Feature Pyramid Neural
AUC	Area Under the Curve
CHALE	Contrast Limited Adaptive Histogram Equalization
CIN	Cervical Intraepithelial Neoplasia
CNN	Convolutional Neural Network
DCNN	Deep Convolution Neural Network
DL	Deep Learning
ELM	Extreme Learning Machine
EVA	Enhanced Visual Assessment
FC	Fully Connected
FP	False Positive
FN	False Negative
GAN	Generative Adversarial Neural Network
GLCM	Gray-Level Co-occurrence Matrix
HCME	Hierarchical Convolutional Mixture of Experts
HDI	Human Development Index
HPV	Human Papillomavirus
HSIL	High Grade Squamous Intraepithelial Lesion
IoU	Intersection Over Union
ISBI	International Symposium on Biomedical Imaging
KNN	K Nearest Neighbor
LMIc	Low- & Middle-Income Countries
LSIL	Low-grade Squamous Intraepithelial Lesion
MLP	Multilayer Perceptron
MSE	Mean-Squared Error
NILM	Negative for Intraepithelial Lesion or Malignancy
PCA	Principal Component Analysis

ReLU	Rectified Linear Unit
RCNN	Region-Based Convolutional Neural Network
RNN	Recurrent Neural Network
RPN	Regional Proposal Network
SGD	Stochastic Gradient Descent
SPFNet	Subspace Pyramid Fusion Network
SPFNet	Series-Parallel Fusion Network
STIs	Sexually Transmitted Infections
TCT	ThinPrep Cytologic Test
TN	True Negative
TP	True Positive
UCI	University of California Irvine machine learning
VA	Variational Autoencoder
VILI	Visual Inspection Lugol's Iodine
VGG	Visual Geometry Group
VIA	Visual Inspection Acetic Acid
WHO	World Health Organization
YOLO	You Only Look Once

CHAPTER 1
INTRODUCTION

This chapter serves as an expansive exploration into the multifaceted landscape of cervical cancer diagnostics and understanding. It encompasses a comprehensive discussion on various pivotal facets, including a detailed portrayal of cervical cancer pathology, an analysis of associated risk factors such as HPV infection and socio-economic disparities, an overview of diverse screening methodologies like Pap smears and HPV DNA tests, and an exploration into the Bethesda System of Classification. Additionally, it delves into the transformative role of deep learning (DL) techniques in enhancing cervical cancer diagnosis, illuminating the impact of advanced computational methodologies on accuracy and efficiency. Furthermore, this chapter navigates through publicly available datasets pertinent to cervical cancer research, highlighting their potential utility for researchers. It also elucidates the motivation behind the research, identifying existing gaps in diagnostic accuracy and computational methodologies. The chapter culminates by articulating the research objectives and outlining the structural organization of the thesis to provide a coherent roadmap for the ensuing discussions and analyses.

1.1 Cervical Cancer

Cervical cancer originates in the cells of the cervix, the lower portion of the uterus (womb) which is connected to the vagina [1,2,3]. Most cervical cancer are caused by the human papillomavirus (HPV), a sexually transmitted infection. Approximately 123,907 women receive a diagnosis of cervical cancer annually, with 77,348 succumbing to the illness [4-5,46]. However, not all HPV infections lead to cervical cancer; some strains of HPV are considered high-risk and can lead to abnormal changes in cervical cells, potentially developing into cancer over time [4,5,6]. The progression from HPV infection to cervical cancer typically occurs slowly, often starting with precancerous changes in the cells of the cervix. These changes can be detected through regular screening tests such as Pap smears or HPV tests, allowing for early detection and intervention before cancer develops. In the initial stages of cervical cancer, symptoms might not manifest, but as the disease advances, individuals may notice irregular vaginal bleeding, pain during sexual intercourse, pelvic discomfort, and other associated symptoms [7-14]. Preventive measures for cervical cancer include HPV vaccination, routine screenings, and practicing safe sex [15-16]. Nearly all instances of cervical cancer stem from persistent infections by one of the 15 carcinogenic HPV genotypes [17-19], following a sequence of stages: initial infection of the cervical transformation zone's metaplastic

epithelium, sustained HPV infection, progression to cervical precancer from persistently infected cells, and eventual invasion through the epithelium's basement membrane [20-25]. HPV vaccines offer protection against HPV infection for girls and young women, yet current HPV vaccination rates remain notably low, even in some developed nations [26,27,28]. Moreover, the benefits of HPV vaccination typically extend to young women under 26 years old, particularly with the 9-valent HPV vaccines [29,30]. Despite vaccination, recommendations from the American Cancer Society advise vaccinated women to undergo regular screening like unvaccinated individuals, recognizing the inability to eliminate risk through vaccination alone [31-33].

Cervical cancer poses a significant global health challenge, in nations with lower and middle economic standings [LMICs] where effective interventions like HPV vaccination and improved screening methods remain less accessible [34,35]. While strides in reducing cervical cancer incidence and mortality have been notable in high Human Development Index (HDI) countries due to quality screening and timely treatments, progress has been slow in LMICs [36], where most cases and deaths occur. The WHO launched an initiative in 2020 to address the cervical cancer issues, aiming to lower incidence rates to below 4 cases per 100,000 women-years worldwide by 2030 [64]. This initiative sets ambitious targets, this includes ensuring that 90% of girls are vaccinated by the age of 15, screening 70% of females at least twice by age 45 with high-performance tests and ensuring treatment for 90% of identified cases. The WHO's strategy emphasizes the need for enhanced surveillance and monitoring to identify gaps and take targeted actions, highlighting the importance of continuous progress in the fight against cervical cancer on a global scale [37-40]. Figure 1 provides an overview of the human papillomavirus (HPV) life cycle, crucial for understanding cervical cancer development. On the left side of the diagram, a representation of a healthy cervix is depicted, serving as the baseline for cervical cancer analysis. Within this context, the progression of cervical cancer is elucidated through three distinct grades of cancerous cells, each delineated by the extent of cellular abnormalities and involvement. Notably, CIN3, positioned at the apex of the severity spectrum, signifies a substantial proportion of cells affected by malignancy. Transitioning towards the far-right section of the image, an abnormal cervix overtaken by cancer is illustrated. Here, the cells exhibit profound deviations from the morphology of healthy cervical cells, showcasing a spectrum of severe deformities and dysplasia. This work focuses specifically on applying deep learning methods to the analysis and classification of cervical cancer images. The research aims to develop and evaluate models tailored to the unique features and diagnostic needs associated with cervical cancer. While the findings and methodologies presented in this work demonstrate

significant potential within this specific context, it is important to note that these results are not directly generalizable to other types of cancer or medical conditions. The methods developed and tested in this thesis are optimized for cervical cancer, and their applicability to other contexts would require further validation and adaptation.
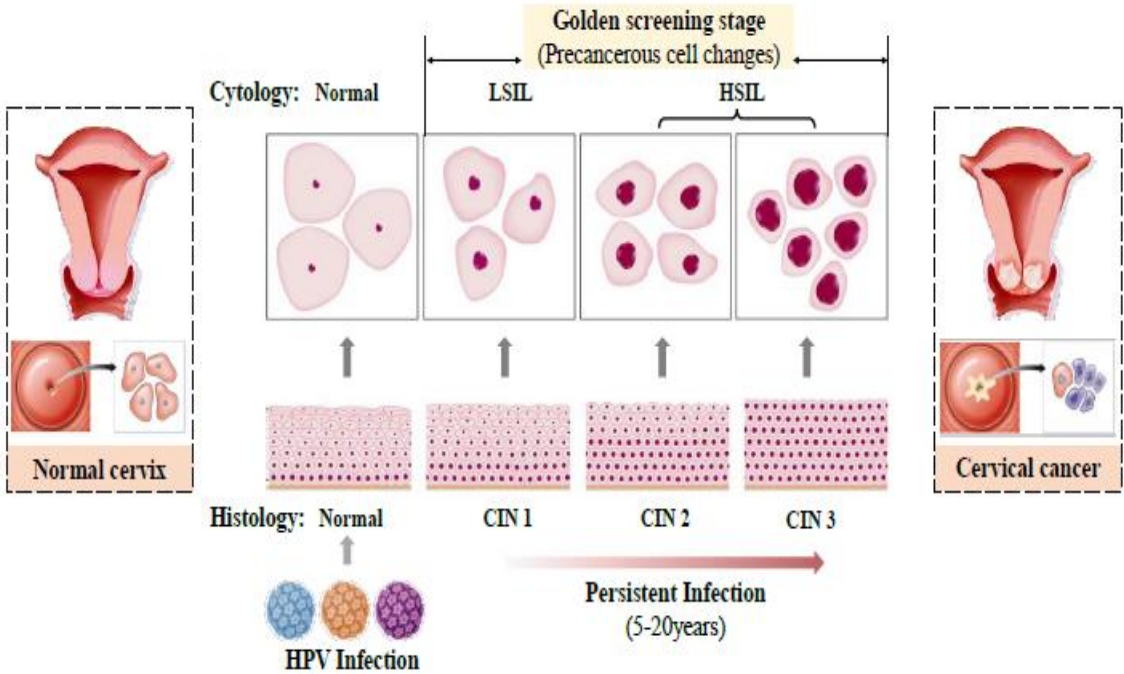
Figure1: Visual representation of HPV life cycle [33].

1.2 Risk Factors

Cervical cancer risk factors encompass a range of biological, lifestyle, and environmental elements that contribute to the likelihood of developing the disease.

1.2.1 HPV Infection

Infection with high-risk HPV strains, like types 18 and 16, is the main risk factor associated with cervical cancer. HPV is transmitted through sexual contact and can lead to cellular changes in the cervix that may progress to cancer if left untreated. Regular screening helps detect HPV-related changes early, preventing their advancement to cancer [40].

1.2.2 Sexual Behaviour

Certain sexual behaviours increase the risk of HPV exposure, which heightens the likelihood of cervical cancer. Engaging in sexual activity at a young age, having n number of sexual partners, or being in a relationship with someone who has multiple partners elevates the risk of HPV. Additionally, a history of other sexually transmitted infections (STIs) may exacerbate this risk [41].

1.2.3 Lack of Screening

Inadequate or infrequent cervical cancer screenings, like HPV tests, can delay the identification of precancerous or cancerous changes in cervical cells. Without timely detection and intervention, these changes might progress to more advanced stages of cancer [41].

1.2.4 Smoking

Tobacco use, particularly smoking, is linked to an increased risk of cervical cancer. Chemicals in tobacco smoke can damage cervical cells, making them more susceptible to HPV infection. Furthermore, smoking reduces the capability to immune system to combat off HPV infections and may also hinder the body's ability to clear the virus [41,42].

1.2.5 Weakened Immunity

Conditions that weaken the immune system, such as HIV/AIDS or the use of immunosuppressive medications, heighten the risk of cervical cancer. A robust immune response is crucial for clearing HPV infections. Individuals with compromised immunity may struggle to control HPV infections, leading to a higher risk of cervical cancer development [43-45].

1.2.6 Socioeconomic Factors

Socioeconomic disparities impact cervical cancer risk. Limited access to healthcare, including preventive measures like HPV vaccination and regular screenings, is associated with higher incidence rates. Socioeconomic factors also influence lifestyle choices and access to quality healthcare services, affecting the ability to prevent and detect cervical cancer early [46].

1.2.7 Oral Contraceptives

Prolonged intake of oral contraceptives may slightly elevate the likelihood of cervical cancer. However, this risk tends to decrease after discontinuing their use [47].

1.2.8 Genetics and Family History

A family history of gynaecological cancers or certain genetic factors might predispose individuals to a higher risk of developing cervical cancer. Genetic susceptibility can influence an individual's ability to combat HPV infections and their progression to cervical cancer [48].

Addressing these risk factors involves comprehensive strategies such as promoting HPV vaccination, encouraging regular cervical cancer screenings, advocating for smoking cessation programs, emphasizing safe sexual practices, improving healthcare accessibility, and conducting genetic counselling for individuals with a family history of gynaecological cancers [49].

1.3 Types of Screening Test

Cervical cancer screening aims to detect precancerous cell changes to prevent the development of cervical cancer through timely treatment. Occasionally, cancer is identified during screening. Detecting cervical cancer early typically facilitates easier and more effective treatment. However, when symptoms manifest, the cancer may have already started spreading, posing challenges for treatment. The various types of screening tests are depicted in figure 2.

1.3.1 HPV Test

This test looks specifically for the existence of high-risk types of HPV in cervical cells. The HPV test may be used alone or in conjunction with a Pap smear (co-testing) for more accurate identification of precancerous changes or early-stage cancer [50].

1.3.2 Pap Smear (Pap Test)

This test involves collecting cells from the cervix, which are then examined under a microscope to identify any abnormal changes in the cervical cells. It's effective in detecting precancerous or cancerous cells. The traditional Pap smear has been largely replaced by liquid-based cytology, where the collected cells are preserved in liquid before examination [51].

1.3.3 Visual Inspection with Acetic Acid (VIA)

These are low-cost screening methods suitable for resource-limited settings. It is also known as Lugol's Iodine (VILI) test.VIA involves applying dilute acetic acid to the cervix and observing it under light to detect precancerous changes. VILI uses iodine solution for a similar purpose. If abnormalities are detected in a Pap smear, HPV test, or during a visual inspection, a cervical biopsy may be performed. This involves removing a small tissue sample from the cervix for further examination under a microscope to confirm the presence of precancerous or cancerous cells. The choice of screening test or combination of tests depends on various factors, including age, risk factors, previous screening history, and healthcare resources available in a particular setting. Generally, regular screenings starting at age 21 or earlier if sexually active, and following recommended guidelines based on age and risk factors, are crucial for early detection and prevention of cervical cancer [52]. Figure 2 depicted below represents the common types of tests used in cervical cancer diagnosis.
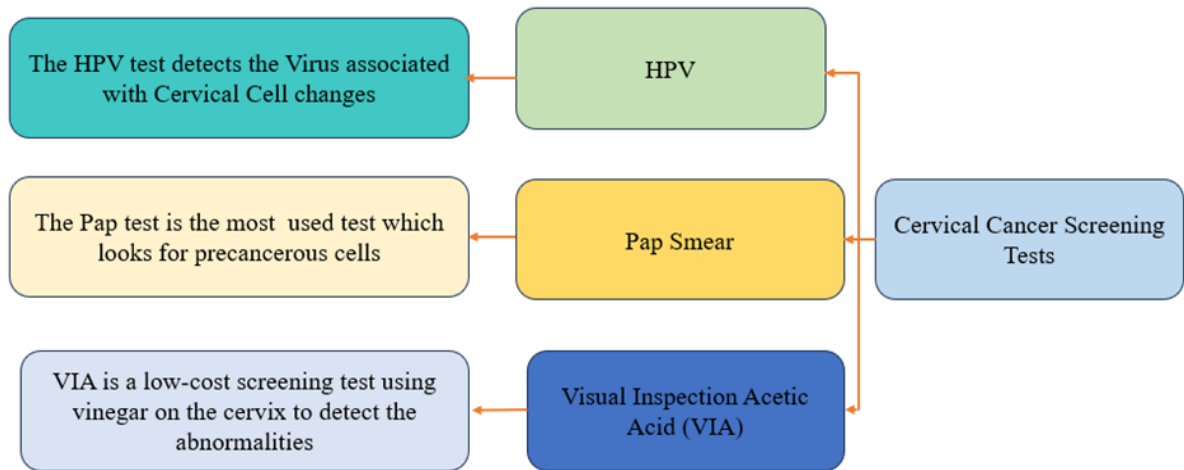
Figure 2: Basic Types of cervical cancer diagnostic tests.

1.4 Bethesda System

It is a standardized classification system used worldwide to interpret and categorize cervical cell abnormalities observed in Pap Smears. Developed to improve consistency and accuracy in reporting cytological findings, this system helps guide clinical management decisions and facilitates communication among healthcare professionals. The system, established in 1988 and subsequently updated in 2001 and 2014, categorizes cervical cell abnormalities into specific groups. Figure 3 represents the normal and abnormal endometrial cells.

1.4.1 NILM

Negative for Intraepithelial Lesion or Malignancy (NILM) indicates a normal Pap smear result, where no evidence of precancerous or cancerous cells is observed [53]. Figure 3a represents the ASC-US cells

1.4.2 Atypical Squamous Cells of Undetermined Significance (ASC-US)

These categories denote minor abnormalities in squamous cells that do not distinctly suggest precancerous changes but warrant further investigation due to their uncertain nature [53]. Figure 3b represents the ASC-US cells.

1.4.3 Low-Grade Squamous Intraepithelial Lesion (LSIL)

LSIL findings indicate mild cellular changes, often associated with human HPV infections, and typically pose a lower risk of progressing to cancer [53-55]. Figure 3c represented below is the LSIL image. It is considered a precursor to cervical cancer, but usually reflects early changes in the cells that are often caused by an HPV (Human Papillomavirus) infection.

6

1.4.4 High-Grade Squamous Intraepithelial Lesion (HSIL)

HSIL denotes significant to severe abnormalities in squamous cells, indicating a higher risk of progression to cervical cancer [53-56]. Figure 3d represents the HSIL of Pap Smear images.

1.4.5 Atypical Glandular Cells (AGC)

AGC findings suggest abnormalities in glandular cells of the cervix, necessitating further evaluation due to the potential for underlying neoplastic changes [53,57]. Figure 3e represents the Atypical Glandular Cell.

1.4.6 Squamous Cell Carcinoma or Adenocarcinoma

These categories represent definitive findings of invasive cervical cancer. The Bethesda System aims to standardize terminology, enhancing the accuracy and reproducibility of Pap smear interpretations, and guide appropriate patient management strategies based on the identified abnormalities [53,58]. Figure 3f represents the Squamous Cell Carcinoma.
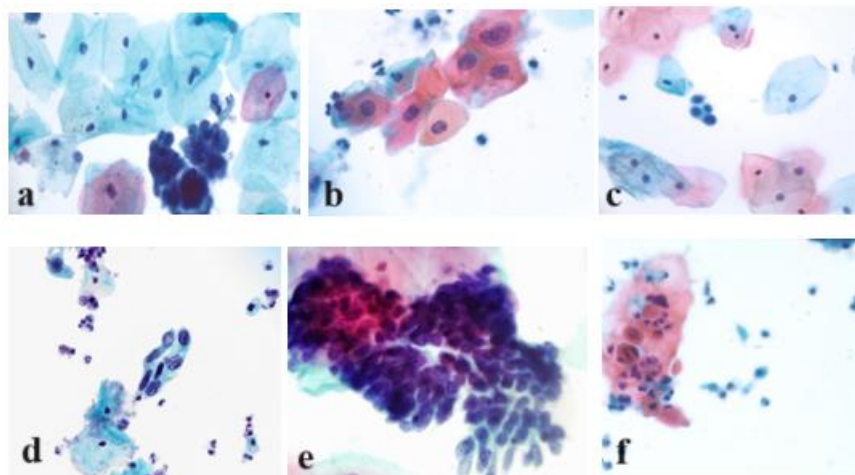
Figure 3: Bethesda system of cell classification [53].

1.5 Role of DL in Cervical Cancer Diagnosis

Over the past decade, deep learning has emerged as an important contributor in the domain of cervical cancer diagnosis. From 2014 to 2022, deep learning-based approaches have experienced remarkable growth and advancement, revolutionizing the way cervical cytology screening is conducted. The early years saw the exploration and development of foundational DL models and architectures, laying the groundwork for subsequent breakthroughs. As DL gained traction, its application in cervical cancer diagnosis expanded rapidly, leveraging large-scale datasets, and enhanced computational resources. From 2016 onwards, there has been a notable surge in research publications, indicating the increasing interest and recognition of DL's potential in this domain. Key areas of focus include object detection for precise identification

of abnormal cells, segmentation, feature extraction assessments, and the integration of DL with other diagnostic modalities for improved accuracy and efficiency. The trends from 2010 to 2022 clearly demonstrate DL's significant impact in advancing cervical cancer diagnosis, offering promising avenues for enhancing early detection, patient outcomes, and ultimately, reducing the global burden of this devastating disease. Figure 4 illustrates a comprehensive overview of DL techniques employed in automatic cervical cancer diagnosis, showcasing the diverse methodologies utilized for enhanced detection and classification. It is observed that CNN is frequently used deep learning approach in automatic diagnosis of cervical cancer.
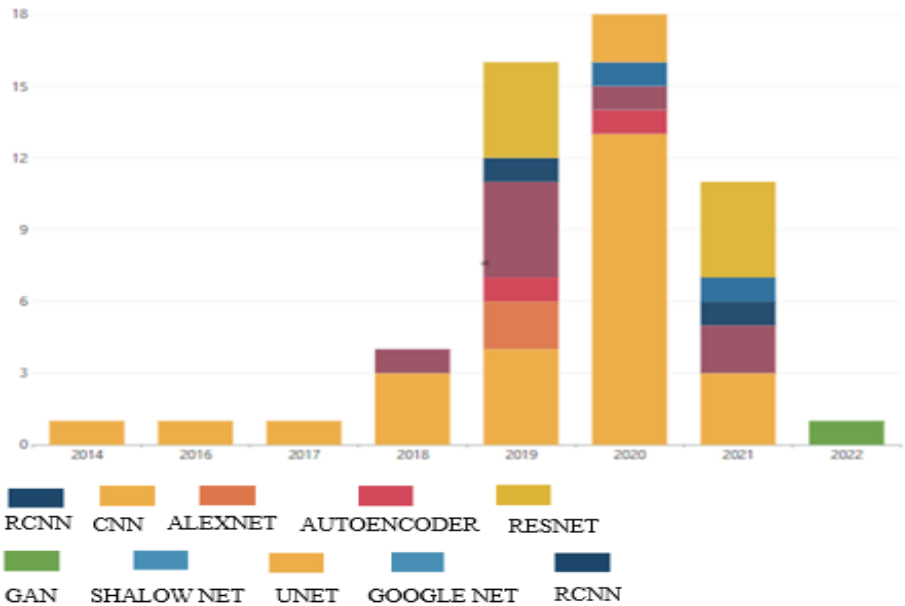
Figure 4: Deep learning trends in cervical cancer diagnosis from 2014 to 2022.

1.6 Public Datasets

The effectiveness of deep learning techniques depends on several factors, including computational resources and the availability of testing and training datasets. This section presents the most used datasets by the authors, but some researchers have also explored datasets like UCI, ISBI challenge database, and Guanacaste database.

1.6.1 Herlev Dataset

It is a publicly available dataset widely used in the domain of cervical cancer diagnosis. It consists of cervical cell images that are collected from Pap smear images, a standard technique for detecting precancerous changes in the cervix. The dataset is composed of both normal and abnormal cells. Researchers often utilize the Herlev dataset to train and evaluate DL models

for cervical cancer diagnosis. By using this dataset, they can develop computer-assisted diagnosis tools that can aid in the accurate and efficient screening of cervical cancer. The database contains a total of 917 images, comprising both normal and abnormal cases. These images are divided into seven separate classes, each based on different cell characteristics. The dataset can be accessed from the mentioned link http://mdelab.aegean.gr/index.php/downloads [59].

1.6.2 SIPakMed Dataset Description

The SIPaKMeD (Serbian Intelligent PAP smear analysis database) is a valuable resource in the domain of cervical cancer diagnosis and research. It consists of cervical cell images obtained from the Pathology department at the Clinical Center of Serbia. This dataset includes both normal and abnormal cell images, making it suitable for studying cervical abnormalities and developing computer-assisted diagnostic tools. The SIPaKMeD dataset is well-annotated, providing class labels for different cell types, including normal cells, LSIL, and HSIL. These annotations facilitate the training and evaluation of machine learning and DL models for cervical cancer classification and segmentation. By using this dataset, researchers can gain insights into the characteristics of cervical cells and explore novel approaches to improve early diagnosis of cervical cancer. The SIPaKMeD Database comprises a total of 4049 images, which were captured using a digital. This dataset can be accessed from https://www.cs.uoi.gr/~marina/sipakmed.html [60].

1.6.3 Intel and Mobile ODT Dataset Description

The Intel and Mobile ODT datasets are two valuable resources used for cervical cancer diagnosis and analysis. The Intel dataset contains cervical cell images captured using Intel's mobile devices. This dataset provides a large collection of images, including both normal and abnormal cervical cells. The dataset is annotated with class labels, allowing researchers to train and evaluate AI based models for cervical cancer detection and classification. The Mobile ODT dataset, on the other hand, consists of cervical images captured using the Enhanced Visual Assessment (EVA) system developed by Mobile ODT. This dataset includes images from both the cervix and the vaginal fornix, providing a comprehensive view of the cervical region. It is a valuable resource for analysing cervical abnormalities and studying various cell characteristics associated with cervical cancer [61]. The Intel and Mobile-ODT datasets, available on Kaggle's website (https://www.kaggle.com/c/intel-mobileodt-cervical-cancer-screening), consist of a total of 1481 training samples, 512 test images, and an additional 4633 images provided for extra training data [62].

1.6.4 ISBI 2014 Challenge Database

Biomedical Imaging (ISBI) 2014 Challenge. This database typically contains medical imaging data or related information that participants use to develop and evaluate algorithms or methods for specific tasks in biomedical image analysis. These challenges aim to advance the state-of-the-art in medical imaging technology and often focus on areas such as image segmentation, registration, or classification [62].

1.6.5 UCI Dataset

This dataset contains information on various risk factors related to cervical cancer, such as demographic factors, behavioural factors, and medical history. The data was collected from patients in hospitals in Brazil and is aimed at studying the risk factors associated with cervical cancer development. Researchers can use this dataset for studying and analyzing the relationships between different risk factors and cervical cancer incidence. It provides valuable information that can contribute to the understanding and prevention of cervical cancer, as well as the development of predictive models for early detection and diagnosis [63].

Cervical cancer diagnosis typically begins with the utilization of various primary datasets, which serve as fundamental resources for training and developing diagnostic models. These datasets encompass diverse cervical cell images, allowing researchers to study normal and abnormal cell patterns. Following the primary datasets, researchers often turn to the Herlev dataset, a well-known and extensively used collection of cervical cell images. The Herlev dataset provides annotated images divided into multiple classes based on cell characteristics, making it valuable for refining and evaluating diagnostic algorithms. Additionally, the SIPaKMeD dataset is a prominent secondary resource, comprising manually selected images from Pap smear samples, which aids in the analysis of cervical abnormalities. Alongside these primary and secondary datasets, researchers have explored other secondary datasets to further enhance cervical cancer diagnosis and contribute to advancements in automated screening methods. The below mentioned figure 5 represents the year wise usage of various public and private datasets. From the stacked bar graph, it is evident that real dataset is used more frequently compared to other datasets, which is followed by the Herlev dataset.
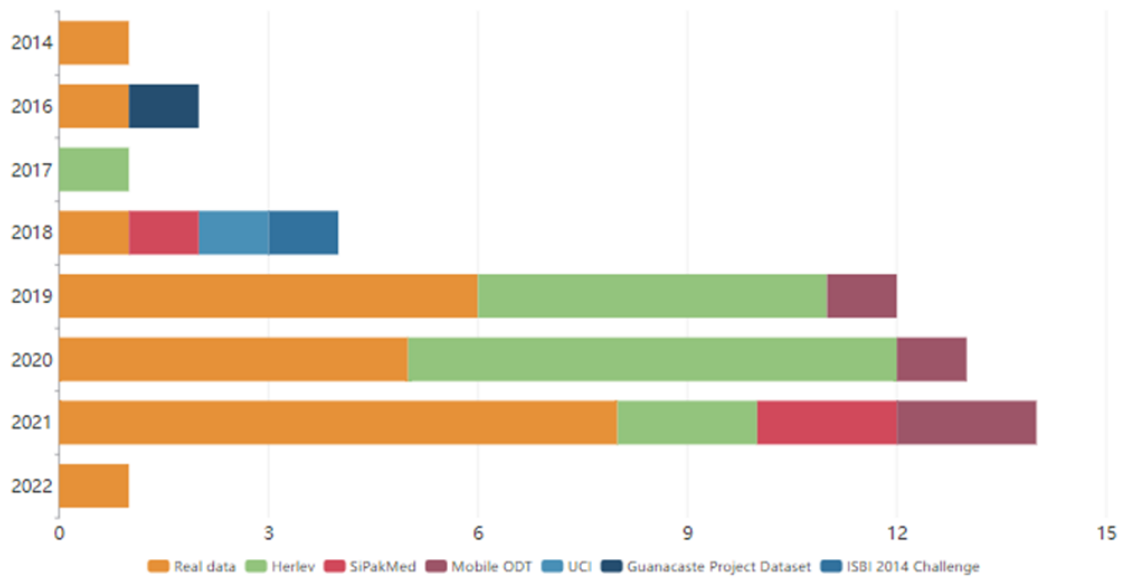
Figure 5: Yearly usage of public and private datasets in cervical cancer diagnosis.

1.7 Motivation

The motivation behind delving into cervical cancer diagnosis using deep learning stems from the urgent need to enhance early detection and accuracy in identifying cervical abnormalities. By using AI based approaches, this research aims to overcome challenges associated with human error in manual screening processes, providing a consistent and objective approach. The scalability of deep learning models facilitates the efficient processing of vast amounts of medical data, optimizing healthcare resources and improving accessibility to screening services. Ultimately, the goal is to significantly impact patient outcomes by enabling timely interventions, reducing false negatives, and contributing to the broader field of medical research, all with the overarching objective of alleviating the worldwide impact of cervical cancer and improving the quality of life for affected individuals. Our primary emphasis will be on the following aspects:

- Focusing on using deep learning to enhance early detection for identifying cervical abnormalities.
- Improving accessibility to cervical cancer screening services through efficient data processing, particularly in resource-constrained regions.
- Commitment to minimizing false negatives in cervical cancer diagnoses through highly sensitive and accurate deep learning algorithms.
- Aspiring to reduce the global burden of cervical cancer through innovative deep learning approaches, improving healthcare outcomes worldwide.

11

1.8 Research Gaps

- In contemporary biomedical research, the unfolding of folded cytoplasm's stands as a challenging frontier, and despite the strides made in image processing and deep learning techniques, a noticeable research gap persists in their collective ability to achieve comprehensive unfolding. Existing methodologies encounter difficulties in preserving the intricate spatial relationships within cytoplasmic structures, grappling with complex geometries, and adapting to the dynamic nature of cellular components during the unfolding process.

- The scarcity of benchmark datasets specifically designed for evaluating cytoplasmic unfolding represents an additional significant research gap. The absence of standardized datasets tailored to the intricacies of cytoplasmic structures hampers the development, testing, and comparison of unfolding algorithms. A lack of diverse and representative datasets impedes researchers from effectively benchmarking their approaches, hindering progress in the field. Establishing comprehensive benchmark datasets that encompass the variability in cytoplasmic structures will be crucial for advancing the accuracy and reliability of unfolding methodologies, contributing to a more robust foundation for future advancements in this domain.

- Medical images, especially those acquired during cervical cancer screening, are often affected by noise, artifacts, and variations in imaging conditions. Deep learning models may struggle to perform reliably in the presence of such challenges. There is a need for research focusing on developing robust deep learning models that can effectively handle noisy and artifact-ridden images.

- Cervical cancer diagnosis often involves multiple modalities such as histopathology, cytology, and imaging (e.g., colposcopy, MRI). There is a gap in research focusing on the development of deep learning models capable of integrating information from multiple modalities to improve diagnostic accuracy and reliability.

- A significant gap exists in pre-existing research efforts: there's a lack of advanced techniques to effectively profile features from folded cytoplasm. Current methods fall short in capturing the detailed characteristics of these folded structures. Addressing this gap is crucial for gaining a deeper understanding of the functional aspects of cytoplasmic components. Developing specialized techniques for accurate

feature profiling will be essential to close this research gap and uncover insights into the role of folded cytoplasm in cellular biology.

- The absence of annotated datasets containing both cytoplasmic and nuclear features poses a significant obstacle in training accurate deep learning models for cervical cancer detection. Current research predominantly focuses on individual cellular components, overlooking the vital interplay between cytoplasm and nucleus crucial for precise cancer classification. Addressing this gap by developing annotated datasets with comprehensive cellular features will enable the training of more effective deep learning models, thereby improving the accuracy of cervical cancer screening methodologies.

1.9 Research Objectives

- To analyze the existing image processing and deep learning algorithms.
- To collect and preprocess the appropriate cervical image dataset.
- To segment cell organelles from individual cells.
- To extract the features from processed dataset.
- To develop a deep learning model for profiling of cells with folded cytoplasm.
- To test and validate the proposed model.

1.10 Thesis Organization

The thesis on automated diagnosis of cervical cancer is meticulously structured into distinct sections, each contributing significantly to the comprehensive understanding of this critical domain. Chapter 2, the Literature Review, serves as a foundational pillar, meticulously surveying and synthesizing prior research and studies pertinent to automated cervical cancer diagnosis. This extensive review contextualizes the research within the existing landscape, elucidating the evolution of methodologies, technologies, and challenges faced in this specialized field. Chapter 3, Methodology, presents an intricate and detailed exposition of the employed approach. This section meticulously elaborates on the research methodology, detailing the selection criteria, data collection methods, technological tools utilized, and the step-by-step process of automated diagnosis. It elucidates the rationale behind the chosen methodologies, providing transparency and reproducibility while showcasing the robustness of the applied approach.

Chapter 4, Results and Comparison, emerges as the focal point where the culmination of research efforts is presented. This section meticulously unveils the findings gleaned from the conducted research, employing a comparative analysis approach. It not only showcases the

achieved results but also compares these findings with existing benchmarks or standards. Through this comparative analysis, it sheds light on the efficacy, strengths, and potential limitations of the devised automated diagnosis system, thereby emphasizing the significance of the contributions made by the research.

Finally, Chapter 5, Conclusion and Future Scope, serves as the intellectual zenith of the thesis. It provides an extensive and in-depth discussion, interpreting the implications of the research findings within the broader context of cervical cancer diagnosis. This section synthesizes the results, critically analyzes their significance, and explores avenues for further research and practical application. It culminates with a comprehensive conclusion, encapsulating the key takeaways, implications, and potential future directions arising from the study.

CHAPTER 2
LITERATURE REVIEW

The current chapter provides an examination and analysis of Pap smear images with a focus, on segmentation, feature extraction, and classification techniques. It delves into these methods using datasets from secondary sources aiming to offer a comprehensive understanding of their applicability and performance. A significant part of this investigation revolves around using deep learning architectures like Mask RCNN, U-Net, VGG16, Google-Net and other cutting-edge approaches. The goal is to clarify their contributions, nuances, and comparative strengths in the context of Pap smear image analysis. By covering a range of methodologies and datasets this chapter aims to facilitate improved insights, for optimizing the accuracy and effectiveness of medical image processing specifically tailored for Pap smear analysis.

2.1 Segmentation

Segmentation serves as the foundational step in the diagnostic pipeline, involving the precise isolation of relevant regions of interest within medical images. The delineation of abnormal cell nuclei or lesions from healthy tissue is fundamental for accurate diagnosis. A diverse array of techniques, ranging from traditional thresholding and edge detection methods to more sophisticated deep learning-based approaches, have been employed to achieve accurate segmentation. The success of subsequent stages greatly hinges on the precision of this segmentation process.

2. 1.1 Deep Learning for Segmentation

This subsection delves into the realm of DL segmentation approaches tailored for cervical cancer diagnosis. These techniques leverage convolutional neural networks and other sophisticated architectures to automatically delineate regions of interest within cervical images, such as the cervix, squamous epithelium, columnar epithelium, and potentially cancerous lesions. By utilizing large, annotated datasets and powerful computational resources, these approaches have unlocked new avenues for early detection and accurate assessment of cervical abnormalities. The section begins by introducing the foundational concepts of deep learning and explaining how they apply to medical image segmentation. Subsequently, it explores a spectrum of state-of-the-art architectures, ranging from U-Net and FCN (Fully Convolutional Network) to more recent innovations like Mask R-CNN, discussing their adaptations and optimizations for diagnosis of cervical cancer.

Song et al. [51] conducted a study on automatic segmentation of cervical cancerous cells using a combination of superpixel and CNN. The CNN achieved a precision of 0.914 ± 0.0202 and a recall of 0.872 ± 0.0008 for detecting nucleus areas, surpassing other methods like Backpropagation Neural Network and Support Vector Machine, demonstrating the effectiveness of deep learning in detecting nuclei and cytoplasm. The process involved pre-processing the images to remove Gaussian and impulse noise using a median filter. A coarse cytoplasmic mask was then created with a global threshold value, followed by fine segmentation using the superpixel-based algorithm. Finally, the CNN was implemented to extract various features from cervix cells. The dataset used for model training was collected from the Sixth People's Hospital of Shenzhen.

Sompawong et al. [65] used the Mask RCNN for cervical cancer diagnosis, utilizing the Herlev dataset for model training. Their technique achieved an accuracy of 89.8%, sensitivity of 72.5%, and specificity of 94.3%. Araújo et al. [6] emphasized a CNN approach for segmenting abnormal cervical cells. The technique ranked images based on the average area of segmented sections to determine the likelihood of aberrant cells. It effectively segmented both free-floating and clustered aberrant cells, employing a CNN and post-processing to remove these regions. The method proved resilient to noise and artifacts commonly found in conventional Pap smear images, offering faster performance without the need for prior segmentation. The results were ranked based on image likelihood.

Rigaud et al. [67] investigated deep learning models capable of auto-segmenting cervix anatomy, yielding equivalent results on two different datasets. The study integrated 2D DeepLabV3 and 3D U-Net. The 2D DeepLabV3 model was designed for single-pass inference, classifying each pixel into either the background or one of the 12 anatomic regions. On the other hand, the 3D U-Net model focused on classifying various body organs. To assess performance, the models were evaluated using metrics such as Hausdorff distance, dice similarity coefficient, and distance-to-agreement. Notably, both models achieved comparable and successful results for cervical cancer therapy segmentation.

Ma et al. [68] addressed the importance of precise and effective radiation target outlining for treating cervical cancer. To achieve autonomous contouring of clinical target volumes in cervical malignancies, authors explored deep learning auto-segmentation systems. A new type of network called VB-Net as an improved alternative to V-Net was introduced. In VB-Net, the letter "B" replaces the convolution, normalization, and activation layers of V-Net. This modification allows a significant reduction in model parameters since spatial convolution

operates on lower-dimensional feature images, enhancing overall efficiency. The model was trained on 535 computer tomography images. The proposed auto-segmentation approach outperformed manual contouring procedures. Consequently, this technique holds promise as a valuable tool to precisely outline radiation targets and improve the efficacy of cervical cancer radiotherapy.

Wang et al. [124] focused on addressing the complexity of classifying pap smear images, which is challenging due to overlapped cells, dust, and contaminants. Traditional feature-based techniques struggle with proper segmentation and feature extraction in this context, the researchers implemented deep learning as a viable alternative. This model comprises ten layers of convolution, succeeded by three fully connected layers. The convolution kernels in this network act as filters to extract edge information, aiding in feature extraction. By abstracting the image layer-wise, the model achieves enhanced feature extraction, resulting in improved discrimination and significantly higher classification accuracy. The researchers trained this model on a dataset consisting of 389 Pap smear images and achieved an impressive accuracy of more than 98%.

2.2 Feature Extraction

The subsequent stage entails the extraction of salient features from the segmented regions. These features encapsulate essential characteristics of the cell nuclei or lesions and serve as the basis for subsequent classification. Morphological, textural, and spatial attributes are among the key features extracted. To preserve relevant information while reducing dimensionality, feature extraction techniques such gray-level co-occurrence matrices, and statistical measures are frequently employed.

2.2.1 DL for Feature Extraction

Traditional methods of feature extraction from medical images often relied on handcrafted algorithms, which were constrained in their ability to reveal the nuanced patterns present in cervical images. The deep learning architectures has revolutionized the process of feature extraction, offering a powerful solution to effectively mine complex information from images and significantly enhance the diagnostic accuracy for cervical cancer. This subsection emphasizes the significance of utilizing deep learning-based feature extraction techniques for cervical cancer diagnosis. Unlike conventional approaches that require manual crafting of features, deep learning networks can automatically learn and extract high-level features from raw image data. This intrinsic capability enables these networks to uncover subtle and previously unnoticed patterns, textures, and structures within cervical images that are indicative

of potential abnormalities. By employing convolutional layers, pooling operations, and other architectural components, deep learning models can hierarchically learn features at various levels of abstraction. This hierarchical feature extraction is particularly advantageous for cervical cancer diagnosis, as the disease often presents intricate variations in tissue appearance and structure. Deep learning networks can capture these nuanced variations, allowing for improved discrimination between normal and abnormal cervical tissues. Furthermore, the ability of deep learning models to adapt and generalize to diverse image datasets is instrumental in addressing the challenges posed by the variability in image quality, patient demographics, and imaging conditions. Incorporating deep learning-based feature extraction into the diagnostic pipeline not only enhances accuracy but also potentially reduces the workload for medical professionals. Automating the process of feature extraction through DL enables faster and more consistent analysis, facilitating earlier detection of abnormalities and streamlining the overall diagnostic workflow.

Xu et al. [69] proposed a deep multimodal network for diagnosing cervical dysplasia, utilizing feature learning based on AlexNet. The network includes five convolutional layers, two connected layers, and a 1000-way SoftMax layer, with transfer learning utilized during training. The CNN feature vector from the fc7 layer has a larger dimension than the non-image feature. The dataset for training was collected from the National Cancer Institute (NCI), and the suggested framework learns complementary features through backpropagation. The model achieved 87.83% sensitivity and 90% specificity, providing accurate diagnosis for cervical dysplasia, surpassing techniques that use only one information source or prior multimodal frameworks. Plissiti et al. [70] introduced an annotated image dataset for cervical cancer analysis, exploring 26 features using a convolutional neural network (CNN) on the SIPaKMeD dataset. The CNN outperformed MLP and SVM in terms of performance.

Hu et al. [71] presented a CNN for automated detection of cervical precancer and cancer. This technique aims to address the lack of affordable HPV vaccination and cervical screening in low-resource areas with high cancer incidence. The proposed system performs two key functions: first, it detects and locates the cervix in input images, and second, it predicts the likelihood of the image being a CIN2 case. The Faster R-CNN algorithm handles object identification, feature extraction, and classification. The method was trained and validated using digitized cervical images obtained through cervicography from a population-based longitudinal cohort of 9406 females in Guanacaste, over a seven-year period. The R-CNN

method provides a faster and more accurate balance, making it an effective tool for cervical cancer detection.

Mousser et al. [72] provided a comprehensive analysis on cervical cancer using Pap smear images. Cervical cancer is a significant societal health concern, and early detection through accurate Pap smear tests is essential. Traditionally, cytopathologists have relied on hand-crafted features to assess the health of cervical structures. However, DL techniques have shown greater promise in classification accuracy. The authors employed four CNN models to extract deep features. These features were then used as input for a multi-layer perceptron classifier. The study was conducted on the DTU/HERLEV database, and the results revealed that ResNet50 outperformed VGG and InceptionV3, achieving an impressive accuracy of 89%. The findings highlight the significance of adopting DL methods in medical images for enhanced early detection and prevention of this life-threatening disease.

Xiang et al. [73] proposed an automatic cervical screening technique using object detection. This approach offers an effective and segmentation-free solution for analyzing cervical cells automatically. The method utilizes YOLOv3, a contemporary object detector, which eliminates the need for hand-crafted features. YOLOv3 consists of two components: Darknet53, a deep architecture for feature extraction trained on ImageNet with 53 layers, and multi-scale feature fusion layers for predictions. Unlike traditional multi-stage approaches relying on segmentation accuracy and hand-crafted features, this method instantly extracts high-level features and detects cervical cells promptly. The model achieved impressive results, with 97.5% accuracy, a mean average precision of 63.8%, and a specificity of 67.8%. This research represents a significant advancement in cervical cancer screening, providing a more efficient and accurate solution for early detection.

Xia et al. [74] introduced Series-parallel fusion network (SPFNet), a novel network topology for cervical cancer cell detection. Unlike standard architectures that use classification models for feature extraction, SPFNet applies multiple combination tactics in the series module and incorporates five distinct head components to optimize the detection task. The approach involves three main steps: feeding the processed image into the network, extracting relevant features, and obtaining different feature maps. These maps with various resolutions are then passed through the Regional Proposal Network, resulting in 256 regions of interest. Subsequently, the R-CNN head is applied to these regions, enabling SPFNet to extract high-level semantic and spatial information's. The SPFNet achieved remarkable results, attaining an average precision of 78.4% in identifying cancerous cells, outperforming traditional detection

methods significantly. This novel network presents a promising approach for cervical cancer detection, demonstrating its superiority in comparison to other existing techniques.

Alyafeai et al. [75] presented a robust deep learning approach for diagnosing cervical cancer, exploring techniques like RCNN, YOLO, and GoogleNet. Their proposed method demonstrated significant efficiency, being 1000 times faster than existing models. It achieved an intersection of union of 0.68 and an AUC score of 0.82, outperforming other techniques by 20 times. In another study by Lee et al. [76], introduced a deep learning method using YOLO V3, a lightweight object recognition algorithm. This approach allows the model to operate efficiently on low-performance devices while maintaining high accuracy. The model uses bounding boxes to identify abnormal cells from normal cells in cervical cancer images. The algorithm's performance was measured using average precision, which reached 73.34%. This method could serve as an auxiliary tool for pathologists in cervical cancer detection.

Khamparia et al. [77] proposed an innovative approach by combining a convolutional neural network with a variational autoencoder (VA) for cervical cancer analysis. The model's encoding phase uses convolution to generate high-dimensional feature vectors, while the decoding phase employs autoencoders to reconstruct damaged cells. Data augmentation and VA generate diverse outputs during training. Experimental results showed that VA provided spatial image features for convolutional networks, leading to the best classification results for cervical cells. The accuracy of classifying the normal class was significantly higher with a 3*3 filter compared to a 2*2 filter. The designed architecture achieved a variational accuracy of 99.2% with a 2*2 filter size and 99.4% with a 3*3 filter size. The Herlev dataset was used for model training. This hybrid variational convolutional autoencoder technique outperformed other state-of-the-art machine learning algorithms, offering promising results for cervical cancer analysis.

Jia et al. [78] developed a novel approach to detect abnormal cervical cells using the YOLO method, ensuring fast and accurate detection. To improve the model's ability to recognize cell features effectively, authors introduced DenseNet and Stochastic-pooling (S3Pool) algorithms into the Darknet-53 feature extraction network. DenseNet connects lower layer features with higher layer features, creating a more cohesive network structure. S3Pool enhances the network's generalization capability, leading to better feature extraction and improved overall performance. The model was trained on Herlev dataset and a primary dataset. This training resulted in an increase in average detection accuracy from 70.65% to 78.87%. Additionally, to handle complicated backgrounds, dense cell clusters, and irregular cells, the model utilized

Focal Loss and balanced cross-entropy functions, enhancing its diagnostic ability. Moreover, the integration of transfer learning a technique where pre-trained models are leveraged as starting points significantly accelerates the training process and enhances classification performance, even when the available cervical cancer dataset is limited. This transfer learning approach takes advantage of knowledge gained from larger datasets, effectively increasing the model's ability to generalize to the nuances of cervical cancer images.

Attallah et al. [79] cervical cancer prevention involves routine screenings, with Pap smears being common but prone to errors. The authors introduced an advanced computer-aided diagnostic (CAD) model leveraging Artificial Intelligence. Unlike existing CADs, it avoids complex pre-segmentation, using three efficient deep learning models for spatial features. Through rigorous analysis, the model achieves 100% accuracy using SVM. The CAD's performance underscores the effectiveness of combining diverse features for improved diagnostic accuracy, as demonstrated through comparative analysis with existing studies.

Kalbhor et al. [80] presented a new approach for automatic cervical cancer diagnosis. The experimental results revealed that ResNet-50 stands out with superior accuracy compared to other models. Simple Logistics achieves the highest accuracy among machine learning algorithms. The authors utilized features from the ResNet-50 pre-trained model with Simple Logistic classification performs best for pap smear image classification.

Chung et al. [81] focused on developing an auto-segmentation model for organs-at-risk (OARs) and clinical target volumes (CTVs) in cervical cancer radiotherapy. Experimental findings indicated acceptable correlation between manual and auto-segmented contours for various structures. Auto-segmentation demonstrated favourable outcomes in terms of reducing heterogeneity and contouring time, with most physicians expressing a preference for the auto-contouring system. The study emphasizes the feasibility and efficacy of the proposed DL model for the early treatment of cancer.

Xu et al. [82] investigated SIPaKMeD and Herlev datasets, categorized into eight types, with 3612 images. Initial attempts with the Faster R-CNN yielded suboptimal classification outcomes. Subsequent analysis led to the development of an improved classification of cervical cells. Comparative evaluations using P-R curves, AP values, and mAP values against prominent detection algorithms demonstrated the superior performance of the proposed model over CenterNet, YOLOv5, and Faster R-CNN. Comparison with recent literature on cervical cell classification revealed higher accuracy and a precision score of 99.81%. While the precision of automatic classification of pathological images using DL algorithms requires

further rigorous testing, the current results offer a valuable preliminary analysis that can aid pathologists in improving their work efficiency.

2.3 Classification

Classification is the final stage of the diagnostic pipeline, involving the allocation of labeled data points into predefined classes based on the extracted features. Machine learning approaches, encompassing support vector machines, random forests, and neural networks, have been harnessed for this purpose. Particularly, CNNs, exhibit remarkable performance due to their ability to automatically learn intricate hierarchical features directly from images. Rigorously trained on annotated datasets, these models accurately classify cervical cells and predict malignancy likelihood. The application of these advanced algorithms not only expedites the diagnostic process but also mitigates inter-observer variability and subjectivity. As the synergy of medical imaging and machine learning progresses, these components continue to evolve, promising heightened accuracy, dependability, and accessibility in cervical cancer diagnosis. This comprehensive approach embodies a significant stride towards bridging the gap between medical expertise and technological innovation.

2.3.1 Deep Learning for Classification

This subsection delves into a comprehensive exploration of various deep learning architectures employed for cervical cancer classification. From convolutional neural networks that excel at image-based classification tasks to recurrent neural networks (RNNs) capable of processing sequential data, a diverse range of models can be adapted and fine-tuned to discern different stages of cervical cancer with a great level of precision**.**

Bora et al. [83] introduced a robust system for efficiently recognizing dysplasia from pap smears. Authors integrated ALexNet architecture for model design and employed DCNN for extracting features. The experiments were conducted on a Herlev dataset and a clinical dataset, with the private database containing 1611 categorized samples. The model operates in various steps. Firstly, a private dataset is created, comprising images from class1, class2, and class3. The system is also validated using the Herlev dataset. Next, features are extracted using AlexNet. In the subsequent phase, the model minimizes various features by selecting a subset of features. Finally, cell classification is performed using the mentioned classifiers. By implementing the feature selection technique, the accuracy of the system can be improved to reach 90-95 percent.

Zhang et al. [84] introduced a CNN-based cervical cell classification system that overcomes the limitations of previous methods. Unlike earlier systems that relied on cytoplasm-nucleus segmentation and manual feature engineering, their technique extracts deep features without the need for segmentation. The approach involves pre-training the ConvNet on a real-world image dataset and then fine-tuning it using the images of cervical cells, which is composed of adaptively re-sampled, coarsely centered patches of image. The experimental results demonstrate its superiority over previous algorithms, achieving an impressive 98.3% accuracy, 0.99 area under the curve, and 98.3% specificity in classification. This method outperforms other approaches on both the Herlev dataset and liquid-based cytology datasets. However, a notable limitation of this research is that it requires 3.5 seconds for single patch classification, which is impractical for clinical settings where real-time processing is necessary.

Rohmatillah et al. [85] introduced an automatic cervical cancer classification system, which operates in three distinct steps. Firstly, different features are extracted using a CNN. In the second phase, the number of features is reduced through the application of PCA and LDA and followed by the classification process. The experimental findings demonstrated that the proposed approach outperforms existing techniques for cervical cancer classification.

Harinarayanan et al. [86] introduced VGG16, for cervical cancer diagnosis. This innovative approach involves a segmentation-free deep learning algorithm for classifying PAP smear images. The proposed method generates a map highlighting important regions within the images, aiding pathologists in their investigations using the network's inherent knowledge. This map is created through subtraction, eliminating the need for traditional segmentation techniques, and significantly reducing the time required to gather training and testing data. By leveraging extracellular information, the algorithm enhances data classification. The intrinsic information within the neural network provides doctors with a map of significant images and regions, aligning with the requirements of assisted screening. This map not only improves result interpretability but also expedites the evaluation process for doctors, making the entire diagnostic process more efficient.

In separate research done by Promworn et al. [87] five DL models were explored to locate abnormal cells from cervical cells. The model training was accomplished using the Herlev dataset. The experimental results revealed that Densnet 161 outperformed other approaches. Guo et al. [88] developed and fine-tuned several deep learning approaches, such as RetinaNet, fine-tuned VGG, and Inception-based models, to improve cervical cancer screening. The authors incorporated an automatic cervix detection method using a pre-trained deep learning

model to provide bounding box annotations for observed cervix regions. When necessary, the image could be cropped to this region for further processing. Subsequently, the leading object identification model, RetinaNet, along with VGG16 as base networks for image sharpness and classification was used. Among various techniques, RetinaNet with ResNet50 proved to be the most effective. The developed deep learning approach demonstrated the capability to identify precancerous cells faster than human specialists, especially in women aged 25 to 49. Improving cervical cancer diagnosis and early detection of lesions.

Hussain et al. [89] explored six deep convolutional neural networks to detect cervical cancer's precancerous and cancerous lesions. The authors utilized both the Herlev dataset, and a hospital-derived pap smear dataset gathered from various medical facilities, containing 1670 liquid cytology-based and 1320 pap smear test samples. The proposed technique resolves inaccurate predictions without needing segmentation or hand-engineered feature extraction processes. The ensemble method proved beneficial as it focused on all dysplasia stages, suggesting its potential for early-stage illness diagnosis. Among the models tested, GoogLeNet attained the topmost accuracy and the smallest log-loss, whereas AlexNet displayed the lowest accuracy and the highest log-loss. Deeper networks, like Googlenet, performed better, whereas Alexnet networks showed a 37% reduction in the number of trainable parameters. This research highlights that fine-tuned Googlenet, Resnet-(50 and 101) models effectively learn more visual deep characteristics of pap smear images compared to other pre-trained models.

Yilmaz et al. [90] performed a comparative examination between traditional ML and DL methods for the cancer classification of cervical cells. Experimental findings revealed that CNN surpassed traditional ML techniques in cervical cancer classification, achieving an accuracy of 93% compared to traditional methods reaching a maximum accuracy of 85%. Notably, features extracted by CNN were superior to manually crafted ones. Furthermore, the CNN model's training time was minimized to 22 seconds per epoch.

Martínez et al. [91] proposed a novel cell merger and CNN approach for accurate cervical cancer analysis. Early diagnosis of cervical cancer is crucial to reduce mortality rates. While most algorithms involve cell pre-processing, this research directly used folded cells to build the CNN model, creating a robust and reliable cervical analysis system. The presented model achieved an accuracy of 88.8%, making it suitable for clinical screening. The dataset comprised cells from ten patients, categorized into four classes: normal squamous, L-SIL, ASC-US, and H-SIL. The cells were further labeled as "review not needed" for squamous cells and "review needed" for L-SIL, ASC-US, and H-SIL to enable proper cervical cancer classification.

Mohammed et al. [92] implemented a CNN technique for classification of cervical cells. Authors evaluated the top 10 pre-trained deep CNN models, including DenseNet 169, ResNet (101 and 152) for classifying cancerous cells. The presented technique involved image acquisition, pre-processing, feature extraction, and finally cell classification. DenseNet169 outperformed the other 9 deep CNN techniques, demonstrating a minimum size of 57 MB, making it easier for deployment on portable devices. The model achieved an impressive, normalized accuracy of 0.990, making it a promising tool for accurate classification of single-cell pap smear images.

Tan et al. [93] introduced an efficient DCNN technique to help medical practitioners in cervical cancer screening. Authors conducted a retrospective investigation of multicentre ThinPrep cytologic test (TCT) images to develop a CNN-based TCT screening system. This approach significantly improves the screening of cervical cancer, addressing the bottleneck in medical resources. The system can identify images and provide a test report in just three minutes, reducing pathologists' workload and freeing up time for complex case investigation. The proposed system achieved a specificity of 34.8%, sensitivity of 99.4% and an AUC of 0.67. Dhawan et al. [94] presented various DL methods for classification of cervix. Authors utilized VGG19, Inception V3, and ResNet50 architectures. VGG19 and Inception V3 were employed for classification, while ResNet50 benefited from transfer learning. The dataset comprised 5287 labeled images collected from Kaggle, pre-processed using a Python inbuilt function. Fine-tuning multiple parameters improved the model's performance, achieving an accuracy of 97.3%. The experimental results indicated that Inception V3 outperformed VGG19 and ResNet50.

Jia et al. [95] developed a novel model by integrating CNN with SVM for the classification of cervical cancer. Experimental findings revealed that LeNet-5 architecture for feature extraction outperformed other existing techniques. On the other hand, Park et al. [96] performed a comparative analysis between ML and DL algorithms for early diagnosis of cervical abnormalities. Authors worked with 4119 Cervicography samples, including normal and abnormal cells. For comparison, authors used the ResNet-50 architecture against machine learning techniques (Random Forest, Extreme Gradient Boost, and Support Vector Machine). Pre-processing included image cropping for uniformity. The SVM classifier achieved an accuracy of 0.84, better than Extreme Gradient Boost and Random Forest, but ResNet-50 surpassed all three classifiers with an accuracy of 0.97. The experimental results showed that ResNet-50 outperformed current machine learning models for cervicography images.

Cao et al. [97] introduced a technique called attention feature pyramid network for spotting cervical cancerous cells. The technique consists of two vital parts. The initial segment emulates the procedure undertaken by medical practitioners when scrutinizing a cervical cytology image. The second module involves fusing refined features to detect abnormal cervical cells at multiple scales. The proposed algorithm achieves an average diagnosing time of 0.04 seconds per image, significantly faster than the average time taken by the doctor. Adweb et al. [98] developed ResNet-based networks for cervical cancer detection and investigated the impact of activation functions on ResNet's performance. Authors designed three ResNet models with distinct activation functions and evaluated them on a cervical image dataset. The findings revealed that ResNets with leaky and parametric rectified linear unit activation functions achieved similar accuracy, with 90.2% and 100% respectively. Both activation functions enhanced the performance of the ResNet-based network for cervical cancer detection.

Elakkiya et al. [99] addressed the challenges in proper segmentation of cervical cells, which has been problematic despite extensive research. To overcome this, authors proposed a blended method using Fine-tuned Stacked Autoencoder and Small-Object Detection GAN. The detection procedure consisted of three steps: data fine-tuning and normalization, cervical lesion identification and classification, and stage identification and cervical cancer prognosis. The approach utilized encoders to extract prominent features from images and object detection methods to locate areas of interest. Bayesian Optimization (BO) was employed to balance the network parameters. The dataset used to train the model was collected from public repositories and hospitals. The presented deep learning architecture utilized heterogeneous data to diagnose and predict cervical abnormalities.

Harangi et al. [100] presented a new screening technique for the detection and segmentation of cervical cells with high resolution, sensitivity, and specificity. The study consists of two phases: cell segmentation and deep learning-based cervical cancer classification. Compared to traditional algorithms, deep learning-based methods, such as FCN-8, FCN-16, and superpixel-based segmentation algorithms, outperformed with higher sensitivity and specificity. The dataset comprised 10,000 images from 6 digitized slides, used as input for the network. The evaluation metrics included sensitivity, specificity, Matthew's correlation coefficient, and intersection over the union.

In a related study, Lin et al. [101] introduced a robust cervical cancer diagnosis approach using GoogleNet, achieving good accuracy in 2 and 7 class classifications compared to other architectures. The experimental studies were conducted on the Herlev dataset. Additionally,

With the use of the Hierarchical Convolutional Mixture of Experts (HCME) method, Gorantla et al. [102] developed the CervixNet technology for the diagnosis of cervical cancer, reaching a remarkable 96.77% accuracy rate and a kappa value of 0.95.

Allehaibi et al. [103] presented a novel approach for cervical cancer classification. Authors utilized a Mask Regional CNN for cell segmentation, followed by classification using a smaller VGG-like Network. The first stage involved Mask R-CNN segmentation to divide cell areas, and the second stage focused on specifying the entire cell region based on the identified segments. Authors used ResNet10 as the backbone for Mask R-CNN, allowing it to exploit spatial information and prior knowledge to automatically segment and create pixel masks for each image component. For classification, a modified VGG-like Net was employed.

Rahaman et al. [104] presented a detailed survey on deep learning methods for diagnosing cervical cancer, a deadly disease affecting females. While the Pap smear examination is frequently employed for early identification purpose, it endures from a significant FP rate due to errors made by humans. To enhance the efficiency DL has been employed automatically segmenting and categorizing cervical cytology images. The study focused on deep learning techniques for segmenting and classifying cervical cytopathology images. MASK-RCNN combined with LFCCRF demonstrated superior performance in segmenting both overlapping and non-overlapping cells. On the other hand, for classification tasks, CNN (AlexNet) and decision-based tree algorithms outperformed other methods.

Dharani et al. [105] implemented an innovative technique for cervical cancer diagnosis that considers both single and overlapped cells. The study involved three main phases: detection, segmentation, and classification. Authors utilized Mask-RCNN, a deep learning technique, for segmenting cell nuclei. Classification was performed using decision trees to categorize cells as either normal or abnormal. The researchers collected the dataset for model training from the Guanacaste dataset. The study was carried out in two phases: firstly, implementing Mask-RCNN for cell segmentation, and secondly, classifying the segmented images based on the calculated area of cytoplasm and nucleus. To achieve multiclass classification using transfer learning, the authors examined conv5T, conv3T, and conv1T. Interestingly, the results revealed that successful segmentation was not a prerequisite for deep learning-based classification. For 2-class classification, the proposed approach achieved an impressive accuracy of 99.3%, and for 7-class classification, the accuracy attained was 93.7%. This research highlights a promising and efficient method for cervical cancer diagnosis.

Chen et al. [106] introduced an autonomous cervical cancer detection system based on deep learning techniques. The architecture of CytoBrain includes four max-pooling layers, two fully connected layers, and ten convolution layers which together extract features from input images and provide prediction probabilities for various classes. CytoBrain utilizes CompactVGG, a thinner and shallower variant of VGG16, to reduce computational cost while maintaining efficiency. The experimental results demonstrate that CompactVGG outperforms VGG11, which is the most efficient VGG network in the family, in terms of training and testing time. Kano et al. [107] examined a 2D and 3D UNet models based on automatic contour segmentation technique for cervical cancer diagnosis. The objective of this approach is to reduce the workload on oncologists. Although it has seen limited application in clinical settings, it shows potential for enhancing the diagnosis of cervical cancer.

Desiani et al. [108] presented an automatic analysis system for cervical cancer screening, focusing on segmenting and classifying the pap smear images. The manual screening process for cell classification is laborious and prone to inaccuracies, prompting the authors to work on an automatic solution for analysing cervical cancer cells. To enhance image quality, pre-processing steps such as Normalization, CLAHE, and Adaptive Gamma Correction were applied. The multi-class classification utilized the SoftMax function, with ANN and KNN used for the classification task. The Herlev database provided the dataset for model training. For segmentation, the UNET architecture, consisting of contracting and expanding paths, was employed for semantic segmentation of cytoplasm, nucleus, and background. Although the segmentation results with CNN were not ideal, they significantly impacted the classification process when using ANN or KNN. The experimental results confirmed the effectiveness of the proposed architecture in accurately classifying Pap smear images as either benign or malignant. The system showcased potential as an efficient and reliable tool for cervical cancer screening.

Kuko et al. [109] developed a novel method for categorizing cervical cancer. The research encompassed four main approaches: data collection, cell extraction, segmentation, and classification of abnormal cells. Authors collected 104 Pap smear images from the University of Southern California Medical Center. The cell extraction process involved converting input images to grayscale and using thresholding to convert them to black and white for improved readability and removal of debris from the area of interest. During segmentation and feature extraction, the cells were broken down into three distinct vector pixels, which were categorized into four different clusters using the K-means algorithm. The authors utilized 33 morphological features, including size measurements of the nucleus and cytoplasm, as well as mean color

values (red, green, and blue) of nuclei channels. The presented architecture consisted of five separate convolutional neural networks (CNNs), each trained on one of the pre-defined clusters to highlight irregularities among the cell clusters. Each convolutional layer had a ReLU activation function and dropout. Ensemble learning was applied with multiple random forests, achieving a promising accuracy of 90.37%. However, CNNs outperformed the ensemble learning approach, achieving an accuracy of 91.63%. This automated cell screening system has the potential to save time compared to manual examination by pathologists, as it can analyze most cells using ensemble learning and deep learning techniques. Notably, deep learning outperformed the ensemble learning approach, significantly reducing sensitivity and specificity variations by 95.47%.

Da et al. [110] implemented a hybrid model for cervical cancer diagnosis, combining ML and DL to overcome their individual limitations. The model starts with pre-processing of input images, which are then fed into a RetinaNet with a ResNet50 for deep learning analysis of the area of interest. Subsequently, feature extraction is performed, and a basic machine learning approach like SVM is used for classification. Two primary datasets were used for model training, and a RetinaNet with a ResNet50 backbone was employed to detect anomalous regions and classify them as low- or high-grade lesions. Nuclei within anomalous regions with a detection score above 50% were segmented using an iterative thresholding approach. The feature extraction module utilized the segmentation findings to locate 29 geometrical, color, and texture features from the nucleus structures, along with 840 additional features from the entire aberrant area. The final classified result was generated using two SVMs. The approach was trained and evaluated on the Herlev dataset. For the 7-class problem, the model achieved 95.9% accuracy, 96.2% sensitivity, and 99.3% specificity. A key advantage of this technique is its reduced reliance on extensive pre-processing compared to other methods. However, a drawback is the relatively high computational cost.

Pacal et al. [111] employed AI based approaches on the SIPaKMeD dataset, comprising cervix images. Experimental findings unveiled that the ViT-B16 transformer approach and EfficientNet-B16 CNN model performed exceptionally well. The max voting ensemble approach, applied to vision transformer techniques, demonstrated the highest performance in cervical cancer diagnosis. The proposed approach proves successful for early detection and holds potential for addressing other medical issues in the future.

Meza et al. [112] presented a method for automatically extracting features from Pap smear images and classifying them to diagnose cervical cancer. The first approach employs

techniques such as ResNet-50 and Simple Logistics which yielded the best performance. The second approach utilizes transfer learning with CNN models (AlexNet, ResNet-18, GoogleNet), showing GoogleNet's superior performance. Utilizing transfer learning to extract features yields greater accuracy compared to pre-trained CNN models independently. The study aims to inspire new methodologies and serve as a foundation for advancements in Pap smear image categorization.

Fekri et al. [113] presented a cervical cancer diagnosis approach with enhanced accuracy by proposing a method that combines VGG-19, ResNet34 and MLP. The classification accuracy is improved by replacing and flattening the classification layers. Results indicate higher accuracy compared to using each network separately and outperforming other state-of-art techniques. This method can be expanded to solve visual pattern classification issues in domains such as skin cancer diagnosis and image retrieval. The authors suggested using GoogleNet for future studies, emphasizing the generalizability and versatility of the proposed approach to cervical cancer diagnosis.

Youneszade et al. [114] provided a comprehensive review of current research on DL solutions utilizing image segmentation and classification techniques for the analysis of cervical screening images. It discusses key components and methods of DL techniques, emphasizing their significance in cytopathology and colposcopy image processing for cervical cancer detection. CNNs are acknowledged for their exceptional performance. Future studies may explore mixed feature selection using VGG19, RCNN, and Faster RCNN.

Chen et al. [115] introduced an independent Risk Score (RS) for outcome prediction in cervical cancer patients, distinct from primary clinical characteristics. Classifying patients into high- and low-risk groups based on the risk score demonstrates notable effectiveness.RS surpasses clinicopathological characteristics in predictive accuracy, attributed to its capacity to incorporate semantic structural tumor features alongside the often-neglected manual tumor information. This superiority is further enhanced by utilizing high-dimensional deep learning imaging features for analysis of Whole Slide Images, providing comprehensive insights into tumor microenvironment properties. In the current landscape of post-surgery prognostic stratification, RS proves advantageous, offering a more nuanced approach compared to the reliance on clinicopathological factors.

Kalbhor et al. [116] cervical cancer entails abnormal cell growth in the cervix, which is the second most prevalent cancer in women. Pap smears, Colposcopies, and HPV tests are a couple of screening techniques to aid in diagnosis. The subjectivity of cancer diagnosis underscores

the importance of pathologists' experience and training. The suggested method refines these models and utilizes diverse machine learning algorithms, where simple logistic regression achieves an accuracy peak at 95.14% when employing the pre-trained AlexNet model.

Li et al. [117] cervical cancer stands as the most widespread form of cancer among women worldwide, necessitates accurate diagnosis for optimal treatment decisions. This study aims to create and validate a CNN-based diagnostic system for identifying cervical malignancies in histology images. The Xception model demonstrated excellent performance, achieving AUC values of 0.98 (SCC) and 0.966 (AC) in internal validation. External validation results were 0.974 (SCC) and 0.958 (AC). Guided Backpropagation and Grad-CAM ensured interpretability by highlighting morphological features, suggesting CNNs as efficient tools for histological image classification in cervical cancer diagnosis.

Kang et al. [118] introduced a 1D hierarchical convolutional neural network (H-CNN), to integrate prior hierarchical classification knowledge with DL in Raman spectroscopy. Comparative evaluations against traditional methods such as decision tree classifier, Bayesian classifier, support vector machine classifier, and CNN. H-CNN's effectiveness in accurate tissue identification was demonstrated by its Macro-Accuracy of 94.91%, Macro-AUC of 97.35%, Macro-Recall of 95.31%, and Macro-F1 of 95.23%. This suggests that Raman spectroscopy may advance as a clinical diagnostic tool for cervical cancer diagnosis.

Sheela et al. [119] cervical cancer, a significant contributor to mortality rates, poses challenges in identification. Therefore, the authors presented three innovative approaches for automatically detecting and categorizing cervical cancer in Pap images. The presented method was evaluated on Herlev dataset. The model achieved an accuracy of 96% which proved superiority of the proposed method over existing intelligent techniques.

Attallah et al. [120] presented CerCanNet, a novel CAD model for automated cervical cancer classification. CerCanNet utilizes three streamlined CNNs with reduced parameters and less complex layers in comparison to established models such as MobileNet, DarkNet-19, and ResNet-18. Unlike usual methods, CerCan·Net uses the final three deep layers from each CNN, making it more straightforward and efficient. These features are merged and subjected to a feature selection process to create a compact set capable of distinguishing Pap smear image categories. The classification process involves three contexts: individual layer-based classification, ensemble-based classification, and feature-selected blended features classification. Context II demonstrates that merging features from CNNs with distinct structures enhances accuracy but increases feature space dimension. Context III validates that

the reduced feature set selected by CerCan·Net improves classification performance while reducing dimensionality. Overall, CerCan·Net proves to be a reliable CAD model, offering potential assistance to cytopathologists in cervical cancer diagnosis and addressing challenges in routine diagnosis.

Del et al. [121] focused on cell segmentation of nuclei and cytoplasm from Pap smear images based on Herlev dataset. Given the substantial impact of cervical cancer on public health in less developed regions, there is a heightened necessity for enhanced screening methods. Thus, leveraging advanced classification architectures like EfficientNet and Feature Pyramid Networks, the proposed method demonstrates robust performance. Validated through 5-fold cross-validation, the approach accurately identifies cervical cell lesions, offering potential enhancements in patient care. Table 1 provides a comprehensive overview of various deep learning approaches used in the automatic diagnosis of cervical cancer, highlighting their specific techniques, such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Autoencoders, and Generative Adversarial Networks (GANs). It also details the datasets employed, like the Herlev Dataset, ISBI Challenge Dataset, and National Cancer Institute (NCI) Dataset. The table presents key metrics for evaluating model performance, including accuracy, precision, recall, F1 score, and AUC-ROC. Accuracy measures overall prediction correctness, precision assesses the proportion of true positives, and recall evaluates the model's ability to detect actual positives. The F1 score balances precision and recall, while AUC-ROC measures the model's ability to distinguish between positive and negative cases. Additionally, the table addresses limitations such as data quality, overfitting, and computational complexity, which can affect the model's reliability and efficiency.

Table 1: Summary of the state-of-the-art approaches in cervical cancer diagnosis.

References	Type of learning	Dataset	DL Models	Performance	Limitations
[51]	Supervised learning	(Primary dataset) Real dataset of 1400 images	CNN	Accuracy=95% Precision=0.91430 Recall =0.87260.0008	Weak staining of abnormal nucleus affected the segmentation process.
[55]	Supervised learning	(Secondary dataset) UCI database which consists of 858 samples	Stacked auto-encoder	Accuracy =97.8%	Training time performance is worse compared to other approaches.
[65]	Supervised learning	(Primary dataset) Liquid-based histological	Mask RCNN	Accuracy= 89.8% Sensitivity =72.5%	Data imbalance

References	Type of learning	Dataset	DL Models	Performance	Limitations
		slides obtained from Thammasat University, Thailand		Specificity =94.3%.	
[66]	Unsupervised learning	(Primary dataset)	A CNN based model called Pap smear images convolutional neural network (PsiNet)	Accuracy=98%	The integrated optimization strategy makes it computationally expensive
[67]	Unsupervised learning	(Primary dataset) CT scans obtained from MD Anderson Cancer Center, USA.	(2D) DeepLabV3+Go ogl e and (3D) Unet	DSC=0.71to0.97	Reduced image quality, Poor performance.
[71]	Supervised learning	(Primary dataset)	Faster R-CNN	Accuracy =91%	A smaller number of CIN3 cases were considered.
[72]	Supervised learning	(Secondary dataset) DTU/HERLEV dataset	VGG, ResNet50 and inception V3	Accuracy=98.87%	Model is evaluated on single performance metric
[73]	Supervised learning	(Primary dataset)	YOLOv3	Sensitivity =97.5% Specificity=67.8% Mean Average Precision(mAP)=6 3.4%	Extra overhead due to integration of extra task specific classifier.
[74]	Unsupervised learning	(Primary and Secondary dataset) Primary data collected from Peking Union Medical College Hospital,China Herlev dataset	A CNN based model called Series-parallel fusion network	Average precision =78.4%	Data imbalance
[75]	Unsupervised learning	(Secondary dataset) Intel & Mobile ODT Dataset NCI Guanacaste Project Dataset provided by American	CNN	AUC=0.82	Data imbalance

References	Type of learning	Dataset	DL Models	Performance	Limitations
		National Cancer Institute			
[76]	Supervised learning	(Secondary dataset) COCO dataset obtained from Microsoft ImageNet Large Scale Visual Recognition Challenge	YOLO V3	Accuracy=86.45 %. Precision= 73.34 %	Small dataset
[77]	Unsupervised learning	(Secondary dataset) Herlev dataset	CNN, Variational autoencoder	Accuracy=99.4%	Minor distortion could lead to the wrong results
[78]	Supervised learning	(Primary and secondary datasets) Images obtained from Guangdong Provincial People's Hospital, Beijing,China Herlev dataset	YOLOv3	Accuracy=98.8, Sensitivity=96.7, Specificity=98.4	Weak generalization
[79]	Supervised learning	(Secondary dataset) Mendeley LBC dataset.	A computer-aided diagnostic model based on deep learning, PCA and SVM.	Accuracy=100%	It is restricted to classification task only.
[81]	Supervised learning	Secondary dataset	U-Net based CNN	DSC= 95% Hausdorff distance=95%	Limited to traditional geometric indices.
[83]	Supervised learning	(Primary dataset) Pap smear images collected from Ayursundram Healthcare Pvt. Ltd and Dr.B.Borooah Cancer Institute, Guwahati, Assam	CNN	Accuracy = 95%	Small dataset
[84]	Supervised learning	(Primary and secondary dataset) HEMLBC dataset obtained	CNN	Accuracy =98.3%, Area under the curve=0.99 Specificity=98.3%	Miss classification of few severe dysplasia and carcinoma cells.

References	Type of learning	Dataset	DL Models	Performance	Limitations
		from People's Hospital of Nanshan, China Herlev dataset			
[85]	Supervised learning	(Secondary dataset) Herlev dataset	CNN	Accuracy= 95.8%, Sensitivity =99.3% Specificity =99.1%	Data imbalance
[86]	Supervised learning	(Primary dataset)	VGG16	Accuracy=81%	Evaluated on single performance metric
[87]	Supervised learning	(Secondary dataset) Herlev dataset	DenseNet161, AlexNet, ResNet101, VGG19 and SqueezNet	Accuracy=94.38% Sensitivity=98.48 Specificity=82.61	High computational cost.
[88]	Supervised learning	(Secondary dataset) Mobile ODT dataset	RetinaNet, Inception V3, VGG	Accuracy=94% Recall=98% Precision=94% F1 score=96%	No hyperparameter information available.
[89]	Supervised learning	(Primary and secondary dataset) The primary dataset was collected from Babina Diagnostic Pvt. Ltd, Imphal Herlev dataset	AlexNet,VGGNet, ResNet, GoogleNet	Accuracy=0.989, Sensitivity=0.978, Specificity=0.979	Data imbalance
[90]	Unsupervised learning	(Secondary dataset) Herlev dataset	CNN and basic machine learning approaches	Accuracy=99%	Computationally expensive
[91]	Unsupervised learning	(Primary dataset)	A blended approach of CNN and cell merger	Sensitivity=0.92 Specificity=0.83	High Computation cost required
[92]	Supervised learning	(Secondary dataset) SIPaKMeD	Deep CNN	Accuracy=0.990 Precision=0.974 Recall= 0.974 F1-score= 0.974	Lack of similar cytology datasets.
[93]	Supervised learning	(Primary dataset) Images were obtained from Shanghai General Hospital, Shanghai First	Faster R-CNN	Sensitivity=99.4 Specificity=34.8% Area under the curve=0.67	Due to overlapping of cells it was difficult to different between normal and abnormal cells.

References	Type of learning	Dataset	DL Models	Performance	Limitations
		Maternity and Infant Hospital Shanghai.			
[95]	Unsupervised learning	(Primary and secondary dataset) Images were obtained fromGuangdong Province People's Hospital,China Herlev	CNN and SVM	Accuracy=99.3 Sensitivity=98.9 and Specificity=99.4	Based on two class classification
[96]	Supervised learning	(Primary dataset) Cervicography images obtained from Ewha Womans University Mokdong hospital, South Korea	XGB, RF, SVM ResNet50	Highest accuracy=97%(Re sNet50)	The cropped images contain vaginal walls which reduce the accuracy of the model.
[97]	Supervised learning	(Primary dataset) Cervical Cytology images obtained from Heilongjiang Maternal and Child Health Hospital, China Harbin Medical University Cancer Hospital, China.	A novel CNN based model called Attention feature pyramid network (AttFPN))	Sensitivity=91.30 %, specificity=90.62 %, Accuracy=90.91%, and AUC=0.934	Subclassification of abnormal cells is missing.
[98]	Unsupervised learning	(Primary and Secondary datasets) Healthy cervical cells were collected from Tripoli Hospital center, Libya Intel and Mobile ODT obtained from Kaggle.	ResNet	Accuracy=100% Sensitivity=97.8% Specificity=98.1% AUC=96.9%	Data Imbalance
[99]	Unsupervised learning	(Secondary dataset) Herlev dataset	A novel model based on GAN called Small-Object Detection	Accuracy= 97.08 Recall= 100 Precision= 95.62	The proposed architecture is not compared with state-of-the-art deep

References	Type of learning	Dataset	DL Models	Performance	Limitations
		colposcopy images	Generative Adversarial Networks (SOD-GAN)		learning-based approaches.
[100]	Supervised learning	(Primary dataset)	Ensemble learning and fully connected neural network	Accuracy=0.8656 IoU=0.5197 Specificity=0.6439 Sensitivity=0.9666 MCC=0.3892	Manual annotation of the dataset.
[101]	Supervised learning	(Secondary dataset) Herlev dataset	AlexNet GoogLeNet, ResNet, and DenseNet	Accuracy=94.5%	Model overfitting
[102]	Unsupervised learning	(Secondary dataset) Intel and Mobile-ODT	CNN (Hierarchical convolutional mixture of experts)	Accuracy=96.77%, kappa score=0.951	The Type of noise is not mentioned.
[103]	Supervised learning	(Secondary dataset) Herlev dataset	Mask RCNN and VGGNet	Accuracy=95.9%, Sensitivity=96.2% Specificity=99.3%, High h-mean of =97.7%	Higher processing power required.
[105]	Supervised learning	(Secondary dataset) Guanacaste dataset	Mask R-CNN	Precision=(0.92±0.06),Recall=(0.91± 0.05),and ZSI=(0.91±0.04)	Not compared with other state-of-art approaches.
[107]	Supervised learning	(Primary dataset)	2D U-Net 3D U-Net	DSC<0.65 HD=2.7to9.6mm	Only one contour extractor was utilized to minimize the impact of inter-contour variation.
[108]	Supervised learning	(Secondary dataset) Herlev dataset	CNN UNET	Accuracy=0.91 Sensitivity= 0.87 Specificity= 0.91 F1 score=0.89	Extensive pre-processing is required.
[109]	Unsupervised learning	(Primary dataset)	CNN and Ensemble learning	Accuracy=90.4% Accuracy=91.6%	Cells are corrupted with debris
[110]	Supervised learning	(Primary dataset) Liquid based slides obtained Fernando Fonseca (HFF), Portugal	A ResNet based model called RetinaNet	Precision=0.20, recall=0.40 F1score= 0.27	The network faces challenges in accurately recognizing the annotated areas with abnormalities.
[111]	Supervised learning	Secondary datastet	SIPAKMED Dataset	Accuracy =9.95% F1-score =90.74%	It is limited to secondary dataset.

References	Type of learning	Dataset	DL Models	Performance	Limitations
[113]	Supervised learning	Secondary dataset	ResNet-34, ResNet-50 and VGG-19	Accuracy= 97.65%	The model is evaluated on single performance metric.
[116]	Supervised learning	Secondary dataset	Alexnet, Resnet-18, Resnet-50 and Googlenet	Accuracy=95.14%	The model is evaluated on single performance metric.
[118]	Supervised learning	Primary dataset (233 real images)	Hierarchical convolutional neural network	F1= 95.23%, AUC= 97.35%	The model is computationally intensive.
[119]	Supervised learning	(Secondary dataset) Herlev Pap smear data	Anisotropic Diffusion Filter, Advance Map-Based Superpixel Segmentation, Auto encoder-based Extreme Learning	Accuracy=96.0%	Extensive Pre-processing is required.
[121]	Supervised learning	(Secondary dataset) Herlev dataset	EfficientNet	0.91 F1 score, 0.85 IoU, 0.91 Precision, 0.92 Recall, and 0.96 Specificity.	It is limited to secondary dataset.
[122]	Supervised learning	(Primary dataset) Data collected from Fourth central hospital of Baoding city, China	CNN	Area under the curve=0.997	Limited dataset

2.4 SUMMARY

The development of successful screening systems can contribute to the decline of cervical cancer. While the Pap smear is a widely used technique for precancerous diagnosis, human analysis of these smears is prone to errors due to its difficulty and time-consuming nature. Hence, creating a computer-assisted diagnosis tool can improve precision and speed in the diagnostic process. This chapter explores recent deep learning techniques for diagnosing and categorizing cervical cancer, aiming to minimize human errors and enhance diagnostic efficiency. The research identifies specific challenges such as the lack of annotated public datasets, high computation costs, huge training time, and the need for ground truth images in supervised approaches. The study emphasizes the importance of classification, segmentation,

and feature extraction approaches in cervical cancer analysis. The classification problem may involve multiple classes depending on the Bethesda system of classification. While 99 percent of algorithms achieve over 90 percent accuracy in classifying normal and abnormal cells, none have achieved 100 percent results. These techniques vary in parameters, model size, architecture complexity, dataset size, labeled images, and total training and testing time. Commonly used techniques include CNN, VGGNet, ResNet, GoogleNet, AlexNet, GAN, and YOLO V3 for classification and feature extraction, with CNN, ResNet, and VGGNet being the most widely used. Researchers have frequently utilized Mask-RCNN and UNET for segmenting the nucleus and cytoplasm in cervical cancer analysis. Despite the superior performance of deep learning approaches in this aspect, Support Vector Machines (SVM) remain commonly employed as a classifier, followed by Multilayer Perceptron (MLP), decision trees, and Random Forest. This comprehensive survey delves into crucial aspects of cervical cancer analysis, including dataset descriptions, the percentage of dataset usage over the past nine years, and performance metrics. By addressing these key areas of interest, the survey serves as a valuable resource for authors and researchers, offering insights into current trends in automatic cervical cancer analysis, dataset selection, and technique implementation for enhanced diagnosis. Its findings provide guidance for future research endeavors, facilitating a deeper understanding of the evolving landscape in cervical cancer diagnosis and paving the way for improved methodologies and outcomes.

CHAPTER 3
METHODOLOGY

This chapter presents the approach and methods employed to achieve the objectives outlined in the preceding chapters. The process consists of data collection, analysis, pre-processing, and the deep learning model utilized for cervical cancer diagnosis. Initiating with data collection, the dataset was sourced from the Sheri-Kashmir Institute of Medical Science Soura (SKIMS) in Jammu and Kashmir. A total of 109 Pap smear slides were obtained, yet only 90 were deemed suitable for image preparation due to over-staining issues, rendering 20 slides unsuitable for experimentation. The gathered data underwent meticulous analysis and pre-processing (Contrast Limited Histogram Equalization and Resizing) to ensure quality and relevance for the subsequent phases. Furthermore, a specialized deep learning models such as Standard U-Net, Residual U-Net, Attention U-Net and a Hybrid model of Residual U-Net and Attention U-Net (DeepSegNet) was employed to facilitate cervical cancer diagnosis, leveraging the refined dataset for accurate and precise evaluations.

3.1 Proposed Methodology

The proposed methodology comprises three integral sub-modules aimed at comprehensive analysis and classification of Pap smear images. The initial phase involves meticulous data collection and preparation, forming the foundational first module. Through rigorous acquisition from diverse sources, a robust dataset of Pap smear images is assembled, ensuring representation across various conditions and abnormalities. This data undergoes thorough preparation, encompassing cleaning, labeling, and annotation to guarantee accuracy and relevance for subsequent analysis. Following the data preparation, the second module focuses on pre-processing the Pap smear images. Techniques like cropping, resizing, and normalization are employed to standardize the images, ensuring uniformity in size, resolution, and quality. This crucial step optimizes the images for subsequent stages of analysis and facilitates consistent interpretation by subsequent algorithms. The final module constitutes the core of the methodology, revolving around segmentation, feature extraction, and classification. Utilizing advanced U-net architectures, the images are segmented to delineate the cytoplasm and nucleus boundaries accurately. This segmentation lays the groundwork for subsequent analysis and classification. Feature extraction is then performed using either the VGG16 model or Cell Profiler, aiming to capture nuanced characteristics and patterns within the segmented regions.

The subsequent step involves employing classification methodologies. Here, the extracted features serve as inputs for classification models. Two distinct classifiers are evaluated for their

performance: XGBoost and Artificial Neural Networks (ANN). Comparative analysis reveals that ANN outperforms XGBoost in the challenging task of 7-class classification, showcasing superior accuracy and robustness in distinguishing between various cell types and abnormalities. This comprehensive approach not only accurately identifies cellular structures but also facilitates the differentiation of diverse cell categories crucial in diagnosing various conditions. The integrated approach of these modules signifies a systematic and comprehensive strategy for the analysis of Pap smear images. From meticulous data collection and preparation to precise segmentation, feature extraction, and ultimately classification, the methodology aims to enhance the accuracy and efficiency of cell classification in Pap smear images, potentially revolutionizing diagnostic processes in cytology. The significance lies not only in the individual components but in their cohesive integration, ensuring a comprehensive and accurate analysis of Pap smear images for improved medical diagnoses. The proposed methodology for automated cervical cancer diagnosis, as illustrated in Figure 6, encompasses a series of critical procedures aimed at effectively analyzing Pap smear images. Initially, the dataset is meticulously labeled to enable seven-class classification, ensuring that each image is appropriately categorized according to the specific type or stage of cervical cancer it represents. Following this, ground truth images of cytoplasm and nucleus are meticulously generated utilizing the label studio tool, serving as reference standards for subsequent analyses. Upon preprocessing the collected Pap smear images, which involves essential steps like cropping, resizing, and normalization to standardize their attributes, attention is devoted to cases where cytoplasm appears folded or fragmented. In such instances, an additional preprocessing step is undertaken, involving the careful rotation of the region of interest to reestablish its connection with the original image, thereby facilitating accurate feature extraction. Subsequently, the segmentation process is initiated, aimed at precisely delineating the cytoplasm and nucleus regions of both normal and abnormal cells within the images. This segmentation step is crucial as it lays the foundation for subsequent feature extraction. By extracting pertinent features from the segmented cytoplasm and nucleus regions, a rich set of attributes is derived, which serves as the basis for subsequent analysis. Leveraging these extracted features, an automatic classification framework is then deployed to categorize cervical cancer instances into one of the predefined seven classes. Overall, this comprehensive methodology integrates meticulous data preprocessing, precise segmentation, feature extraction and sophisticated classification techniques to deliver a robust and reliable automated cervical cancer diagnostic solution.

The novelty of this work lies in the creation of a Deep-SegNet architecture (semantic segmentation model), which integrates residual blocks and attention blocks. This model excels in segmenting complex cell structures, maintaining cell boundary integrity more effectively than existing models. Besides its precise segmentation capabilities for cell structures, Deep-SegNet can also be applied to segment other complex structures in satellite images. The different types of layers used are convolutional layers, Max pooling layers, Up-sampling layers, Attention gates, and Gating signals. Each convolutional block contains two convolutional layers. Max pooling layers are used for down-sampling in the encoder. The up-sampling layers are used for up-sampling in the decoder. Attention blocks integrate attention mechanism. Gating signal layers generates gating signals for attention blocks. The ReLU and sigmoid activation functions are used. Apart from this the main components of the model are encoders, decoders, attention gates, skip connections, and residual blocks. The total number of encoders used is 5, with 4 decoders, 4 attention gates, 4 skip connections, and 9 residual blocks. Each encoder and decoder block uses residual blocks, resulting in a total of 9 residual blocks (5 in the encoders and 4 in the decoders). One of the interesting aspects of Deep-SegNet is its ability to preserve the cytoplasmic and nuclear boundaries for proper feature extraction from the cytoplasm and nucleus. Additionally, the development of the Deep-CervixNet model, which restructures folded structures, represents a significant breakthrough not previously achieved in this field. Deep-CervixNet is composed of three primary modules. The first module is responsible for data collection, followed by data preparation. The next module handles the preprocessing of Pap smear images. The last module comprises three sub-components: the first one performs segmentation of the cytoplasm and nucleus, including folded cells; then, features are extracted from the cytoplasm and nucleus, which are provided as input to an artificial neural network to classify cells as cancerous or non-cancerous according to the Bethesda system (TBS) of classification. The Bethesda system is a standardized system which is used to report the cervical cytology diagnosis. This system was initially introduced in 1988 and subsequently revised in 1991, 2001, and 2014. For cervical cytology, TBS categorizes findings into several groups, including specimen adequacy, general categorization (such as negative for intraepithelial lesion or malignancy), and epithelial cell abnormalities, which range from atypical squamous cells to squamous cell carcinoma.
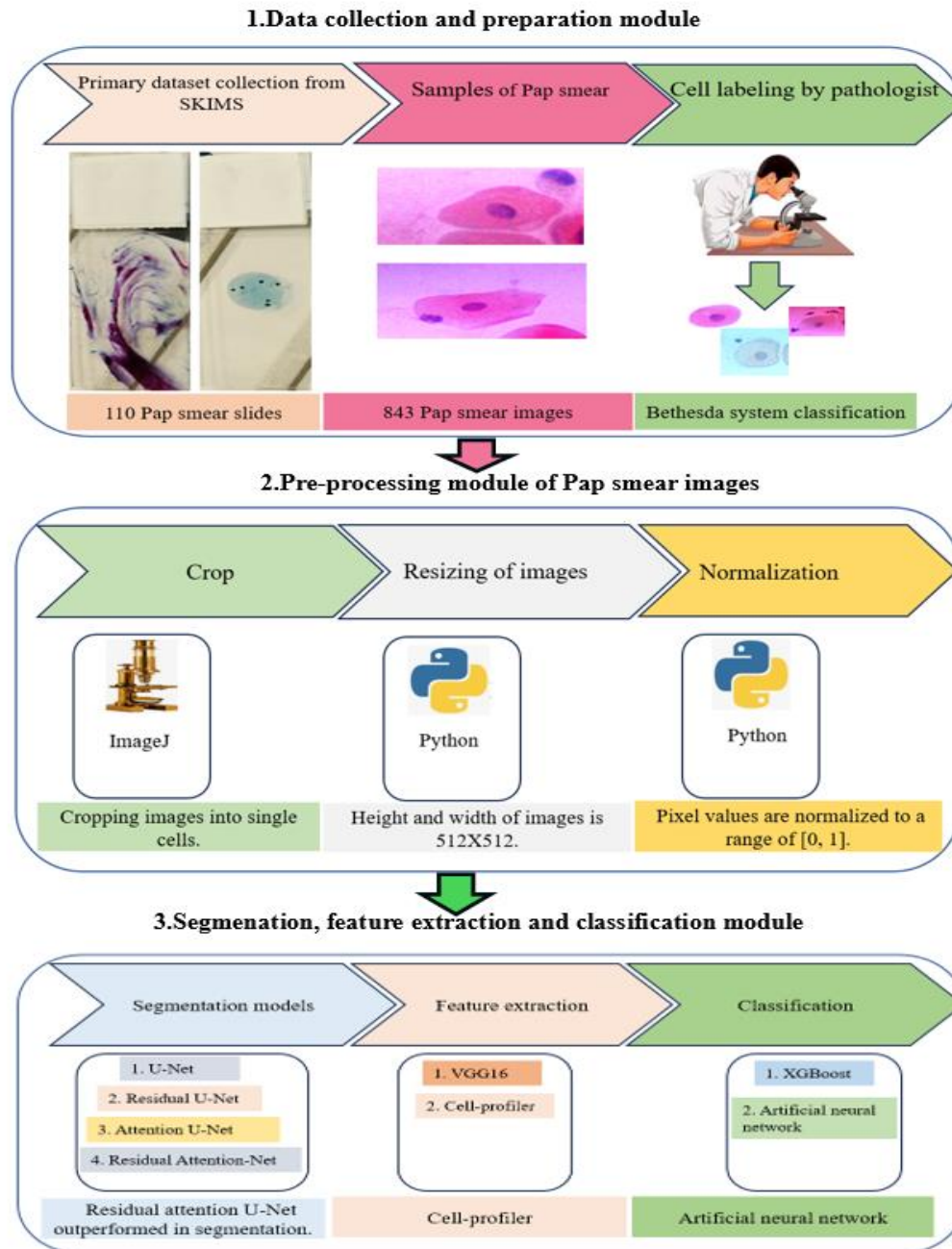
Figure 6: Proposed methodology for diagnosis of cervical cancer.

3.2 Dataset for Study

The success of a Machine learning (ML) techniques heavily depends on the quality of the training dataset. It's believed that with a suitably precise database, a machine can learn and replicate human behaviour. Cervical cancer screening using machine learning algorithms heavily relies on the quality and comprehensiveness of the training dataset. The choice of dataset significantly impacts the performance and accuracy of the developed AI models. In this

section, the information of the dataset used in the research study are presented. The dataset employed in this study follows the 2001-Bethesda system of classification, which provides standardized criteria for reporting cervical cell abnormalities. It consists of 843 cells that represent the data from approximately 109 clinical cases. To compile this dataset, slides of pap-smear tests were collected from medical healthcare institutions in northern India, specifically from the Sheri Kashmir Institute of Medical Sciences, Soura (SKIMS), Kashmir. Care was taken to ensure patient confidentiality and adherence to medical ethical guidelines throughout the data collection process. The slides were examined under a multi-headed microscope of Labomed (LX 300 Microscope) equipped with a digital camera. Images of the slides were captured at 100X magnification to maintain uniformity. Pre-processing techniques were applied to isolate individual cells from cell clusters and enhance their brightness, color, and contrast using Contrast Limited Adaptive Histogram Equalization for improved cytoplasmic and nuclear feature recognition. Two trained Cyto-pathologists inspected and cross-validated the diagnoses, ensuring the accuracy and reliability of the database. Samples with conflicting diagnoses were excluded from the dataset. Out of a total of 843 images, 487 are of the normal category, whereas 356 are of the abnormal category. Each tuple contains 40 attributes and is uniquely identified by a primary key. Out of these attributes, 19 describe cytoplasmic features, 19 describe nuclear features, one records the nuclear-cytoplasmic area ratio, and the last attribute identifies the cell's class. In consultation with medical experts, 39 morphological features (19 each from cytoplasm and nucleus) were identified for cell profiling, extracted using the open-source software Cell Profiler. This software, developed by the BROAD Institute and MIT, allows quantitative measurements of medical image phenotypes. This software allows for the creation of tailored pipelines to quantitatively measure the features of biological objects in input images. The process begins with importing Pap smear images depicting cytoplasm and corresponding nucleus regions. Cell Profiler's object processing module identifies the primary objects, which are the cytoplasm and nucleus. Once identified, advanced algorithms in Cell Profiler measure and extract relevant features from both cytoplasm and nucleus regions, providing valuable insights into cellular properties, including size, shape, and textural characteristics. Upon completing the analysis of cells using the cell profiling tool, the resulting data was organized in a spreadsheet format. All database properties were integrated to facilitate seamless correlation between digital data and respective clinical cases. A comprehensive description of the extracted features is provided in the subsequent section.

Table 2 highlights a summary of the various categories of Pap smear images and their respective quantities. The dataset contains seven classes, with the initial four representing healthy cells and the latter three indicating unhealthy cells. Therefore, the screening task can be approached as a binary classification problem (healthy vs. unhealthy) or as a more detailed classification into seven distinct categories.

Table 2: Total number of Pap smear images in the prepared dataset.

Class	Category	Cell Type	Number of cell images	Sub total
I	Normal	Superficial Squamous	172	487
II		Intermediate Squamous	125	
III		Parabasal Squamous	98	
IV		Basal Squamous	92	
V	Abnormal	LSIL	170	356
VI		HSIL	146	
VII		SSC	40	

The dataset employed in this research study is a valuable resource for training and evaluating machine learning algorithms for cervical cancer screening based on AI approaches. The dataset is characterized by its comprehensiveness, adherence to standardized classification criteria, and the valuable input of skilled cyto-pathologists, guaranteeing its precision and trustworthiness. The dataset's efficiency has been verified through the successful execution of diverse machine learning algorithms in the screening procedure. The cells presented in the table exhibit variations in several characteristics, including cytoplasmic-nucleus ratio, solidity, eccentricity, and more. Class I consists of 172 normal Superficial Squamous cells, while Class II comprises 125 Intermediate cells. Class III consists of 98 Parabasal Squamous cells, followed by Class IV with 92 Basal Squamous cells. Class V contains 170 LSIL cells, and Class VI consists of 146 HSIL cells. Finally, Class VII comprises a limited pair of cells, totalling 40, known as Squamous cell carcinoma (SSC). Figure 7 illustrates the dataset preparation for model training, comprising a total of 843 images.
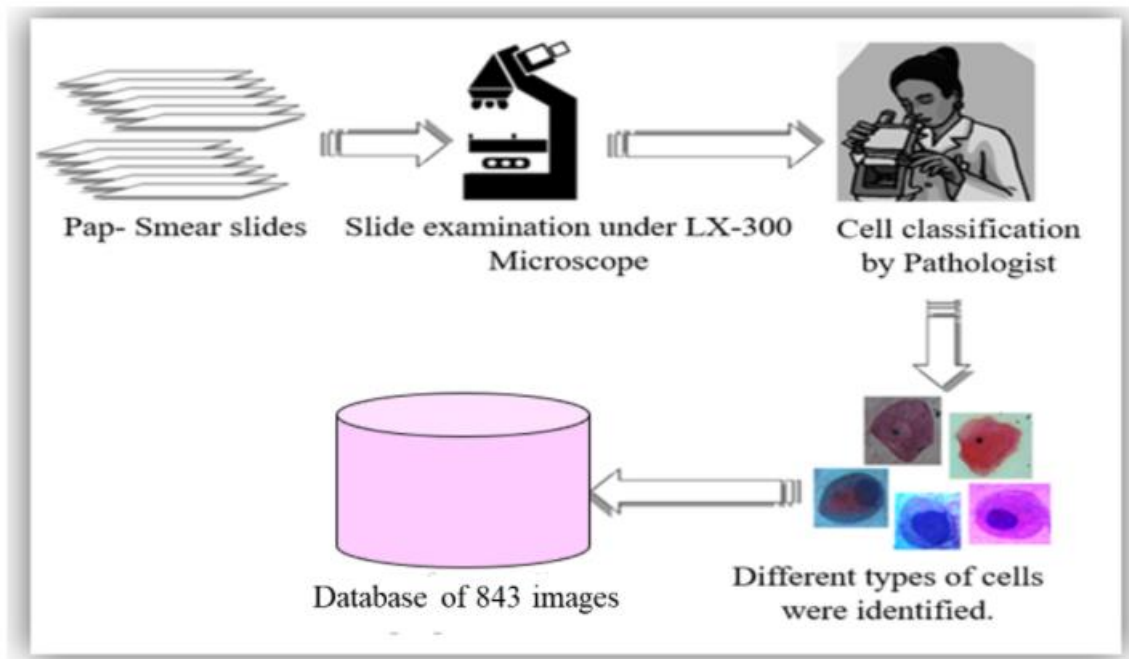
Figure 7: Dataset preparation with 843 images for model training.

3.2.1 Description of the Collected Cells

3.2.1.1 Superficial Squamous

The outermost layer of non-keratinized epithelium consists of flat squamous cells that appear polygonal and lack visible nuclear details due to nuclear degeneration. The late proliferative and ovulatory phases of the menstrual cycle are characterized by increased concentrations of these superficial squamous cells.

3.2.1.2 Intermediate Squamous

The polygonal cells range in size from 1256 to 1618 µm [40]. They have vesicular nuclei that are noticeably bigger. When a cervix is estrogen-influenced, the cells in the superficial and intermediate layers have more cytoplasm as they get older and have more glycogen in them.

3.2.1.3 Parabasal Squamous

The lowermost layer of the squamous epithelium constitutes this type of cells, commonly seen in individuals with low estrogen levels, like those in premenstrual, postpartum, or post-menopausal stages or using estrogen-inhibiting medications. These cells have a compact, even basophilic cytoplasm enveloping a 50 µm nucleus [41].

3.2.1.4 Basal Squamous

These cells reside in the deepest layer of the skin, called the epidermis, which forms the skin's outermost surface. When these cells undergo malignancy, they can result in a skin cancer called

46

basal cell carcinoma. This is the most common kind of skin cancer, and it typically manifests as a tiny, glossy bump. Although basal cell carcinoma generally grows slowly and seldom metastasizes to other body areas, untreated cases can inflict considerable harm on the surrounding tissue.

3.2.1.5 Low-Grade Squamous Intraepithelial Lesion

An LSIL denotes an area containing abnormal cells that have an increased amount of cytoplasm. The nuclear membrane exhibits notable anomalies in LSIL, and the nucleus is three times larger than in typical intermediate cells.

3.2.1.6 High-Grade Squamous Intraepithelial Lesion

An HPV-related anomaly in squamous cells is known as an HSIL. HSIL is recognized as a precancerous disorder that is treatable, even though not all cases develop into malignancy. There are three grades for HSILs: CIN 1, CIN 2, and CIN 3. CIN 1 is a lower-grade lesion that affects about one-third of the epithelial cells and has a decreased chance of developing into cancer. Moderate dysplasia, a higher-grade lesion affecting the upper two-thirds of the epithelium, is seen in CIN 2. Severe dysplasia affecting the top two-thirds of the epithelium is seen in CIN 3. The different cell types from the prepared dataset are shown in Figure 8.

3.2.1.7 Squamous Cell Carcinoma

Squamous cell carcinoma is one of the types of cancer. It occurs when the squamous cells that line the surface of the cervix become cancerous. This type of carcinoma is responsible, for several cancer cases worldwide. Squamous cell carcinoma usually develops because of HPV infection, especially from high-risk strains, like HPV 16 and HPV 18. Doctors often identify it by examining tissue samples under a microscope, which reveal cellular characteristics that indicate this cancer. SCC is characterized by large, irregularly shaped cells with hyperchromatic nuclei and an increased nuclear-to-cytoplasmic ratio. These cells often exhibit keratinization, appearing orange or yellow due to keratin production. The presence of SCC in Pap smear samples is highly significant, as it indicates invasive cervical cancer, necessitating prompt medical intervention. Early detection through cytological screening is crucial for improving patient outcomes. Squamous cell carcinoma can cause local tissue invasion, leading to severe damage and potential metastasis to other organs if untreated. It poses a significant risk of morbidity and mortality, especially in advanced stages of the disease.
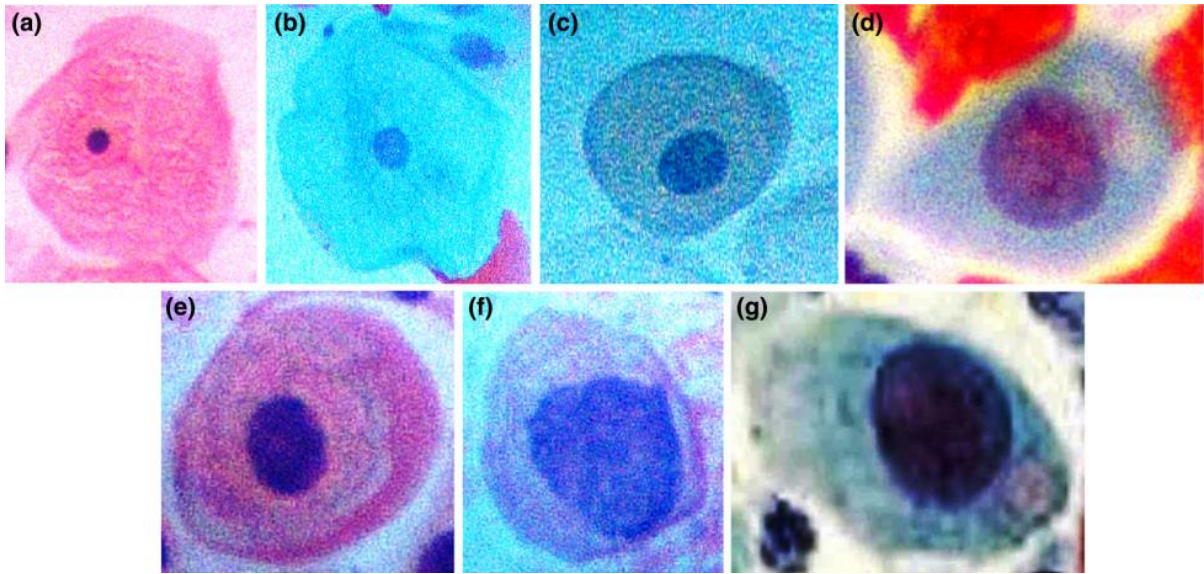
Figure 8: Cell samples a) Superficial squamous cell, b) intermediate squamous cells, c) parabasal squamous cell, d) basal squamous cell, e) low-grade squamous intraepithelial lesion, f) high-grade squamous intraepithelial lesion, g) squamous cell carcinoma.

3.3 Pre-Processing Methods

Data pre-processing means refining and converting unprocessed data into a processed format. This initial phase is crucial for data readiness before model construction. To maintain the balance in the data all 843 images were resized to a common size of 128×128 pixels. This resizing helps streamline model training and ensures consistent input data. Maintaining uniform image sizes eliminates potential variations that could hinder model performance. The following sections will elaborate on methodology, experimental outcomes, and analysis from Pap smear image segmentation.

3.3.1 Resizing

The resizing process of 843 images for input into the segmentation network was executed meticulously and effectively. Each image was carefully adjusted to ensure compatibility with the network's input requirements, guaranteeing optimal performance and accuracy during the segmentation task. By maintaining the integrity of the images through proper resizing techniques, potential distortions or loss of critical information were mitigated, preserving the quality of the dataset. This meticulous approach underscores the commitment to robust data preprocessing, laying a solid foundation for the subsequent segmentation analysis.

3.3.2 Contrast Limited Histogram Equalization

Contrast Limited Adaptive Histogram Equalization (CLAHE) stands as a pivotal image processing technique renowned for its ability to enhance image contrast while mitigating the risk of amplifying noise or artifacts. Unlike traditional Histogram Equalization, CLAHE operates on a localized scale, dividing images into smaller regions, or tiles, to independently adjust their contrast levels. By doing so, it alleviates the uniform contrast adjustment across the entire image, catering to specific regions with varying luminance characteristics. Crucially, CLAHE incorporates a contrast limitation mechanism, preventing excessive amplification of contrast within individual tiles. This constraint ensures that the enhancement process retains natural image features while significantly boosting the visibility of details in both dark and bright areas. Widely utilized in medical imaging, satellite imagery, and diverse fields requiring precise contrast enhancement without compromising image integrity, CLAHE continues to be a cornerstone technique in digital image processing workflows.

3.4 Software Used

3.4.1-Label-Studio

Label Studio is a data labeling tool available as open-source software, facilitates the annotation of diverse data formats such as audio, text, images, videos, and time series via an intuitive user interface. Its functionality extends to exporting labeled data into multiple model formats. It serves as a versatile solution for both refining existing training data and preparing raw data, culminating in the enhancement of machine learning models' accuracy and performance. The software operates under the Apache 2.0 LICENSE © Heartex, covering the period from 2020 to 2022. The role of Label Studio extends beyond its annotation tools; it serves as a pivotal platform for training machine learning models. The meticulously labeled datasets generated within Label Studio serve as invaluable training data, enabling AI based approaches to learn and recognize cytoplasm and nucleus within images with remarkable accuracy. This tool can be downloaded from labelstudio.com. Since the proposed work is based on supervised learning, so the ground truth images were required to generate the individual masks for cytoplasm and nucleus almost 1686 masks were generated for both cytoplasm and nucleus for the proper model training. Figure 9 is an array of labeling interfaces that enable the creation of precise ground truth images, tailored to specific annotation requirements across tasks like object detection, segmentation, and key point identification. The label values for a give class must be of unique or binary (0 /1 or 2/255). The various steps to install the label include:

- Pip install-u label-studio

- Conda env list
- Create a separate environment and activate the environment using below mention command
- Conda activate truth

The below mentioned images a, b, c, d are the various interfaces for generating ground truth. Among these images semantic segmentation mask is used to generate the ground truth for cytoplasm and nucleus. The semantic segmentation mask generating interface is depicted in 9a.

Figure 9: Different types of labeling interfaces for generating ground truth images.

The below mentioned images represents the parabasal image and the process of generating the mask for nucleus and cytoplasm respectively. From the visual representation, it is evident that if three classes are to be generated, then three sets of labels must be created. However, the segmentation process that is carried out in this research is based on two classes cytoplasm and nucleus. Thus, two labels are required to generate the cytoplasm and nucleus of Pap smear images. The first image acts as an input to the label studio then after activating the project cytoplasm and nucleus masks are generated. Figure 10 mentioned below represents how the ground truth images are generated using label-studio for nucleus and cytoplasm of the respective cells. The first image in Figure 10 serves as the input image, followed by the second image, which represents the ground truth image for the nucleus. Subsequently, cytoplasmic labeling of the same cell is depicted in the third image.

Figure 10: Labeling of nucleus and cytoplasm in parabasal cells: a) Input Image b) Nucleus labeling c) Cytoplasm labeling.

3.4.2 Image J

ImageJ is an image processing software that is enriched with abundant functions. To download ImageJ, users can access the official website or repository where they'll find versions compatible with different operating systems. Once downloaded, the software offers an intuitive interface with various plugins and functionalities tailored for tasks such as particle analysis, cell counting, and image segmentation. For academic purposes or publications, referencing ImageJ typically involves citing the software version, the NIH as the developer, and providing the appropriate publication reference or DOI (Digital Object Identifier) for the specific ImageJ version used. ImageJ is a comprehensive software capable of handling, modifying, examining, processing, storing, and printing images in various formats. Notably, ImageJ supports image stacks, where a sequence of images shares a single window, and its multithreaded nature enables concurrent execution of time-intensive operations on multi-CPU hardware. Offering versatile analysis capabilities, ImageJ computes area and pixel value statistics for user-defined selections and objects based on intensity thresholds. It facilitates measurements of distances, angles, creation of density histograms, and line profile plots. Its capacity to handle multiple images simultaneously is constrained solely by available system memory. The below mentioned figure 11 belongs to the intermediate squamous cells which is imported into the ImageJ software to display the histogram of image, measurement of an image and edges of an image for further analysis. Similarly, there are other relevant operations like cell count, filters, enhance contrast and many more. Moreover, this tool is beneficial for visualizing 3D images, performing 3D measurements, tracking objects in 3D space over time, and integrating with other scientific software and databases for streamlined workflows.

51
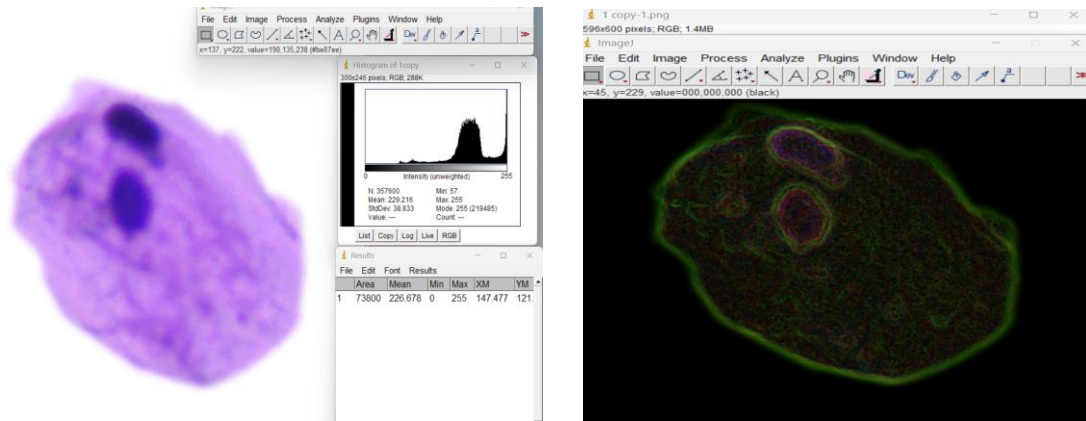
Figure 11: Edge detection and area calculation with ImageJ.

3.4.3 Fiji Plugin

Fiji, an acronym for "Fiji Is Just ImageJ," presents a comprehensive distribution of ImageJ2 bundled with a rich assortment of plugins and enhancements for enhanced functionality. Often described as "ImageJ with Batteries. Fiji provides user-friendly installation packages designed for different operating systems. These packages come pre-equipped with a JRE optimized for Fiji's use, although bare distributions are also available for those preferring to use their existing JRE. Being an Open-Source project, Fiji's entire source code is accessible for download. The installation process for Fiji is uncomplicated as it operates as a portable application, allowing it to function independently from its directory. Installation merely involves downloading and unpacking the distribution Fiji doesn't necessitate an installer, modify system directories, or store configurations in system databases like the Windows registry. This portability facilitates easy relocation by simply moving or copying the directory tree.

Upon launching Fiji, users encounter the main window, which houses the Menu bar, Tools bar, and a Status bar for system notifications and updates. The updater, a standout feature of Fiji, triggers upon launch or via the Help menu. This tool scans the local installation, checking for updates or new features in the global Fiji repositories. Users can easily fetch and apply the latest updates and patches to their Fiji installation, ensuring access to the most recent enhancements and bug fixes.

3.4.4 Cell Profiler

Cell Profiler is a freely accessible modular software specifically designed for image analysis, capable of managing an extensive volume of images. The software comes equipped with pre-existing methods tailored for various cell types and assays. Moreover, it serves as an open-

source platform, fostering collaboration among image analysis experts for the creation, testing, and enhancement of new analysis techniques. Addressing diverse needs, CellProfiler incorporates advanced algorithms adept at accurately identifying crowded cells and non-mammalian cell types. Its modular and adaptable framework allows for the analysis of novel assays and phenotypes. Being open source, the software grants visibility into its underlying methodology, enabling modifications and improvements by the community. Additionally, Cell Profiler offers a user-friendly interface, leverages computer clusters when available, and streamlines the often laborious and non-transferable image analysis steps, such as image formatting and combining multiple analysis steps. Although primarily optimized for two-dimensional high-content screening images, Cell Profiler's support for time-lapse and three-dimensional image stack analysis is limited; however, researchers keen on these areas can potentially develop compatible modules for such analyses. The configurations of the pipeline's modules, along with their settings, are preserved and can be utilized to replicate the analysis or shared among collaborators. The Cell Profiler website [125] offers numerous sample pipelines, serving as helpful templates for initiating new analyses. Figure 12 illustrates the automated process of distinguishing and labeling individual cells within images. Cell Profiler utilizes image segmentation algorithms to accurately identify and outline cellular structures, assigning unique labels to each object. This precise identification allows for subsequent quantitative analysis of various parameters, such as size and intensity, thereby enhancing the accuracy and efficiency of cellular analysis in the context of cervical cancer diagnosis. Moreover, it supports various segmentation algorithms, including thresholding, watershed segmentation, and morphological operations, for accurately identifying cellular structures.
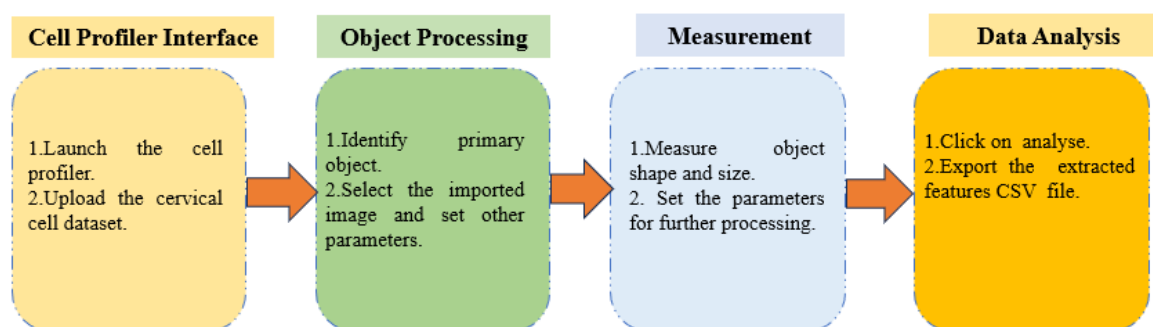
Figure 12: Cell Profiler pipeline for extracting cytoplasm and nucleus features.

# of accepted objects	1
10th pctile diameter	825.0 pixel:
Median diameter	825.0 pixel:
90th pctile diameter	825.0 pixel:
Area covered by objects	41.3 %
Thresholding filter size	1.0
Threshold	0.101
Declumping smoothing filter size	6.7
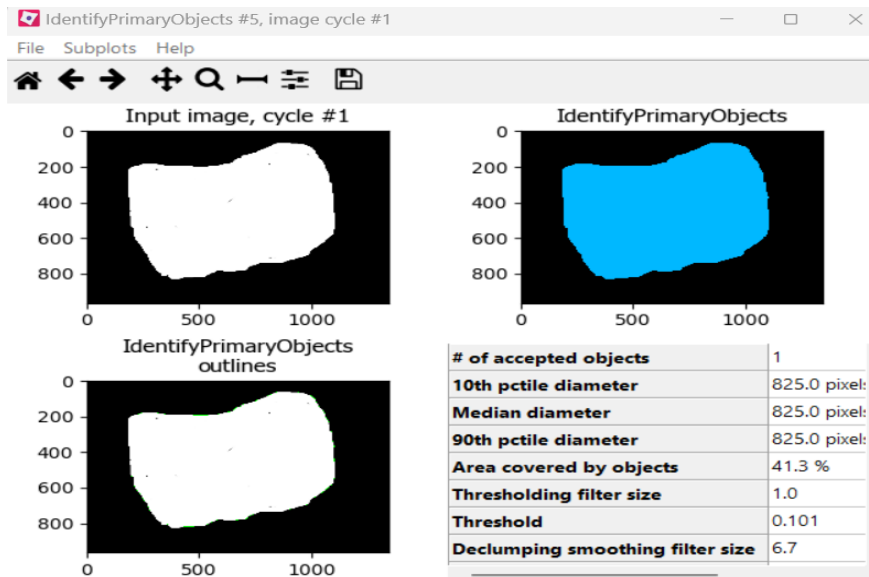
Figure 13: Object identification of intermediate cells for feature extraction.

3. 5 Types of Segmentation

Image segmentation is a widely used approach in digital image processing and analysis to divide an image into several sections or parts, usually according to the properties of the picture's pixels. Foreground and background separation as well as pixel clustering based on color or form similarity are two examples of image segmentation techniques. To identify and label individual pixels in an image or voxels in a 3D volume that reflect a tumor in a patient's brain or other organs, for instance, is a common use of image segmentation in medical imaging. Figure 14 displays segmentation methods for defining the region of interest (ROI). Semantic segmentation classifies pixels into categories, such as tumors. Instance segmentation identifies individual objects, allowing for multiple tumor recognition. Panoptic segmentation merges both methods, labeling all pixels while distinguishing instances, thereby enhancing detail and accuracy in medical image analysis.
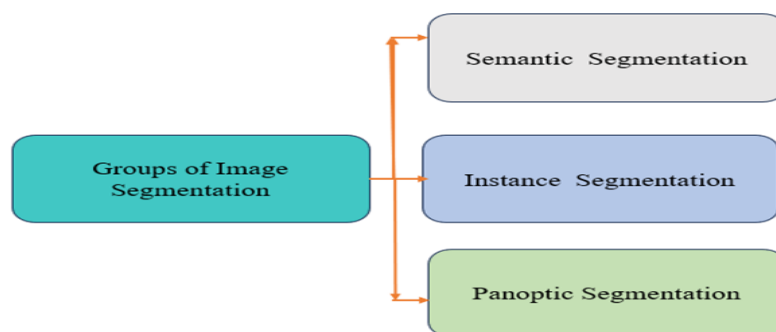
Figure14: Segmentation methods for obtaining region of interest.

3.5.1 Semantic Segmentation

Semantic segmentation means labeling each pixel in a picture with a specific category, dividing the image into separate parts, with each part representing something meaningful in the picture, several methods in semantic segmentation have succeeded through the application of deep neural networks. By providing abundant images with their pixel-level labels for training, these networks learn how semantic labels relate to diverse visual features. This learning process progressively narrows the gap between high-level meanings and detailed visual traits, enhancing the network's understanding of various semantic concepts. The various deep learning architectures used in semantic segmentation are U-Net, VGG (Visual Geometry Group), variants of VGG network are VGG-13, VGG-16, and VGG-19. ResNet, DenseNet, ResNeX, MobileNet are other architectures used in semantic segmentation.

3.5.2 Instance Segmentation

A particular form of image segmentation called instance segmentation focuses on identifying and distinguishing between each individual instance of an object that appears in an image. In addition to detecting all instances of a class, instance segmentation also can distinguish between distinct instances of each segment class. The various architectures based on instance segmentation are RCNN, Fast RCNN, Multipath Network, Faster RCNN, Mask RCNN, Non-Local Neural Networks, (Path Aggregation Network) PANet, YOLACT, Mask Scoring RCNN, and Tensor Mask.

3.5.3 Panoptic Segmentation

Panoptic segmentation combines instance and semantic segmentation to distinguish between elements in a scene, categorizing them into two groups: stuff and things. Stuff encompasses uncountable areas like the sky, pavements, and grounds. The various techniques that fall under panoptic segmentation are Unified Panoptic Segmentation Network (UPSNet), (FPSNet), Efficient Spatial Pyramid of Dilated Convolutions (EPSNet), (VPSNet).

3.6 Segmentation Models

The segmentation models that are used to segment the cytoplasm from nucleus include standard U-Net, residual U-Net, attention U-Net, and residual attention U-Net. The detailed description of the architectures is presented in the preceding sub-sections.

3.6.1 Standard U-Net

Olaf Ronneberger presented the U-Net architecture in 2015 [25]. It is a CNN with an interconnected encoding path and decoding path. The input data is compressed by a series of

encoders in the encoding path, which is on the left. The decoding path, shown on the right, is made up of decoders that recover and extend spatial information. To keep input-output sizes balanced, pixel-level categorization is carried out at each level. Bottleneck layers are then utilized to create the final feature maps. Important components of the U-Net architecture include skip connections, which create direct links between layers to provide inputs, ReLU acting as the nonlinear activation function, and max-pooling layers, which reduce dimensions by choosing the maximum value in a defined region. Figure 15 illustrated below represents the U-Net architecture.
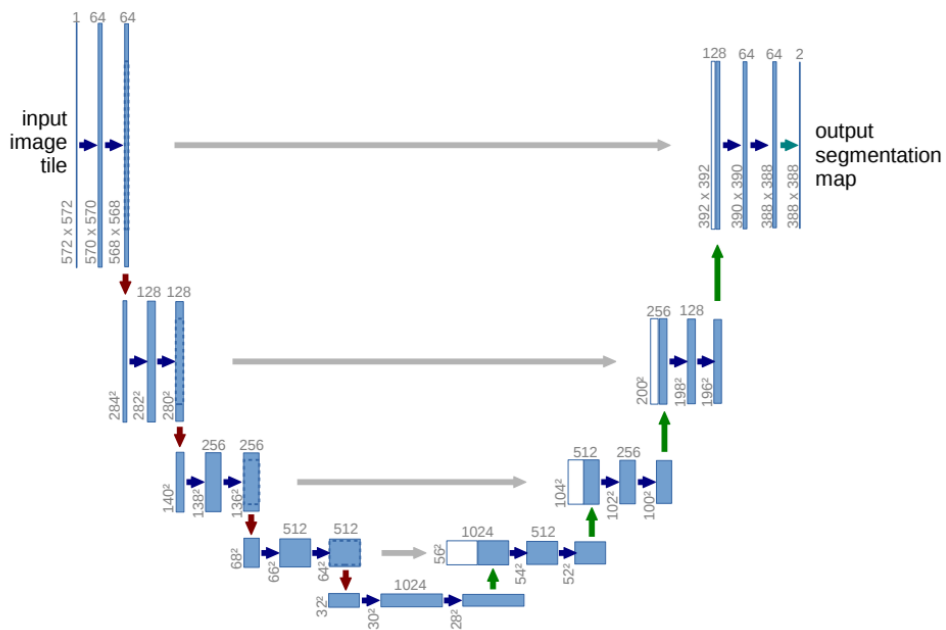
Figure 15: U-Net Architecture (32x32 Pixels) [25].

3.6.2 Residual U-Net

The residual U-Net is an enhancement of the standard architecture introduced by Olaf Ronneberg; the architecture is designed using a U-shaped deep fully connected convolutional CNN. Then, Residual layers are incorporated at every stage within both the compressing and expanding pathways. In the contracting path, each stage comprises a 3×3 convolutional filter set and its accompanying residual layer. At the conclusion of every stage, the resulting images undergo downsizing through a 2×2 average-pooling. Consequently, the input patch size reduces by a factor of two, a process beneficial for enhancing feature propagation due to the typically smooth circular shape of small extracellular vesicles in 2D TEM images. Post downsizing, the feature map size at each stage increases twofold, progressing from 32 to 256 in the contracting path. By merging the features at the same level from the contracting path

with an unsampled version of the features from the level before it in the expanding path, the expanding path creates its feature maps. The remaining blocks in the expanding path are not symmetrical like those in the shrinking path. More specifically, the merging of the expanding and high-level features serves as the direct input for these residual layers. A separate 1×1 convolutional layer is incorporated into the residual block in this configuration, which lowers the quantity of feature mappings while keeping the fine features from the contracting path. Furthermore, the first 3×3 convolution in the residual layer reduces the size of the feature map, which aids in the formation of the residual connection. Figure 16 is the block diagram of the residual U-Net.
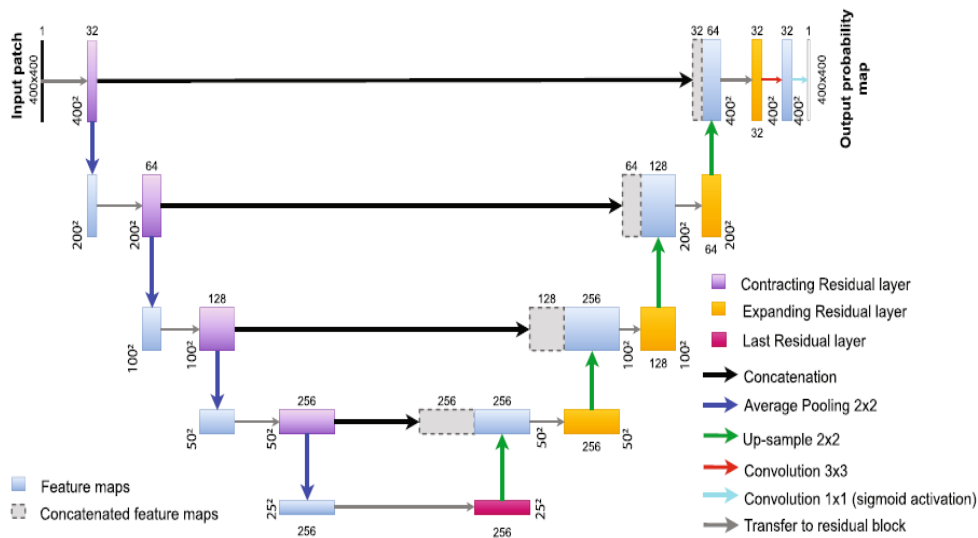
Figure16: Residual U-Net architecture [18].

3.6.3 Attention U-Net

The attention gate module is inserted at the skip links between the encoder and decoder in Attention U-Net, which is an extension of regular U-Net that emphasizes key aspects and mutes' details of unimportant regions. An attention mechanism's primary goal is to quickly connect a set of key-value pairs to a query and generate an output in the most efficient way possible. In the attention mechanism, vectors that represent key, query, values, and outputs are used. Weights are used to determine how well a query and its matching key work together. The output is then obtained by computing the weighted sums of the values. These weights represent the importance of the inputs (keys) with respect to a certain output (the query). As x^1 originates from earlier layers, it holds richer spatial information. Input g, known as the "gating signal," is derived from a deeper network layer and encompasses enhanced feature representations and

contextual details for identifying the focal region. Figure 17 depicts the block diagram of attention U-Net.
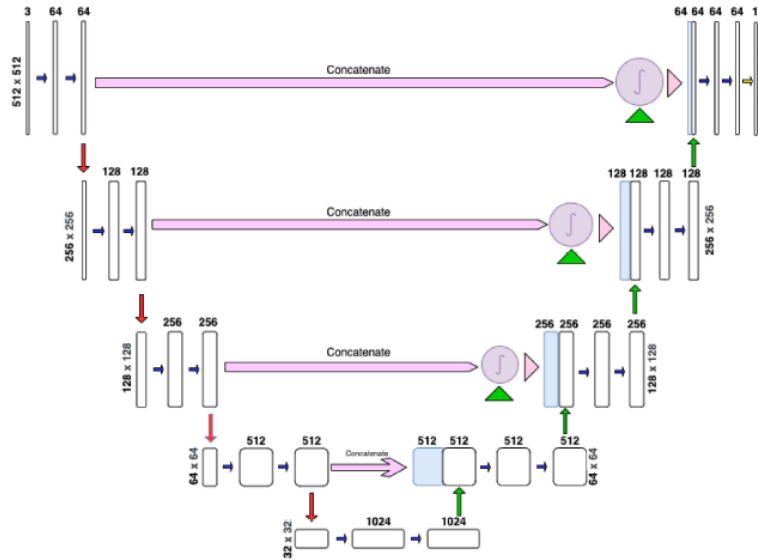
Figure 17: Attention U-Net architecture for segmenting the region of interest [2].

3.6.4 Residual-Attention U-Net

The residual attention U-Net is modified here called DeepSegNet, which is composed of two components. The first component is made up of residual blocks and another module is attention. The integration of residual and attention blocks enhances the performance of the network. Both modules enable the network to focus on important areas while retaining intricate details.
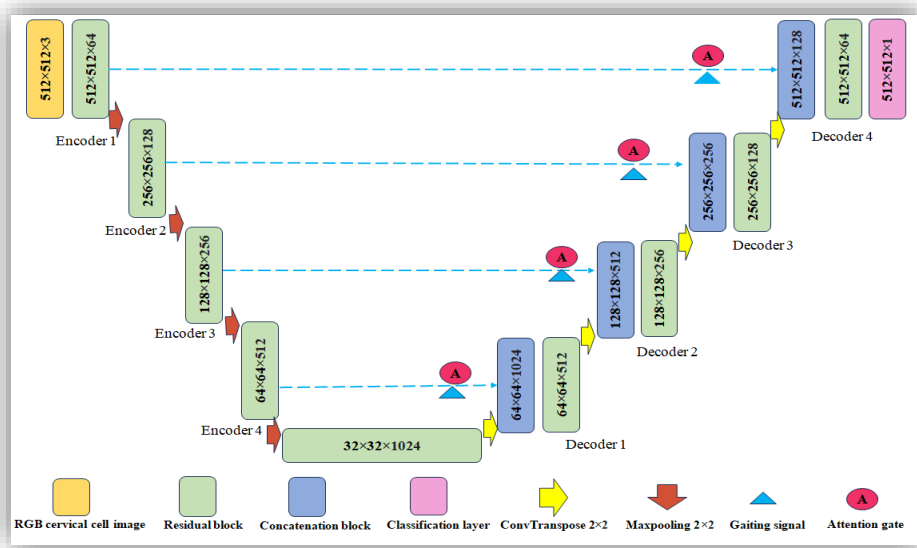
Figure 18: Residual attention U-Net architecture.

Attention mechanisms within deep learning models guide the network in allocating its resources by assigning varying importance to distinct sections of the input. Incorporating attention blocks into the U-Net design empowers RAU-Net to highlight crucial features and diminish irrelevant or noisy data, consequently enhancing the accuracy of segmentation. Figure 18 is the visual representation of the Residual Attention U-Net.

- **Attention Module**

The attention mechanism, akin to human visual focus, effectively filters out irrelevant image regions [20], thus reducing false positives. In this study, it enables parameter updates in spatial areas crucial for cytoplasm and nucleus segmentation. A connection is established between the encoder and decoder through a skip connection, utilizing an attention gate. This gate takes two inputs: one from the encoder, containing contextual and spatial information, and the other from the decoder layer below. The gate's output is subsequently directed to the decoder for concatenation. For a visual representation of our model's attention gate, please refer to Figure 19.
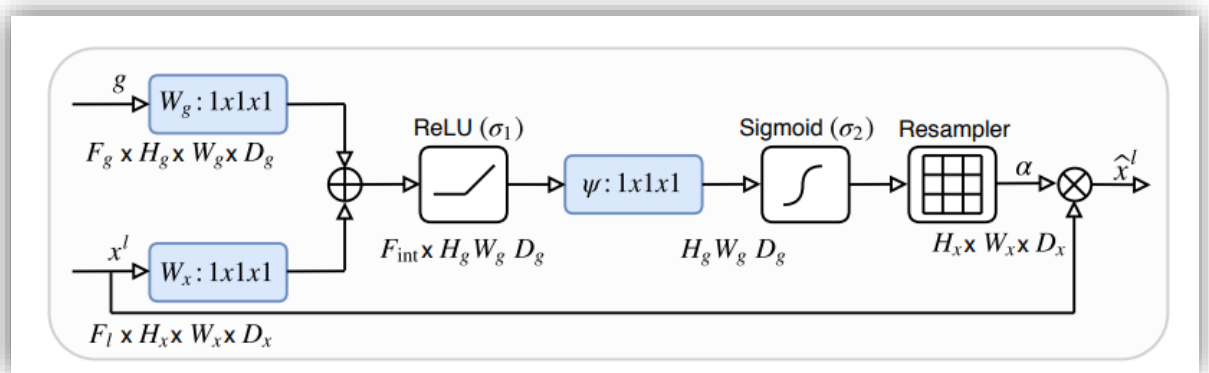
Figure 19: Block diagram of attention module[54].

- **Residual block**

A residual block is a sequence of layers arranged in a manner where the output of one layer is combined with the output of a deeper layer within the block. After this combination, a non-linear operation is applied to the result alongside the output of the corresponding layer in the main path. The main difference between the traditional network and the residual block is that instead of F(X) the F(X)-X is used. In networks containing residual blocks, each layer passes its output to the subsequent layer and to layers approximately 2-3 steps ahead. Here in, R(X)=F(X)-X or F(X)=R(X)+X. Therefore, the layers are trying to learn the residual R(X).

Thus, the output is added up with the identity which is shortcut connection then passed to the next layer. Figure 20 mentioned below depicts the structure of various residual blocks used.
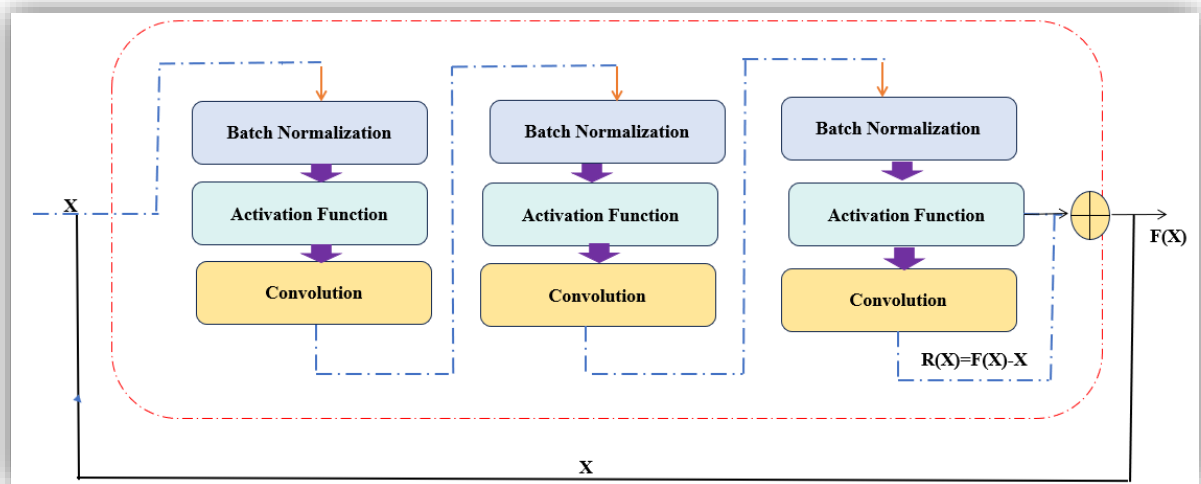
Figure 20: Structure of different residual blocks utilized in the architecture.

- **Algorithm of DeepSegNet a modified architecture of Residual Attention U-Net**
  BEGIN
    1. Input: Pap smear images with cytoplasm and nucleus.
    2. Preprocess the images (e.g., resizing, normalization, etc.)
    3. DEFINE Residual Attention U-Net
    4. INITIALIZE encoder layers
    5. DOWN-SAMPLE using max pooling
    6. Integrate residual blocks
    7. APPLY attention modules
    8. UP-SAMPLE using transposed convolutions
    9. ADD residual blocks
    10. FINALIZE with convolution layer
    11. INITIALIZE model parameters
    12. DEFINE loss function
    13. SELECT Adam optimizer and learning rate schedule
    14. SPLIT dataset into training, validation, and testing sets
    15. FOR each epoch DO
    16. FOR each batch of training data DO
    17. PROPAGATE input through the network

60

18. COMPUTE loss between predicted segmentation masks and ground truth

19. COMPUTE gradients of the loss w.r.t. model parameters

20. UPDATE model parameters using optimizer

21. END FOR

22. COMPUTE validation loss

23. ADJUST learning rate, if necessary, based on validation performance

24. END FOR

25. GENERATE segmentation masks for test images

26. EVALUATE segmentation accuracy using appropriate metrics

27. VISUALIZE segmentation results

28. ANALYZE performance

29. DEPLOY the model for segmenting cytoplasm and nucleus in new pap smear images

END

- **Model Hyperparameter Tunning**

The manual method of hyperparameter tuning was applied with several sets of values. It uses a trial-and-error approach, tracking trial outcomes and using them as feedback to find the combination of hyperparameters that performs best. The implemented filters are 16, 32, 64, 128, and 256, and the kernel size is 3. The Adam optimizer is used, batch normalization is enabled, and dropout is set at 0.5. Twenty epochs are followed by a batch size of 32. The selection of these hyperparameters resulted in high performance of the trained models. The chosen hyperparameters were selected based on their proven effectiveness in deep learning techniques and to balance computational efficiency with model performance. Filters (16, 32, 64, 128, and 256) capture features at various levels of abstraction, while a kernel size of 3 effectively extracts spatial features. The Adam optimizer, known for its adaptive learning rate, and batch normalization stabilize and accelerate training. A dropout rate of 0.5 prevents overfitting, and twenty epochs ensure sufficient learning without excessive fitting. A batch size of 32 balances gradient estimate noise and memory usage. These choices, fine-tuned through a trial-and-error approach, resulted in high performance for the trained models.

3.7 Feature Extraction

Feature extraction entails the process of diminishing dimensionality of raw data, breaking it into smaller, more manageable groups for easier processing. Large datasets, characterized by numerous variables, demand substantial computing resources. Feature extraction addresses

this challenge by selecting and combining variables to create features, significantly reducing the data volume. These derived features remain easy to process yet accurately represent the original dataset with precision and authenticity. The architecture used for feature extraction is VGG16.

3.7.1 VGG16

VGG stands for Visual Geometry Group, the different variants of VGG are VGG11, VGG13, VGG16, and VGG19. The architectures vary in terms of the number of layers they possess. VGG11, and VGG13 contains 11 and 13 convolutional layers, which are comparatively fewer than VGG16 and VGG19. VGG16 is based on 16 layers convolutional layers and last three are fully connected layers. Each convolutional block is made up of multiple 3x3 convolutional layers followed by a max-pooling layer, contributing to the network's ability to learn rich representations of input images. VGG16 has been used widely as a base architecture for various tasks, such as transfer learning. VGG19 consists of 19 layers. It is a variant of VGG16, three extra convolutional layers are added into it, to make it deeper. Since deeper models tend to learn more complex features but requires more computational features. In an ImageNet dataset, VGG16 obtained 92.7% top 5 test accuracy. This model is trained on 14 million images which contain 1000 classes. The input to the cov1 layer maintains a fixed size of 256 x 256 RGB image. A sequence of convolutional (conv.) layers processes the image using 3×3 filters. Some configurations incorporate 1×1 convolution filters, functioning as a linear transformation of input channels followed by non-linear processes. Convolution occurs with a fixed stride of 1 pixel, maintaining spatial resolution through 1-pixel padding for 3×3 conv. layers. Among the conv. layers, five max-pooling layers perform spatial pooling using a 2×2-pixel window and a stride of 2. Following the convolutional layers of varying depths across architectures, three Fully Connected layers are employed: the first two possess 4096 channels each, while the third facilitates 1000-way ILSVRC classification with 1000 channels, representing individual classes. The soft-max layer is made up of the last layer, followed by ReLU which is encoded in every hidden layer. Notably, barring one exception, the networks do not include Local Response Normalisation (LRN) due to its negligible impact on ILSVRC dataset performance, yet it significantly increases memory usage and computation time. Following figure 21 is the architecture of VGG16 with various convolutional layers.
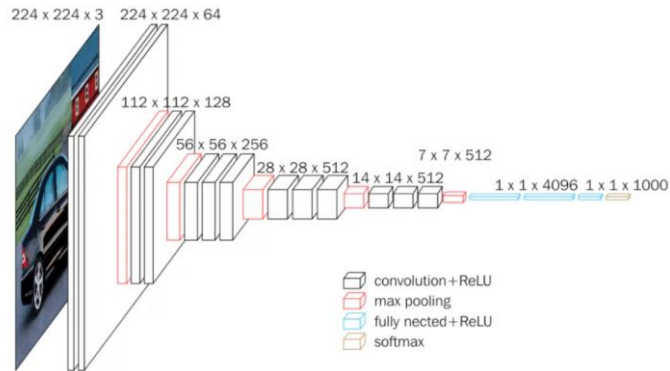
Figure 21: VGG16 architecture for feature extraction [123].

3.8 Classifier

In machine learning, a classifier is an algorithm that autonomously organizes or groups data into one or multiple predefined "classes." The various types of classifiers are decision tree, Support Vector Machine, Naïve Bayes classifier, K Nearest Neighbor, Artificial Neural Network, Random Forest, Logistic Regression, and Ensemble Methods. These classifiers play an important role as they facilitate the automated decision making, enable predictions, and unveil patterns within data that might not be immediately apparent. The two classifiers chosen for the classification purpose include ANN and XGBoost.

3.8.1 Artificial Neural Network

In this classification module, a three-layer artificial neural network architecture, as depicted in Figure 22, is utilized. The initial layer, serving as the input, comprises several nodes that receive data from various cervical cells. At the other end, the output layer consists of one node that provides a confidence level for a sample's classification. To normalize the network's output and restrict each output value within the 0 to 1 range, a final output function known as SoftMax is employed. Notably, the sum of the output nodes' values always equals 1, enabling straightforward interpretation as probabilities or confidence scores. Consequently, the predicted class corresponds to the one with the highest confidence level. The intermediate layer, also known as the hidden layer, is completely connected and resides between the input and output layers. All its nodes receive input from each of the seven input nodes. It processes this information and transmits the results to the three output nodes. Different configurations with varying numbers of nodes in the hidden layer were explored in this study to strike a balance between architecture complexity and effective classification. Thus, to perform the

63

multiclass classification of pap smear images, an approach based on ANNs utilizing features extracted from Cell profiler is proposed. The input features are derived from different cellular regions, including intermediate squamous, superficial squamous, basal, parabasal, LSIL, and HSIL. The ANNs are trained on a comprehensive dataset of pap smear images, enabling them to learn the complex patterns and relationships between the extracted features and the respective classes. The ANN's architecture is carefully designed to ensure effective classification performance. By leveraging this approach, the aim is to enhance the accuracy and efficiency of Pap smear image classification, facilitating automated screening and detection of cervical abnormalities. The dataset is split into training and validation subsets, ensuring proper evaluation of the ANN's performance. During the training phase, the ANNs learn to map the extracted features to their respective classes through the backpropagation algorithm and gradient descent optimization. Evaluation metrics such as recall, and F1-score, accuracy, and precision, are employed to assess the ANN's performance on the validation set. Hyperparameter tuning is conducted to optimize the network's performance by adjusting parameters such as learning rate, batch size, and regularization techniques. Table 3 provides a comprehensive overview of the model parameters essential for carrying out the classification process. The classification is based on multiclass classification according to the latest Bethesda System of classification.
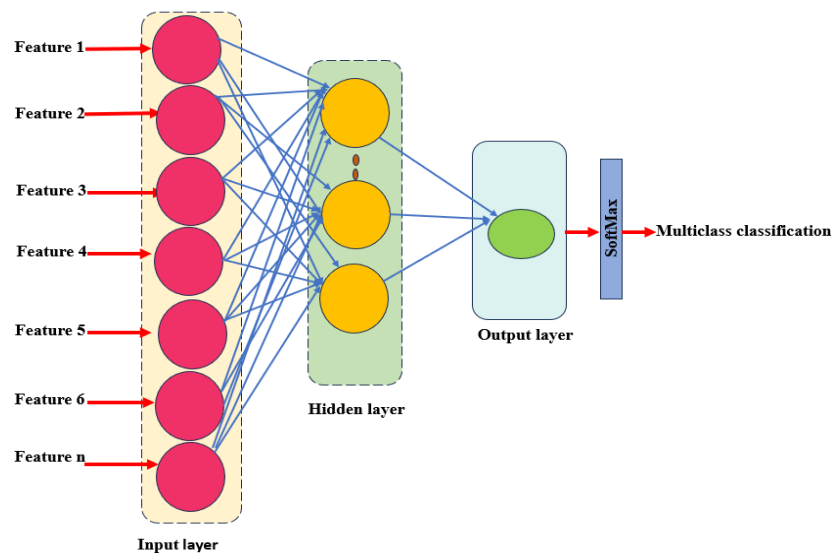
Figure 22: Multiclass cervical cell classification using MLP.

Table 3: Artificial neural network model parameters.

Parameter Name	Value
Epochs	100
Batch size	32
Verbose	1
Optimizer	Adam optimizer
Loss function	Categorical cross entropy
Activation function	ReLU and SoftMax
Test and train	20:80

3.8.2 XGBoost

Cervical cancer diagnosis relies heavily on the accurate and early classification of Pap smear images. In this context, we introduce a state-of-the-art approach that leverages the combined strengths of deep learning, specifically the VGG16 convolutional neural network architecture, and gradient boosting through XGBoost. VGG16 is a well-established neural network architecture celebrated for its effectiveness in feature extraction from diverse image types, including Pap smear images. The VGG16 architecture is characterized by multiple layers of convolution, each followed by a non-linear activation function, allowing it to progressively learn intricate and abstract visual features. When a Pap smear image is input into VGG16, it undergoes a series of convolutions with learnable filters designed to detect patterns across a spectrum of scales, from elementary features such as edges and corners to more complex textures and shapes. As the image traverses through the network, subsequent convolutional layers combine these lower-level features to construct higher-level and more sophisticated image representations. A distinctive feature of VGG16 is its repeated use of 3x3 convolutional filters, which enables deeper networks with a more compact receptive field. This architectural choice enhances the ability to capture intricate details while reducing the spatial dimensions of the input. In the method, 8-bit 256x256-pixel cervical cell images are employed as input. The pre-trained VGG16 with 'ImageNet' weights is adopted, bypassing the dense layers as the focus remains solely on feature extraction. These pre-trained weights are used to extract a rich set of features. Given that 843 cervical cell images of size 256x256x3 are used, VGG16 extracts approximately 32,768 features (8x8x512). Subsequently, XGBoost, a robust gradient boosting classifier, is applied to perform multiclass classification.

This holistic approach blends the power of VGG16 for feature extraction with the precision of XGBoost for multiclass classification, offering an effective tool for the early and accurate detection of cervical cancer. The process commences with meticulous data preprocessing, which involves resizing Pap smear images to a consistent 256x256-pixel dimension and ensuring color normalization. Further, labels are systematically assigned to images based on their folder structure, providing the necessary ground truth for training. Feature extraction with VGG16 takes centre stage, capitalizing on its ability to capture intricate patterns, textures, and structures. The pre-trained weights facilitate the automatic distillation of essential features from Pap smear images. Notably, VGG16's top classifier layers are pruned, preserving only the convolutional layers for this feature extraction step. The extracted features are subsequently reshaped into high-dimensional feature matrices, encapsulating the nuanced information crucial for precise classification. These matrices serve as the foundation for the final classification step with XGBoost, an algorithm well-suited for the intricacies of high-dimensional data. Through this amalgamation of deep learning and gradient boosting, the approach holds immense promise for enhancing the diagnostic accuracy of cervical cancer via Pap smear image analysis. A diagrammatic representation of multi-level image classification of cervical cells into cancerous or non-cancerous cells is depicted in figure 23. The VGG16 model summary is presented in table 4.
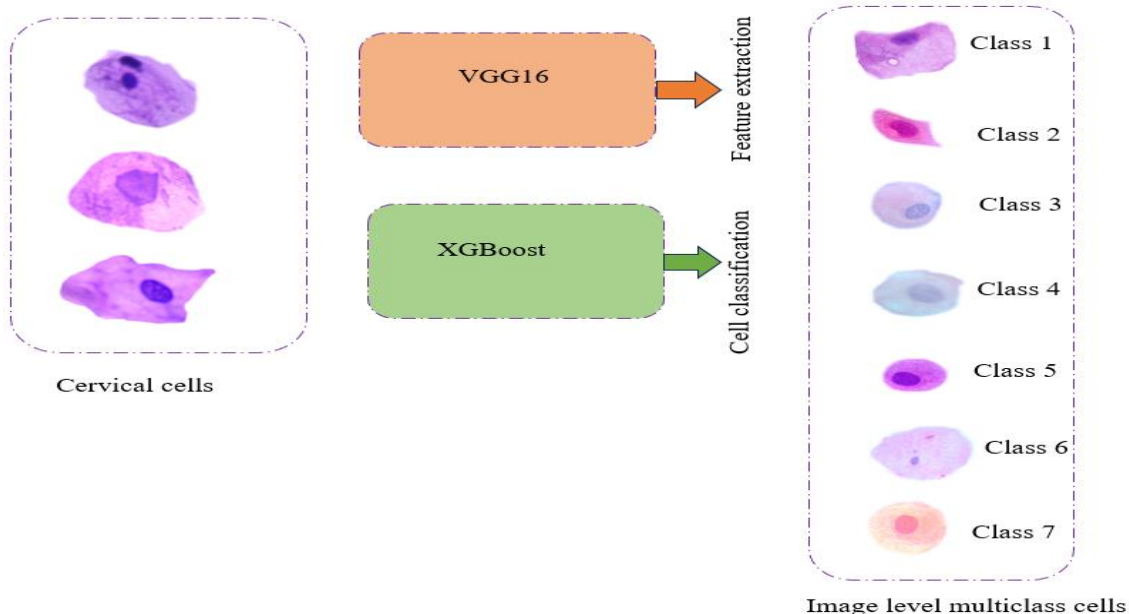
Figure 23: Multi-level classification of cervical cells using VGG16-XGBoost.

Table 4: VGG16 model summary of Convolutional and Maxpooling layers

Layer (type)	Output Shape	Parameter
input_1 (Input Layer)	[(None, 256, 256, 3)]	0
block1_conv1 (Conv2D)	(None, 256, 256, 64)	1792
block1_conv2 (Conv2D)	(None, 256, 256, 64)	36928
block1_pool (MaxPooling2D)	(None, 128, 128, 64)	0
block2_conv1 (Conv2D)	(None, 128, 128, 128)	73856
block2_conv2 (Conv2D)	(None, 128, 128, 128)	147584
block2_pool (MaxPooling2D)	(None, 64, 64, 128)	0
block3_conv1 (Conv2D)	(None, 64, 64, 256)	295168
block3_conv2 (Conv2D)	(None, 64, 64, 256)	590080
block3_conv3 (Conv2D)	(None, 32, 32, 256)	0
block4_conv1 (Conv2D)	(None, 32, 32, 512)	1180160
block4_conv2 (Conv2D)	(None, 32, 32, 512)	2359808
block4_conv3 (Conv2D	(None, 32, 32, 512)	2359808
block4_pool (MaxPooling2D)	(None, 16, 16, 512)	0
block5_conv1 (Conv2D)	(None, 16, 16, 512)	2359808
block5_conv2 (Conv2D)	(None, 16, 16, 512)	2359808
block5_conv3 (Conv2D)	(None, 16, 16, 512)	2359808
block5_pool (MaxPooling2D)	(None, 8, 8, 512)	0
	Total parameters: 14,714,688 Trainable parameters: 0 Non-Trainable parameters:14,714,688	

3.9 Justification of the Chosen Architectures

The choice to use the Residual Attention U-Net model for cytoplasm and nucleus segmentation in Pap smear images and an ANN for classification is based on several benefits that these methods offer in the field of cytological analysis. The intrinsic ability of the Residual Attention U-Net to tackle intricate structural recognition tasks makes it a reliable option. The attention mechanisms incorporated into this architecture allow for selective emphasis on important characteristics and efficient noise suppression, which guarantees correct cellular boundary delineation, which is essential for precisely segmenting nuclear and cytoplasmic regions. Additionally, the U-Net architecture's residual connections allow for smoother information flow, which successfully addresses the vanishing gradient problem and makes learning complex structural elements easier a particularly important aspect. The subsequent use of an ANN for classification benefits from its ability to identify intricate relationships within the segmented regions, ensuring precise classification of specific cellular components such as cytoplasm and nucleus. This approach combines the precise segmentation capabilities of the

Residual Attention U-Net with the ANN's strength in learning from diverse features, resulting in a synergistic methodology that aims for accurate cytological analysis in pap smear images. Furthermore, this strategy's effectiveness is not only rooted in its inherent strengths but also in its potential adaptability and scalability. It has the capability to expand and adapt to handle diverse cell types, varying image resolutions, and other complexities commonly encountered in cytological image analysis.

3.10 SUMMARY

The chapter elaborates on the diverse array of architectures meticulously employed in executing experimental setups. These architectures were specifically tailored for various tasks involving Pap smear image analysis, encompassing segmentation, classification, and feature extraction. The ensemble comprises a spectrum of sophisticated models, including the standard U-Net, Residual U-Net, Attention U-Net, Residual Attention U-Net called Deep-SegNet, VGG16 architecture, Artificial Neural Network, and XGBoost. Each model was strategically chosen to contribute to the comprehensive classification of cervical cells across seven distinct classes: class I, class II, class III, class IV, class V, class VI, and class VII. The utilization of these deep learning and machine learning models serves a distinct purpose within the experimental paradigm. The standard U-Net, renowned for its efficacy in biomedical image segmentation, forms the bedrock of the segmentation task. Its ability to accurately delineate cellular structures and boundaries in Pap smear images lays the groundwork for subsequent analyses. In parallel, the Residual U-Net and Attention U-Net variants enhance segmentation precision by incorporating residual connections and attention mechanisms, respectively. The Deep-SegNet, a fusion of these features, further refines the segmentation process by selectively attending to vital regions while leveraging residual connections for enhanced information flow. Moving beyond segmentation, the VGG16 architecture, with its deep convolutional layers, emerges as a pivotal tool for feature extraction. Its hierarchical representation of image features enables the extraction of high-level features crucial for differentiating between cervical cell classes. Meanwhile, the integration of an Artificial Neural Network and XGBoost specifically caters to the classification phase. The ANN's capacity to discern intricate relationships within segmented regions aids in precise classification, while XGBoost, a robust boosting algorithm, complements this process by enhancing the model's accuracy and interpretability. The collective application of these models amalgamates intricate techniques and methodologies, aiming to achieve comprehensive and precise classification of cervical cells. By leveraging the distinctive strengths of each architecture, this multifaceted approach not only strives for

accurate classification across the seven cell classes but also seeks to establish a robust framework adaptable to the nuances and complexities inherent in cytological image analysis.

CHAPTER 4
RESULTS AND DISCUSSION

This chapter extensively evaluates standard U-Net, residual U-Net, attention U-Net, and Residual Attention U-Net performance in segmenting cytoplasm and nucleus. After successful segmentation, the focus shifts to feature extraction, employing VGG16, and subsequent classification using ANN, and XGBoost. Moreover, this chapter employs a diverse range of performance metrics such as accuracy, precision, recall, and others to comprehensively evaluate the capabilities of these techniques in accurately identifying cellular components and performing precise classifications based on the extracted features. The meticulous integration of these algorithms allows for robust and accurate predictions, enabling the identification of potentially malignant cellular patterns with high precision.

4.1 Segmentation Result

This section showcases visual representations of segmented cytoplasm and nucleus derived from different neural network architectures: the Standard U-Net, Residual U-Net, Attention U-Net, and Residual Attention U-Net. Through these visualizations, the diverse segmentation outputs of each model become apparent, highlighting their distinct approaches in delineating cellular structures within pap smear images. The comparative analysis of these representations offers valuable insights into the varying degrees of accuracy, detail preservation, and boundary delineation achieved by these different architectures, paving the way for a comprehensive understanding of their respective strengths and limitations in cellular segmentation tasks.

4.1.1 Standard U-Net

In this subsection, the focus centers on delineating the segmentation process applied to the cytoplasm and nucleus within Pap smear images. The visual representation provided vividly illustrates the challenges encountered in employing the standard U-Net architecture for achieving accurate segmentation. Through the displayed images, a noticeable limitation becomes apparent: the standard U-Net model struggles to effectively segment both the cytoplasm and nucleus with precision. Despite its widespread use and success in various segmentation tasks, the U-Net architecture encounters difficulties in achieving satisfactory segmentation results in this context. Notably, the evaluation metrics such as Intersection over Union (IoU) and accuracy demonstrate constrained performance, emphasizing the necessity for an enhanced segmentation approach tailored to the complexities inherent in Pap smear image analysis. The discrepancies between the anticipated and obtained segmentation outcomes underscore the need for a refined methodology that can accurately delineate the cytoplasmic

and nuclear boundaries, ensuring improved reliability in diagnostic interpretations. This observation sets the stage for further exploration and development of advanced segmentation techniques, aimed at augmenting the accuracy and efficacy of Pap smear image analysis for enhanced clinical applications and pathological assessments. Figure 24 is the segmented cytoplasm and nucleus from cervical cells. Images a, b, c, d, e, and f depict parabasal and intermediate squamous cells. The first three represent parabasal cells, followed by three squamous cells. Correspondingly, a1, b1, c1 are the ground truth cytoplasm images of a, b, and c, while d1, e1, f1 are the ground truth cytoplasm images of d, e, and f. Additionally, a2, b2, c2 represent the ground truth nuclei for a, b, and c, whereas d2, e2, and f2 are the ground truth nuclei for d, e, and f. A1, A2 denote the segmented cytoplasm and nucleus of cell a, B1, B2 represent those of cell b, and C1, C2 depict those of cell c. Similarly, D1, D2 are the segmented cytoplasm and nucleus of cell d, E1, E2 are for cell e, and F1, F2 are for cell f. The visual representation of these images reveals that the U-Net model does not preserve the cellular structure while segmenting Pap smear images.
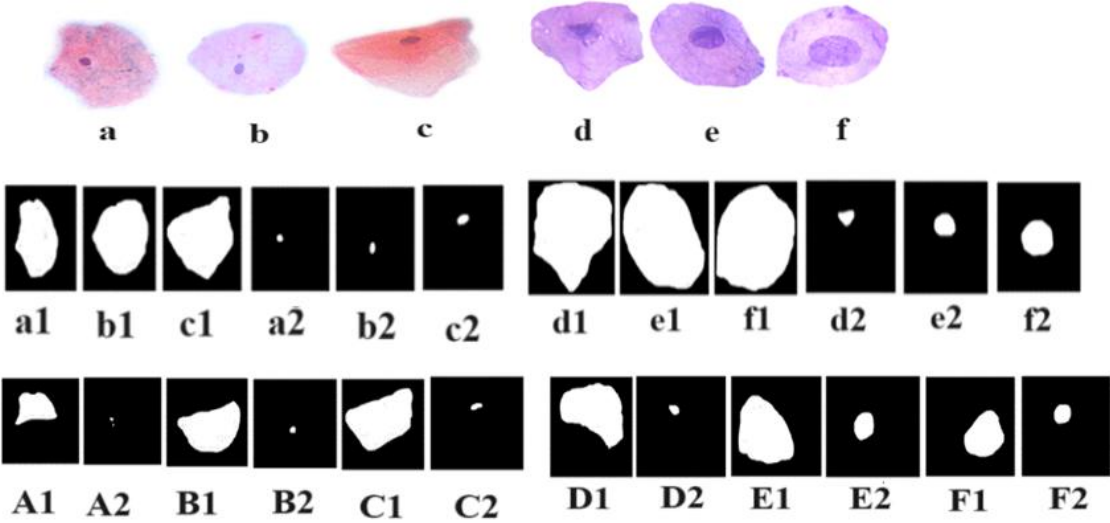
Figure 24: Segmented cytoplasm and nucleus using standard U-Net.

4.1.2 Residual U-Net

This subsection highlights the comparative analysis between the standard U-Net and the Residual U-Net architectures in their application to Pap smear image segmentation. Notably, the Residual U-Net demonstrates superior performance in segmenting select pap smear images, showcasing improved accuracy and delineation of both cytoplasmic and nuclear regions compared to its standard counterpart. However, despite this progress, it becomes evident that

71

the Residual U-Net still encounters challenges in achieving consistent and precise segmentation across the entirety of the Pap smear images. While the Residual U-Net exhibits promising advancements over the standard U-Net by yielding better segmentation results in specific instances, the variability in outcomes across different images highlights the necessity for further refinement and enhancement in the segmentation methodology. Figure 25 represents the segmented cytoplasm and nucleus with residual U-Net. This observation underscores the ongoing need for innovation and optimization within the Residual U-Net framework to ensure consistent and accurate segmentation of all pap smear images, aiming to establish a more robust and reliable tool for pathological assessments and diagnostic applications in clinical settings. Images a-f depict parabasal and intermediate squamous cells. The first three are parabasal cells, followed by three squamous cells. Correspondingly, a1-c1 are the ground truth cytoplasm images for a-c, and d1-f1 are for d-f. Similarly, a2-c2 represent the ground truth nuclei for a-c, and d2-f2 for d-f. Segmented cytoplasm and nuclei for each cell are denoted as A1-A2, B1-B2, C1-C2, D1-D2, E1-E2, and F1-F2. The visual representation reveals that the residual U-Net model fails to preserve cellular structure in segmenting Pap Smear images.
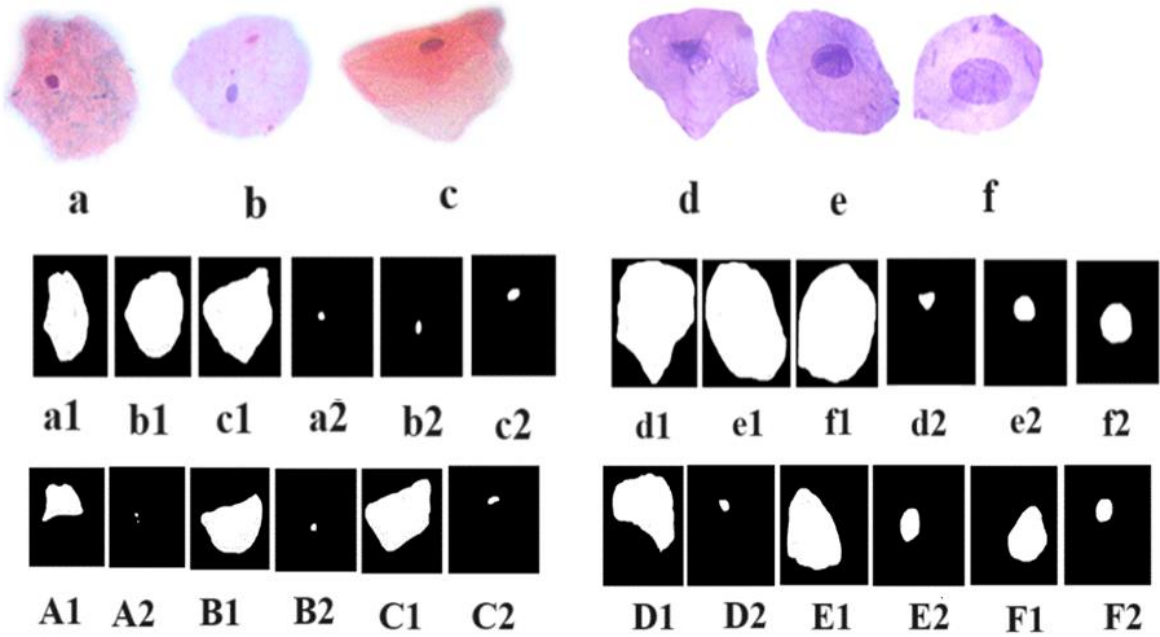
Figure 25: Segmentation of cytoplasm and nucleus using residual U-Net.

4.1.3 Attention U-Net

Despite the advancements brought by the Residual U-Net architecture in the segmentation of cytoplasmic and nuclear regions within Pap smear images, its performance, as indicated by

metrics such as Intersection over Union (IoU) and accuracy, remains below the desired standards. While it exhibits improved segmentation compared to traditional methods like the standard U-Net, there are persistent challenges in accurately defining the boundaries of these crucial cellular components consistently across the entire dataset. The IoU and accuracy scores obtained through the Residual U-Net analysis reveal inconsistencies and inaccuracies in segmenting cytoplasmic and nuclear structures within the images, hindering its reliability for clinical interpretations and pathological assessments. These limitations emphasize the critical need for further enhancements and refinements in the Residual U-Net model to achieve more consistent and precise segmentation results across diverse Pap smear images. Advancements in this area are crucial for bolstering the accuracy and reliability of automated analysis tools in the realm of cytopathology, ultimately improving diagnostic capabilities and patient care in clinical settings. Figure 26 is the visualized images of segmented cytoplasm and nucleus using attention U-Net. The cell types depicted in images a–c are parabasal cells, while images d–f show squamous cells. Ground truth images of the cytoplasm (a1-f1) and nucleus (a2-f2) correspond to each cell. For cells a-f, the labels A1-A2, B1-B2, C1-C2, D1-D2, E1-E2, and F1-F2 correspond to the segmented cytoplasm and nucleus images. The outputs obtained from the Attention U-Net model did not effectively preserve the boundaries between cytoplasmic and nuclear regions.
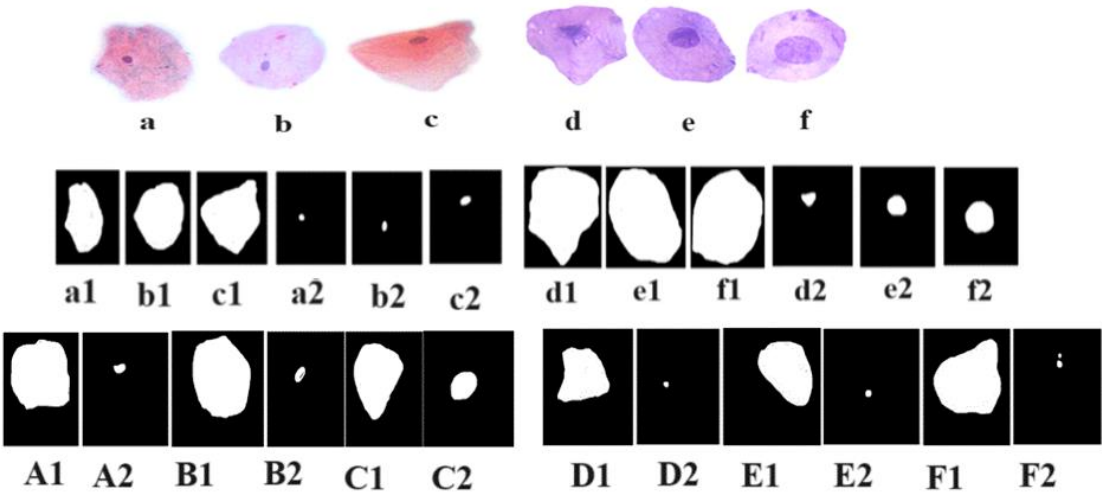
Figure 26: Segmentation of cytoplasm and nucleus using attention U-Net.

4.1.4 Residual Attention U-Net (Deep SegNet)

The segmentation of cytoplasm and nucleus utilizing the Residual Attention U-Net represents a substantial leap forward, notably improving Intersection over Union (IoU) and accuracy

73

metrics while preserving intricate details and edges within the images. This advanced architecture, the Residual Attention U-Net, demonstrates significant success in accurately delineating the boundaries of cytoplasmic and nuclear regions, showcasing superior preservation of fine details and structural edges compared to previous models. The IoU and accuracy scores obtained through the Residual Attention U-Net analysis signify a marked enhancement in the precision and consistency of segmentation, ensuring a more comprehensive preservation of intricate cellular structures and edges within pap smear images. This notable achievement underscores the efficacy and potential of advanced architectures, particularly the Residual Attention U-Net, in bolstering the accuracy and reliability of automated segmentation tools for cytoplasmic and nuclear analysis. Such advancements hold promise for refining diagnostic interpretations in cytopathology and advancing clinical practices by providing more detailed and accurate insights into cellular morphology and pathology.
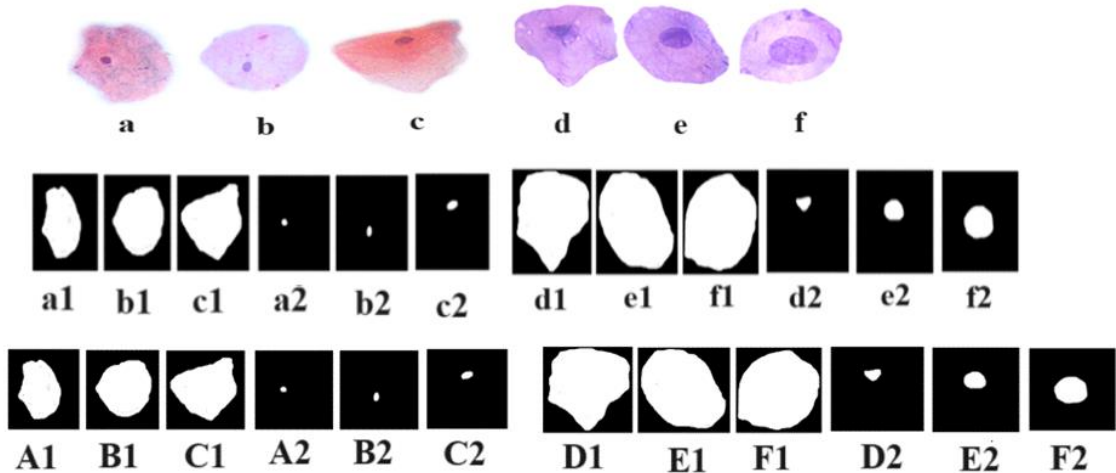
Figure 27: Segmented cytoplasm and nucleus with residual attention U-Net.

Figure 27 is the segmentation of cytoplasm and nucleus achieved using residual attention U-Net. Images a-f depict parabasal and intermediate squamous cells. The first three represent parabasal cells, followed by three squamous cells. Correspondingly, a1-c1 are the ground truth cytoplasm images of a-c, while d1-f1 represent the ground truth cytoplasm images of d-f. Additionally, a2-c2 represent the ground truth nuclei for cells a-c, whereas d2-f2 represent the ground truth nuclei for cells d-f. A1-A2 represent the segmented cytoplasm and nucleus of cell a, B1-B2 represent those of cell b, and C1-C2 depict those of cell c. Similarly, D1-D2 represent the segmented cytoplasm and nucleus of cell d, E1-E2 represent those of cell e, and F1-F2 represent those of cell f. The visual representation of these images reveals that the Residual

attention U-Net model preserve the cellular structure while segmenting Pap Smear images compared to other standard models. Figure 28 represents the training and validation of IOU and Loss of the same architecture. Table 5 is the performance comparison of various deep learning models

Table 5: Performance comparison of segmentation models.

Technique	Cytoplasm		Nucleus	
	IoU	Accuracy	IoU	Accuracy
U-Net	90.23	91.57	91.23	90.27
Residual U-Net	70.53	73.10	69.53	75.10
Attention U-Net	71.64	73.23	71.02	76.23
Deep SegNet	97.36	99.04	96.36	99.84

The standard U-Net model achieves an Intersection over Union (IoU) and accuracy of 90.23 for cytoplasm segmentation and 91.23 and 90.27 for nucleus segmentation, respectively. This performance is comparatively better than that of the Residual and Attention U-Net models. However, the complexity of these models increases with the integration of individual residual and attention blocks.
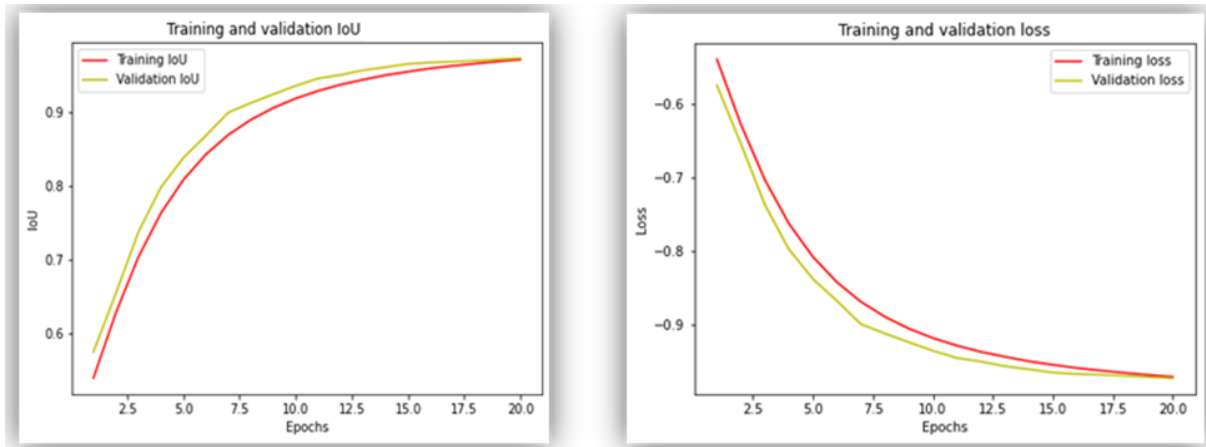
Figure 28: Training and validation IoU and Loss of residual-attention-U-Net.

Interestingly, when both modules are integrated together into the standard U-Net architecture, there is a drastic improvement in performance for cytoplasm and nucleus segmentation. The integrated model achieves an IoU of 97.36 and 96.36 for cytoplasm and nucleus, with accuracy

rates of 99.04 and 99.84 for cytoplasm and nucleus, respectively. These results were obtained after training the models for 20 epochs.

4.2 Feature Extraction and Classification

In an automated cervical screening approach, such as DeepCervix-Net, the morphological features of cervical cells are analyzed meticulously and efficiently. These features, selected in consultation with trained cytopathologists and cyto-technicians, are calibrated in terms of pixel count, as indicated in Table 6. Morphological examination involves assessing size, shape, orientation, and the proportion of the cell area shared by the nucleus and cytoplasm. This study's formalization of these features using computer techniques and image processing algorithms replaces the subjective manual interpretation by cytopathologists, minimizing the possibility for human errors (Bamford and Lovell 1996; Lezoray and Cardot 2002) [56]. Various parameters, such as area, perimeter, form factor, solidity, extent, center, eccentricity, major and minor axis length, orientation, compactness, and nucleus-cytoplasm ratio, are examined to assess various aspects of cellular abnormalities. Table 6 illustrates the morphological features of the cytoplasm and nucleus extracted with Cell Profiler, and figure 29 visually demonstrates the feature extraction process on 843 pap smear images. A detailed description of the features is provided in the following section.
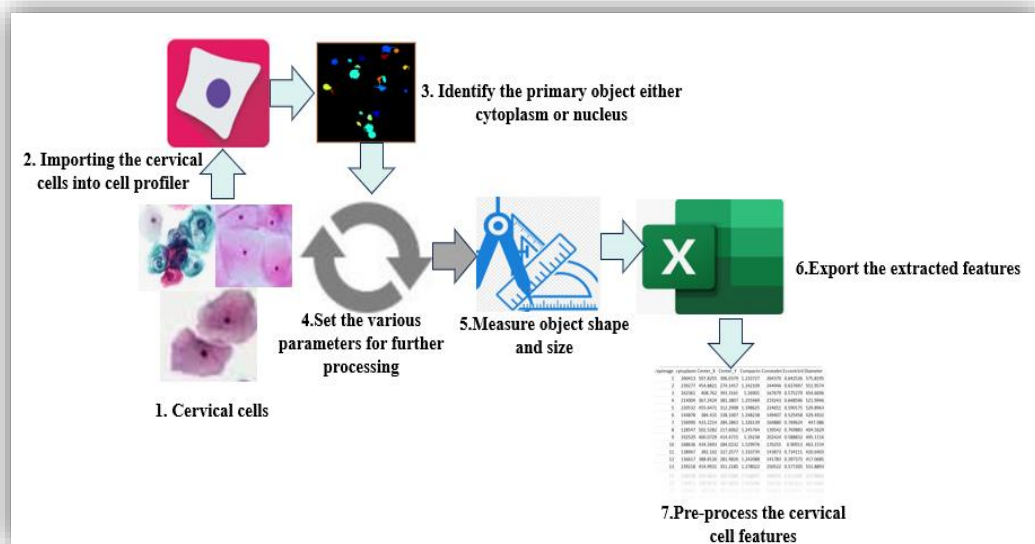
Figure 29: Visual representation of feature extraction from 843 Pap smear images.

4.2.1 Description of Extracted Features from Cell Profiler

4.2.1.1 Area of Cytoplasm-Nucleus

The area of cytoplasm and nucleus represents the total pixel count within their boundaries. The prognosis of malignancies can be significantly influenced by the size of the cytoplasm and nucleus. Normal cervical cells typically have a larger cytoplasmic area and a smaller nucleus compared to abnormal cervical cells like CIN-II and CIN-III, where the nucleus is almost as large as the cytoplasm. Meyer-Arendt and Humphreys [4] noted that cancerous cells are smaller in size compared to normal superficial and intermediate cells.

4.2.1.2 Cytoplasmic and Nuclear Perimeter

The cytoplasmic and nuclear perimeter refers to the number of pixels that make up the boundary of these regions. In this context, it's worth noting that normal cervical cells typically have a larger cytoplasmic perimeter but a smaller nuclear perimeter. Conversely, cervical intraepithelial neoplasia cells exhibit increased perimeter values for both cytoplasm and nucleus.

4.2.1.3 Form Factor

The form factor is a metric that quantifies the roundness of an object. It's determined by the formula: $4\pi \times area/perimeter^2$. When an object has a form factor value of 1, it means it has a perfectly circular shape. The form factor is a useful measure for identifying irregular and asymmetric cells by assessing how much they deviate from a normal circular shape.

4.2.1.4 Solidity

Solidity is a measure that compares the area under the convex hull (the smallest convex polygon encompassing all object points) to the object's total area. When Solidity equals 1, it signifies a smooth and regular surface, typical of Normal and Atypical Squamous Cells of Undetermined Significance (ACS-US) category cells. In contrast, cells from the High-grade Squamous Intraepithelial Lesion (HSIL) category, including moderate (CIN-II) and severe dysplasia (CIN-III), exhibit low Solidity values due to their irregular and non-convex surfaces.

4.2.1.5 Cytoplasmic and Nuclear Extent

Cytoplasmic and nuclear extent represents the proportion of the total figure area covered by the object. It's calculated by dividing the cytoplasmic or nuclear area by the overall figure area.

4.2.1.6 Center of Cytoplasm-Nucleus X-axis and Y-axis

The X and Y coordinates of the cytoplasm-nucleus centre represent the point farthest from the cytoplasm/nucleus edge. These coordinates are significant for characterizing the shape of the

cytoplasm and nucleus, as well as measuring the relative position of the nucleus in relation to the cytoplasm.

4.2.1.7 Eccentricity

Eccentricity is a parameter that measures the roundness or elongation of the cytoplasm and nucleus. It's determined by the ratio of the foci of an ellipse to its major axis length. As Cervical Intraepithelial Neoplasia progresses, the nucleus becomes more distorted, leading to an increase in eccentricity, especially in very elongated nuclei with a value approaching 1. Therefore, eccentricity is a useful indicator for identifying irregular cell shapes, particularly in the high-grade squamous intraepithelial lesion category, including moderate (CIN-II) and severe dysplasia (CIN-III).

4.2.1.8 Major Axis and Minor Axis Length of Cytoplasm-Nucleus

The major axis and minor axis lengths of the cytoplasm-nucleus represent the pixel measurements of the longest and shortest axes, respectively, of the best-fitting ellipse around the object. These measurements are instrumental in calculating other characteristics such as elongation and eccentricity.

4.2.1.9 Cytoplasmic and Nuclear orientation

Cytoplasmic and nuclear orientation indicates the angle between the X-axis and the major axis of the best-fitting ellipse around the biological object.

4.2.1.10 Cytoplasm/Nuclear Compactness

Cytoplasmic/nuclear compactness is a parameter used to standardize shape measurements, widely applied for characterizing morphological changes in biological entities [30,31]. It's determined by calculating the ratio of the perimeter squared to the ellipse's area. Observations reveal that ecto-cervical cells, including mature superficial and intermediate cells, exhibit a more compact nucleus but less compact cytoplasm in comparison to immature squamous cells like parabasal and basal cells, which are characterized by both compact cytoplasm and nucleus.

4.2.1.11 Nucleus-Cytoplasm Ratio

The presence of malignancy can be identified through the nucleus-to-cytoplasm ratio. Abnormal cells, such as atypical squamous cells with a potential for HSIL, CIN-II, and CIN-III, tend to exhibit elevated nucleus-to-cytoplasm ratios, which may approach a 1:1 ratio. Conversely, normal cells or atypical squamous cells of undetermined significance typically have lower nucleus-to-cytoplasm ratios, which may approximate ratios of 1:4 or 1:6 Das et al. [5]. The MaxFeretDiameter, MinFeretDiameter, maximum radius, mean radius, and median radius exhibit greater values in normal polygonal-shaped cervical cells, whereas they are

smaller for cells in the dysplastic category, such as LSIL, CIN-II, and CIN-III. These 39 listed features of cytoplasm and nucleus encompass the critical morphological attributes necessary for a comprehensive profile of individual cervical cells, thereby aiding in precise and accurate automated screening for cervical cancer.

Table 6: Various features extracted from cytoplasm and nucleus using cell-profiler.

Cytoplasmic Features
1. Cytoplasmic and Nucleus Area
2. Centre of Cytoplasm and Nucleus-axis
3. Centre of Cytoplasm and Nucleus Y-axis
4. Compactness of Cytoplasm and Nucleus
5. Eccentricity of Cytoplasm and Nucleus
6. Extent of Cytoplasm and Nucleus
7. Form factor of Cytoplasm and Nucleus
8. Major Axis Length of Cytoplasm and Nucleus
9. Maximum Radius of Cytoplasm and Nucleus
10. Mean Radius of Cytoplasm and Nucleus
11. Median Radius of Cytoplasm and Nucleus
12. Minor Axis Length of Cytoplasm and Nucleus
13. Cytoplasmic and Nucleus Orientation
14. Cytoplasmic and Nucleus Perimeter
15. Solidity of Cytoplasm and Nucleus
16. Location of Centre of Cytoplasm and Nucleus X-Axis
17. Location of Centre of Cytoplasm and Nucleus-Axis
18. Maximum Farthest Diameter of Cytoplasm and Nucleus
19. Minimum Farthest Diameter of Cytoplasm and Nucleus.
20. Cytoplasmic and Nucleus ratio

Table 7 provides the mapping of seven classes used for multi-class classification. Initially, there were twenty attributes associated with cytoplasm and nucleus in pap smear images. These attributes were combined into a single attribute identified by the primary key "Pap smear," which spanned across 843 entries.

Table 7: Cell mapping for cervical cell classification

Pap Smear	Cytoplasmic Attributes	Nucleus Attributes	Class
1	20 Attributes	20 Attributes	Class I
2	20 Attributes	20 Attributes	Class II
3	20 Attributes	20 Attributes	Class III

Pap Smear	Cytoplasmic Attributes	Nucleus Attributes	Class
4	20 Attributes	20 Attributes	Class IV
5	20 Attributes	20 Attributes	Class V
6	20 Attributes	20 Attributes	Class VI
7	20 Attributes	20 Attributes	Class VII

Following this consolidation, the data was processed based on these 40 attributes, resulting in the classification of instances into one of the seven predefined classes.

4.3 Evaluation Metrics

The proposed methodology is assessed through a range of metrics, including Intersection over Union (IoU), accuracy, specificity, sensitivity, and the F1 score. The IoU measures the degree of overlap between a predicted region and a ground truth region. In mathematical notation predicted segmentation (often referred to as "A") and the ground truth segmentation (often referred to as "B"). Here's what each part of the formula represents:$|A \cap B|$ is the size of the intersection between the predicted and ground truth segmentation masks. It measures the number of pixels that are correctly classified as part of the object in both the prediction and ground truth. $|A \cup B|$: The size of the union of the predicted and ground truth segmentation masks. It measures the total number of pixels that are classified as part of the object in either the prediction or the ground truth (or both). Accuracy quantifies the ratio of accurate predictions made by a model. Specificity is a performance metric used in binary classification tasks, particularly in medical and diagnostic applications. It evaluates the capability of a model to accurately detect negative instances (true negatives) among all real negative instances. Sensitivity, also known as recall, measures the true positive cases among all positive cases. The F1 score combines both precision and recall into a single score. Precision quantifies the proportion of true positive predictions among all the instances that the model predicted as positive. The definitions of these metrics are as under:

$$IoU\ (A,\ B) = |A \cap B|\ /\ |A \cup B| \qquad (1)$$

$$Accuracy = (Number\ of\ Correct\ Predictions)\ /\ (Total\ Number\ of\ Predictions) \quad (2)$$

$$Specificity = (True\ Negatives)\ /\ (True\ Negatives + False\ Positives) \qquad (3)$$

$$Sensitivity = (True\ Positives)\ /\ (True\ Positives + False\ Negatives) \qquad (4)$$

$$F1\ Score = 2 * (Precision * Recall) / (Precision + Recall) \qquad (5)$$

$$Precision = (True\ Positives) / (True\ Positives + False\ Positives) \qquad (6)$$

True positives (TP) refer to correctly identified instances of a condition or feature, indicating accurate detection by the diagnostic method. False positives (FP) occur when non-target instances are incorrectly labeled as positive, leading to potential overdiagnosis. False negatives (FN) represent missed detections, where actual instances are not identified, which can result in undetected conditions. Together, these metrics are essential for evaluating the performance of diagnostic tests, impacting clinical decision-making and patient outcomes.

4.4 Muti Class Classification

The confusion matrix provides a detailed assessment of a seven-class classification system that employs the VGG16 architecture for feature extraction and XGBoost as the classifier, evaluated on Pap smear images, which are vital for the early detection and diagnosis of cervical cancer. The analysis reveals specific challenges in classification performance among the different classes. Class I emerged as the most problematic, with 12 instances of misclassification, indicating potential complexities in accurately identifying this category. Class II followed closely with one error, suggesting that while it is generally well classified, some ambiguity remains. Classes III and IV each recorded four misclassifications, reflecting difficulties in distinguishing these categories from others. Conversely, Class V demonstrated flawless classification with no errors, indicating a strong model performance for this category. Class VI faced two misclassifications, which may point to overlaps with other classes, warranting further investigation. In contrast, Class VII exhibited remarkable accuracy, suggesting that the model effectively identifies this class with minimal errors. This high level of precision may be attributed to distinct morphological features that allow for clear differentiation from other classes. Such effective classification reinforces the model's potential in clinical settings, enhancing the reliability of cervical cancer diagnoses and contributing to improved patient outcomes through timely intervention.

Figure 30: Multiclass classification of segmentation free Pap smear images.

The confusion matrix presented in figure 31 is generated from the Deep-CervixNet model which highlights its efficacy in the realm of seven-class classification for cervical cancer diagnosis. With a meticulous examination of its performance, it becomes evident that this model surpasses its predecessors. Notably, classes I and II showcase flawless classification, indicating the model's robust ability to accurately identify these categories. Class III exhibits only two misclassifications, while class IV displays a marginal improvement with a total of three misclassifications. Class V, although presenting eight misclassifications, still demonstrates a notable advancement compared to previous models. Impressively, both class VI and class VII show minimal misclassifications, with zero and one respectively. These findings collectively underscore the superior performance of the Deep-CervixNet model compared to traditional approaches like VGG16 and XGBoost in the complex task of seven-class classification for cervical cancer diagnosis. The Deep-CervixNet model demonstrates enhanced accuracy and robustness in distinguishing between various classes, addressing the challenges faced by conventional methods. Its advanced architecture and feature extraction capabilities enable it to capture subtle differences in Pap smear images, leading to improved diagnostic precision. This improvement is crucial for effective early detection and treatment planning, ultimately contributing to better patient outcomes in cervical cancer management.
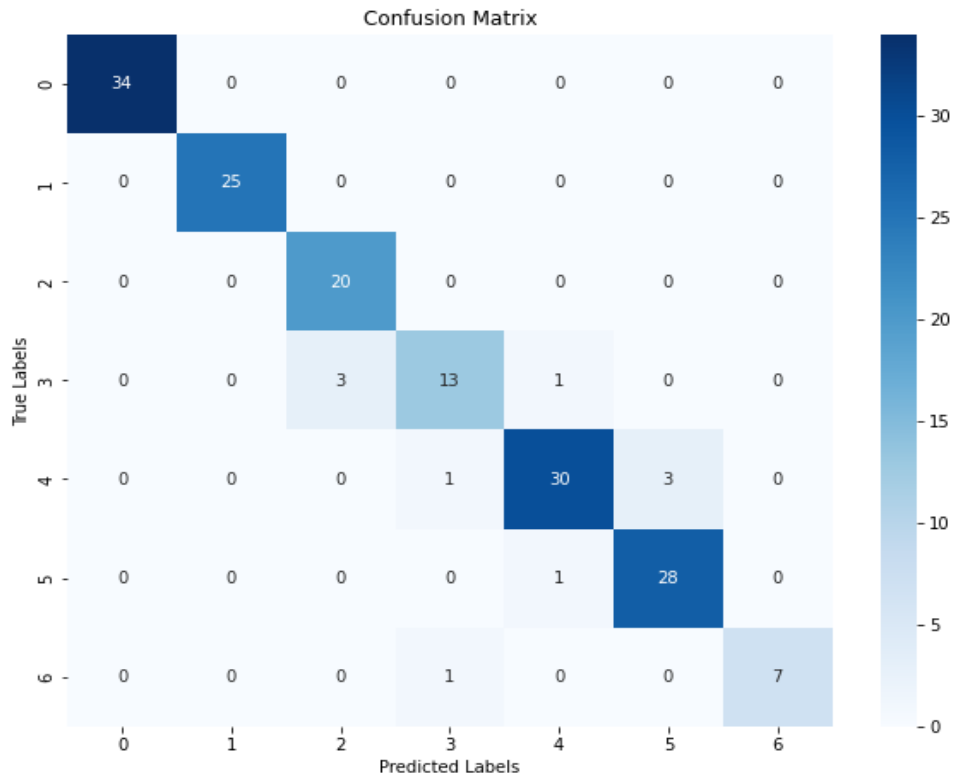
Figure 31: Multiclass classification of Deep-CervixNet architecture.

The performance evaluation of the seven classes, conducted using VGG16 as the feature extractor and XGBoost as the classifier, yields insightful metrics across various aspects of classification accuracy. For class I, an Accuracy of 0.65, Precision of 1, Recall of 0.65, Specificity of 1, and F1 score of 0.79 are observed. Class II exhibits impressive performance with an Accuracy of 0.96, Precision of 1, Recall of 0.96, Specificity of 1, and F1 score of 0.98. Similarly, class III displays notable scores, with an Accuracy of 0.80, Precision of 1, Recall of 0.80, Specificity of 1, and F1 score of 0.89. Class IV showcases an Accuracy of 0.76, Precision of 0.52, Recall of 0.76, Specificity of 0.92, and F1 score of 0.62, indicating some challenges in precise classification within this category. Class V demonstrates perfect performance across all metrics, including Accuracy, Precision, Recall, Specificity, and F1 score, all at 1. Class VI achieves an Accuracy of 0.93, Precision of 1, Recall of 0.93, Specificity of 1, and F1 score of 0.96, indicating robust classification capability. Lastly, class VII presents an Accuracy of 1, Precision of 0.42, Recall of 1, Specificity of 0.93, and F1 score of 0.59, suggesting room for improvement particularly in precision. These metrics collectively provide a comprehensive understanding of the model's performance across diverse classes in cervical cancer diagnosis, offering valuable insights into its efficacy and areas for refinement. The summary of these classes is presented in table 8.

Table 8: Performance analysis of seven classes of cervical cells classified using VGG16 and XGboost.

Class	Accuracy	Precision	Recall	Specificity	F1 score
Class I	0.65	1	0.65	1	0.79
Class II	0.96	1	0.96	1	0.98
Class III	0.80	1	0.80	1	0.89
Class IV	0.76	0.52	0.76	0.92	0.62
Class V	1	1	1	1	1
Class VI	0.93	1	0.93	1	0.96
Class VII	1	0.42	1	0.93	0.59

The reported performance metrics in table 9, including Accuracy, Precision, Recall, Specificity, and F1 score, were achieved through Deep-CervixNet, applied to Pap smear images for cervical cancer diagnosis. Specifically, the evaluation yielded the following results across different classes: For class I, an Accuracy of 1, Precision of 1, Recall of 1, Specificity of 1, and F1 score of 1 were observed. Class II exhibited an impressive performance with an Accuracy of 1, Precision of 1, Recall of 1, Specificity of 1, and F1 score of 1. Similarly, class III displayed notable scores, with an Accuracy of 1, Precision of 0.87, Recall of 1, Specificity of 0.98, and F1 score of 0.93. Class IV showcased an Accuracy of 0.76, Precision of 0.87, Recall of 0.76, Specificity of 0.99, and F1 score of 0.81. Class V demonstrated Accuracy of 0.88, Precision of 0.94, Recall of 0.88, Specificity of 0.98, and F1 score of 0.91. Class VI achieved an Accuracy of 0.97, Precision of 0.90, Recall of 0.97, Specificity of 0.98, and F1 score of 0.93. Finally, class VII achieves an accuracy of 0.88, precision of 1, recall of 0.88, specificity of 1, and F1 score of 0.93. These comprehensive metrics revealed that Deep-CervixNet outperformed than VGG16-XGBoost model for cervical cancer diagnosis. The VGG16-XGBoost method is a segmentation-free classification approach for Pap smear images. In this method, the VGG16 model, pre-trained on the ImageNet dataset, is used for feature extraction, followed by multi-class classification using XGBoost. In contrast, Deep CervixNet employs semantic segmentation of the cytoplasm and nucleus. Features are then extracted from both the cytoplasm and nucleus to classify the images into seven classes. Thus, it is evident that the classification of Pap smear images is enhanced using segmentation compared to segmentation-free classification.

Table 9: Performance analysis using Deep-CervixNet classification of Pap smear images.

Class	Accuracy	Precision	Recall	Specificity	F1 score
Class I	1	1	1	1	1
Class II	1	1	1	1	1
Class III	1	0.87	1	0.98	0.93
Class IV	0.76	0.87	0.76	0.99	0.81
Class V	0.88	0.94	0.88	0.98	0.91
Class VI	0.97	0.90	0.97	098	0.93
Class VII	0.88	1	0.88	1	0.93

Comparing the performance of VGG16 and XGBoost against Deep-CervixNet in the context of cervical cancer diagnosis reveals nuanced differences in classification efficacy. VGG16, coupled with XGBoost, demonstrates commendable performance across various metrics, accurately classifying instances within different classes. However, when contrasted with Deep-CervixNet, a more specialized architecture tailored specifically for cervical cancer diagnosis, distinct advantages emerge. Deep-CervixNet showcases superior precision, recall, and F1 scores across multiple classes, indicating its robustness in identifying intricate features indicative of cervical cancer. While VGG16 and XGBoost offer a general-purpose framework, Deep-CervixNet's specialized design allows it to leverage domain-specific insights and nuances, leading to enhanced classification accuracy and diagnostic reliability. Moreover, Deep-CervixNet's tailored architecture potentially streamlines the feature extraction and classification processes, optimizing performance for the specific task at hand. Overall, while VGG16 and XGBoost provide a solid foundation for classification tasks, Deep-CervixNet's specialized design and superior performance metrics underscore its potential as a promising tool in cervical cancer diagnosis. Further exploration and validation are warranted to fully ascertain its clinical utility and integration into diagnostic workflows. The table 10 mentioned below is the comparison of Deep-CervixNet and VGG16-XGBoost for cervical cancer diagnosis. Acc, as used herein, stands for accuracy of the Deep-CervixNet Model and VGG16-XGBoost for multiclass classification of cervical cells.

Table 10: Average performance comparison of Deep-CervixNet Model and VGG16-XGBoost.

	Deep-CervixNet Model					VGG16-XGBoost				
Class	Acc	Precision	Recall	Specificity	F1 score	Acc	Precision	Recall	Specificity	F1 score
Class I	1	1	1	1	1	0.65	1	0.65	1	0.79
Class II	1	1	1	1	1	0.96	1	0.96	1	0.98
Class III	1	0.87	1	0.98	0.93	0.80	1	0.80	1	0.89
Class IV	0.76	0.87	0.76	0.99	0.81	0.76	0.52	0.76	0.92	0.62
Class V	0.88	0.94	0.88	0.98	0.91	1	1	1	1	1
Class VI	0.97	0.90	0.97	098	0.93	0.93	1	0.93	1	0.96
Class VII	0.88	1	0.88	1	0.93	1	0.42	1	0.93	0.59

Figure 32 depicts the training and validation accuracy, as well as the training and validation loss, of the proposed approach executed over 100 epochs. This visual representation offers valuable insights into the model's performance and learning dynamics throughout the training process. Moreover, the proposed model was initially validated using the holdout validation approach. However, due to certain limitations, an additional external validation was conducted using independent Pap smear images, which were not employed during model training or testing. This process provided a more rigorous assessment of the model's generalization ability. By observing trends in accuracy, one can discern the model's capacity to correctly classify instances in both the training and validation datasets across successive epochs. Similarly, the fluctuations in loss provide a measure of the model's convergence and optimization progress, with lower values indicating improved fitting to the data. Analyzing these metrics collectively offers a comprehensive understanding of the proposed approach's training behavior, aiding in fine-tuning parameters and optimizing performance for the task at hand.
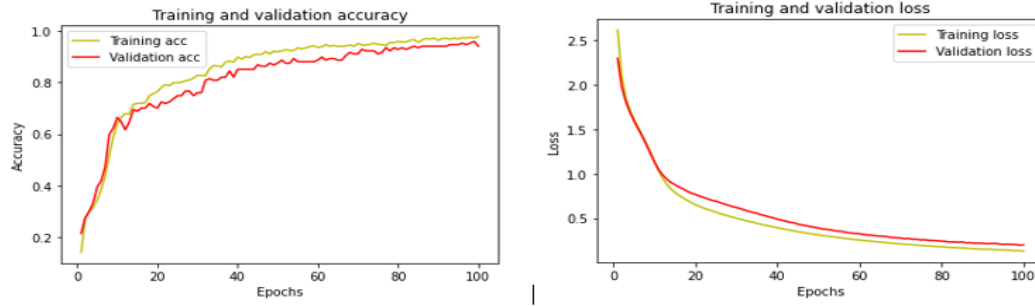
Figure 32:Training and validation accuracy and loss graphs for Deep-CervixNet.

4.5 Profiling of Folded Cytoplasm

The proposed approach presents a promising methodology for profiling folded cytoplasm in a comprehensive three-fold process. This innovative method involves segmenting the folded area, semi-stitching with Fiji software, and effectively profiling the features within the cytoplasm, marking a significant stride in cellular analysis, and understanding. The initial step of this approach is based on precise segmentation of the folded cytoplasmic areas. This crucial stage demands meticulous delineation and isolation of these intricate folded regions within cellular images. Leveraging advanced image processing algorithms and techniques, the segmentation process identifies and separates the folded cytoplasm, forming the foundation for subsequent analysis. Following segmentation, the semi-stitching process, facilitated by Fiji software, is employed to assemble, and reconstruct the segmented folded areas. Fiji, known for its robust capabilities in image processing and analysis, facilitates the semi-stitching by integrating segmented regions, creating a cohesive visual representation of the previously isolated folded cytoplasmic areas. This semi-stitching stage aims to create a coherent and comprehensive visualization for further analysis. The final stage of this approach revolves around effectively profiling the features within the folded cytoplasmic regions. This pivotal step involves a detailed analysis of the reconstructed areas, extracting and characterizing key features present within these folded regions. By meticulously examining the structural attributes, morphological nuances, and textural intricacies, this profiling process provides comprehensive insights into the nature and composition of the folded cytoplasm. This three-fold approach offers a novel and systematic methodology to delve into the complexities of folded cytoplasmic areas within cellular structures. Its segmentation, semi-stitching, and feature profiling collectively aim to unravel the intricacies and nuances inherent in these folded regions, enabling a deeper understanding of cellular structures and functions. Moreover, the application potential of this approach extends across various domains, including but not limited

to medical research, cellular biology, and pathological studies. By providing a structured framework to profile folded cytoplasm, this approach holds promise in elucidating the significance and implications of these structures in cellular dynamics, disease pathology, and functional analyses. In essence, the proposed approach presents a systematic and innovative methodology to profile folded cytoplasm, integrating segmentation, semi-stitching, and detailed feature profiling. Its application heralds a significant advancement in the understanding of cellular complexities, paving the way for enhanced insights into cellular structures and functions across diverse scientific domains.

In the realm of biomedical imaging, a nuanced and intricate challenge lies in deciphering the intricate folds of cytoplasmic structures within microscopic images. This study endeavors to address this complexity through a meticulously designed four-stage unfolding process. At the forefront of this methodology is the implementation of a Residual Attention U-Net, a specialized neural network tailored for image segmentation. This initial stage involves training the network to discern and precisely delineate the folded regions within cytoplasmic structures. The efficacy of this segmentation process is paramount in isolating the areas of interest, as it lays the groundwork for subsequent analyses. Following the precise segmentation, the unfolding process advances with the integration of a Fiji stitching module, renowned for its sophisticated approach to image stitching and blending. This module operates on multiple steps, notably including the identification of keypoints, serving as crucial reference points for seamlessly merging the segmented portions. The culmination of this step is the creation of a cohesive and unfolded representation of the cytoplasmic structure. This unfolded image, now free from the complexities of folding, becomes the canvas for the extraction of meaningful features. Feature extraction, the cornerstone of subsequent analyses, is executed with precision on the unfolded image. Leveraging advanced segmentation techniques, this stage ensures that only the most relevant and biologically significant information is extracted. The objective is to highlight the intricate details and patterns from the unfolded cytoplasmic structures, providing a foundation for a deeper understanding of cellular dynamics and functions. This comprehensive unfolding strategy serves as a critical enabler for sophisticated image analysis. By optimizing the accuracy of feature extraction, it enhances the overall effectiveness of the methodology. The elimination of the intricacies associated with folded structures allows for a more focused exploration of the biological intricacies inherent in cytoplasmic components. The insights garnered through this process contribute not only to the advancement of biomedical imaging techniques but also hold the potential to shed light on fundamental cellular processes,

offering a richer and more nuanced perspective in the realm of life sciences. In summary, this research endeavors to unravel the complexities of cytoplasmic structures, presenting a methodological framework that transcends the confines of folded imagery to reveal the hidden intricacies within the microscopic landscape. The below mentioned figure 33(a) represents the folded cytoplasm, (b) segmented folded part, (c)ground truth of folded cells (d) stitched Pap smear cell, (e) unfolded cytoplasm. Figure 34 represents the folded cytoplasm of Intermediate Squamous Cells (ISC), (b) segmented folded part, (c)ground truth of folded cells (d) stitched Pap smear,

e) Unfolded cytoplasm.
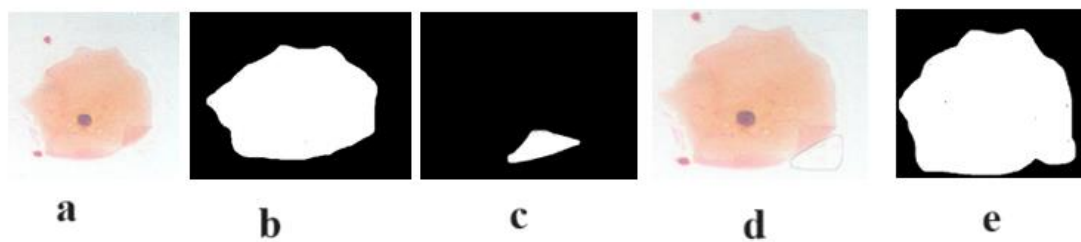
Figure 33: Unfolding of Intermediate Squamous images, a) Input image b) Ground truth image c) Segmented ROI d) stitched ROI part with input image e) Unfolded cytoplasm.
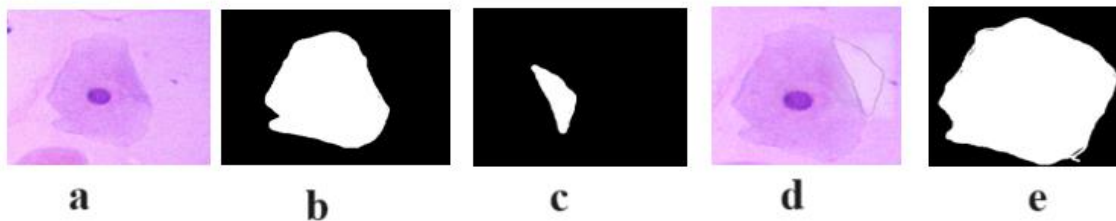
Figure 34. Unfolding of Superficial Squamous cells folded cytoplasm a) Input image b) Ground truth image c) Segmented ROI d) stitched ROI part with input image e) Unfolded cytoplasm.

Table 11: Range difference in folded and unfolded cytoplasmic features.

Features	Folded cell features	Unfolded cytoplasmic features	Difference
Cytoplasmic Area	105573	125231	19658
Centre of Cytoplasm X-axis	459	500	41
Centre of Cytoplasm Y-axis	517	517	0
Compactness of Cytoplasm	1.358512589	1.252156806	-0.1

Features	Folded cell features	Unfolded cytoplasmic features	Difference
Eccentricity of Cytoplasm	0.684710433	0.516674081	0.17
Extent of Cytoplasm	0.703632365	0.735917024	0.03
Form factor of Cytoplasm	0.736099178	0.798622022	0.06
Major Axis Length of Cytoplasm	437.7833183	435.0121187	-2
Maximum Radius of Cytoplasm	147.8715659	173.4041522	28
Mean Radius of Cytoplasm	53.2711685	61.18841165	8
Median Radius of Cytoplasm	47.80167361	54.12947441	7
Minor Axis Length of Cytoplasm	319.063145	372.4496162	53
Cytoplasmic Orientation	9.358100784	-2.936661335	6.4
Cytoplasmic Perimeter	1342.496608	1403.751442	61
Solidity of Cytoplasm	0.963732131	0.968807779	0.005
Location of Centre of Cytoplasm X-Axis	296.8549629	313.0198833	17
Location of Centre of Cytoplasm Y-Axis	310.6512839	303.1752601	-7
Maximum Farthest Diameter of Cytoplasm	445.4267617	450.4686449	5
Minimum Farthest Diameter of Cytoplasm	323.8943178	362.6801896	39
Nuclear cytoplasmic ratio	245.354	287.193	42.559

4.6 Comparison with State-of Art Approaches

Table 12 presents a comparative analysis of the proposed methodology for diagnosing cervical cancer alongside other state-of-the-art approaches. This table allows for a thorough examination of the performance metrics, such as sensitivity, specificity, and accuracy, of each method. By scrutinizing these key indicators, researchers and practitioners can gain valuable insights into the strengths and weaknesses of various diagnostic techniques. This comparison

not only highlights the effectiveness of the proposed methodology but also provides valuable context within the broader landscape of cervical cancer diagnosis.

Alyafeai et al. [75] presented a pipeline to integrate two pre-trained deep learning models for cervix detection and tumor classification. The first model achieves cervix region detection significantly faster than state-of-the-art models, with an IoU accuracy of 0.68. The second model utilizes self-extracted features for tumor classification, leveraging lightweight CNNs. This classifier surpasses existing models in both classification accuracy and speed. The model also achieves a Recall of 59.70±12.08, followed by specificity of 77.43±10.57.

Haraz et al. [21] employed machine learning algorithms to classify cervical cancer cells from Pap smears into five distinct types using the SIPaKMeD database. The dataset comprises 4045 isolated Pap smear cells categorized by experts into superficial-intermediate, Parabasal, Koilocytotic, Dyskeratotic, and Metaplastic cells. A pipeline was introduced to enhance algorithm accuracy and ease of implementation by employing a specified feature extractor and suitable preprocessing steps. The study concluded that machine learning can enhance Pap smear screening results, with SVM achieving the highest accuracy (0.968), followed by Neural Network at 0.958 and KNN at 0.941. The model also achieved an overall score of 0.968, with a Precision of 0.968, Recall of 0.968, and Specificity of 0.992 for cervical cancer diagnosis.

Kalbhor et al. [16] introduced a novel hybrid approach that integrates deep learning architectures with machine learning classifiers and fuzzy min-max neural networks for feature extraction and Pap smear image classification, respectively. Pretrained deep learning models, including AlexNet, ResNet-18, ResNet-50, and GoogleNet, are utilized. Benchmark datasets from Herlev and SIPaKMeD are employed for experimentation. The highest classification accuracy of 95.33% is achieved with the fine-tuned ResNet-50 architecture, followed by AlexNet, on the SIPaKMeD dataset. Based on the findings, SVM achieved superior performance metrics compared to NN and KNN. SVM demonstrated an F1 score, recall, specificity, and AUC of 0.968, 0.968, 0.992, and 0.998, respectively, with a low log loss of 0.103. In contrast, Neural Network yielded an AUC of 0.997, precision of 0.958, F1 score, recall, and log loss of 0.958, 0.958, and 0.141, respectively, with specificity at 0.989. Meanwhile, KNN exhibited an AUC of 0.993, F1 score, recall, and specificity of 0.941, 0.941, and 0.985, respectively, with a higher log loss of 0.340.

Table 12: Comparison of Deep-CervixNet with state-of-art approaches.

Author	Dataset	Technique	IoU	Acc	Precision	Recall	F1 score	Specificity	Acc
Alyafeai et al. [75]	Cervi gram images	CNN	0.68	68.25 ±9.74	-	59.70± 12.08	-	77.43±10 .57	-
Haraz et al. [21]	SIPaKM eD	SVM	-	-	0.958	0.958	0.958	0.989	96.8 %
Kalbhor et al. [16]	Herlev and SIPaKM eD	AlexNet, ResNet-18, ResNet-50, and GoogleNet	-	-	0.958	0.968	0.968	0.992	95.33 %
Proposed Method	Primary dataset	Deep Cervix Net	97.36	99.8 %	0.94	0.94	0.99	0.99	98%
	Herlev dataset		0.95	0.91	0.95	0.95	0.95	0.95	95.5 %

From the experimental results it was observed the Deep Cervix Net achieved an accuracy of 98%, demonstrated remarkable precision, correctly identifying 94% of positive cases, and achieving a high recall rate of 94%, ensuring minimal false negatives. Moreover, its exceptional specificity of 99% indicated its ability to accurately identify negative cases. The F1 score of 94% further underscores the model's effectiveness in achieving a balance between precision and recall. The results clearly show that the Deep Cervix Net outperforms other methods in several key areas. It's not just more accurate; it's also better at precision, recall, and specificity. This means it's good at correctly identifying both positive and negative cases, minimizing errors. Overall, these findings suggest that the Deep Cervix Net represents a significant step forward in diagnosing cervical cancer. It offers promise for improving how to detect this disease, ultimately leading to better outcomes for patients. The Deep Cervix Net's advanced capabilities make it a viable option for underdeveloped countries with limited resources, potentially reducing the economic burden of late-stage cervical cancer detection and treatment. The proposed model, evaluated on the Herlev dataset, shows approximate values of IoU 0.95, Segmentation Accuracy 95%, Classification Accuracy 95.8%, Precision 95.5%, Recall 95.5%, F1 Score 95.5%, and Specificity 95.5%.

CHAPTER 5

CONCLUSION AND FUTURE SCOPE

This chapter focuses on the conclusion and future scope of "Development of Deep Learning Based Approach for Early Diagnosis of Cervical Cancer". It serves as a comprehensive consolidation of the insights garnered throughout this research endeavor, outlining key takeaways and future directions in the realm of cervical cancer detection.

5.1 Conclusion

Cervical cancer remains a significant global health concern, necessitating precise and timely detection methods to mitigate its impact. The development of DeepCervix-Net represents a groundbreaking advancement in this area, introducing an AI-based system meticulously engineered to automate the intricate analysis of cervical cell images. This approach, employing a meticulously structured three-stage process - semantic segmentation, feature extraction, and classification stands as a beacon of promise in revolutionizing early detection strategies and subsequent medical interventions. At its core, DeepCervix-Net's strength lies in its meticulous approach. Through semantic segmentation, leveraging a modified U-Net architecture, DeepSeg-Net, it precisely isolates cell structures, providing a foundation for subsequent analysis. Subsequently, employing Cell Profiler, the system extracts intricate features that enrich the classification task, enabling the identification of potential cancerous and precancerous cells. The final stage, employing the innovative Artificial Neural Network Architecture model, ensures precise classification aligned with the latest Bethesda System of Classification, facilitating early detection and informed decision-making. The efficacy of DeepCervix-Net is evident in its exceptional overall accuracy rate of 98%, surpassing existing approaches. This approach not only enhances efficiency but also offers a transformative potential to reshape cervical cancer detection. By harnessing the capabilities of deep learning techniques and deploying a comprehensive multi-stage approach, DeepCervix-Net paves the way for a future where mortality rates can potentially decrease, offering improved health outcomes and a better quality of life for women worldwide. As research in this field evolves, the integration of AI-driven advancements, as demonstrated by DeepCervix-Net, promises a brighter future in the fight against cervical cancer. It brings renewed hope for improved diagnostic accuracy and better-informed medical interventions, ultimately aiming to reduce the burden of this disease on a global scale. DeepCervix-Net stands not just as a technological innovation but as a potential transformative approach in the ongoing battle against cervical cancer, promising a future where early detection becomes more precise, accessible, and

lifesaving for countless individuals worldwide. However, this work encounters several challenges and limitations: generating extensive ground truth images for each cytoplasm and nucleus is labour-intensive and requires expert knowledge, while high computational power is needed for model training. Additionally, the complexity of double segmentation processes to accurately handle folded parts of the cytoplasm adds to the computational demands. A significant challenge is the careful delineation of nuclear boundaries, which is crucial for effective segmentation but complicated by variability in nuclear shapes, overlapping structures, and image quality. Ensuring precise and consistent annotations is time-consuming and impacts the reliability of the training data, affecting the model's performance. Furthermore, while the model's generalizability was evaluated through the benchmark Herlev dataset, achieving an efficiency of 95.8%.

5.2 Future Scope

- This study can be extended to other diseases as well. For such extension, relevant factors need to be identified after detailed study of concerned literature and consultation with medical experts.

- The study reported in this thesis can be further improved by enriching the database used for training of proposed artificial neural network, to include a greater number of clinical cases so that the database becomes more diverse.

- The work presented in this thesis is limited to cytological features of cervical cells, however if some genetic features are also incorporated in the study, the overall results can become more robust.

- The Deep-CervixNet model proposed in this work can be embedded into a hardware device to makeup a portable cervical cancer screening tool.

- This work can be tested on diverse datasets such as the Cervical Cancer Risk Prediction Database (CRIC), the International Symposium on Biomedical Imaging 2015 Challenge (ISBI 2015), and the Breast Cancer Histology Challenge Dataset (BACH) to evaluate its robustness and improve its applicability, particularly in cervical cancer diagnosis.

- Future work will focus on real-world implementation, including integration into healthcare systems, addressing computational challenges, overcoming real-time image capture difficulties, and ensuring scalability and cost-effectiveness for clinical use

- The presented unfolded cytoplasm is limited to semi-stitching, in future it can be further improved to fully automatic stitching to profile the cytoplasmic features.

5.3 SUMMARY

This section is not just a conclusion but a stepping stone for future research and exploration. It seeks to inspire ongoing efforts, driving innovation and advancements in the use of cutting-edge technologies for the early diagnosis of cervical cancer. By addressing the challenges identified in current methodologies and exploring the untapped potential of new technologies, the goal is to significantly improve the efficacy and accessibility of diagnostic tools, ultimately leading to better healthcare outcomes in cervical cancer diagnosis. Efforts could focus on making these advanced diagnostic tools more accessible to low-resource settings, where cervical cancer incidence and mortality rates are often highest, by developing cost-effective and user-friendly diagnostic platforms. Interdisciplinary collaboration among cytologists, oncologists, data scientists, and engineers will be crucial in driving innovations, addressing ethical, regulatory, and implementation challenges. This collaborative approach will accelerate the development of sophisticated diagnostic systems, paving the way for a future where cervical cancer diagnosis is more accurate, efficient, and accessible, ultimately aiming to reduce the global burden of cervical cancer through early detection and treatment.

LIST OF PUBLICATIONS

- N. Nazir, B. S. Saini, and A. Sarwar, "A Novel Approach in Breast Cancer Diagnosis with Image Processing," in Proceedings of the International Conference on Intelligent Vision and Computing, pp. 468-476, 2022.

- N. Nazir, B. S. Saini, and A. Sarwar, "Early Diagnosis of Cervical Cancer Using AI: A Review," in Proceedings of International Conference on Recent Innovations in Computing, pp. 105-116, 2023.

- N. Nazir, A. Sarwar, B. S. Saini, and R. Shams, "A Robust Deep Learning Approach for Accurate Segmentation of Cytoplasm and Nucleus in Noisy Pap Smear Images," Computation, vol. 11, no. 10, pp. 1-20, 2023.

- N. Nazir, A. Sarwar, and B. S. Saini, "Recent developments in denoising medical images using deep learning: An overview of models, techniques, and challenges," Micron, vol. 180, pp. 103615, 2024.

BIBLIOGRAPHY

[1]   R. L. Siegel, K. D. Miller, N. S. Wagle, and A. Jemal, "Cancer statistics, 2023," CA Cancer J. Clin, vol. 73, no. 1, pp. 17–48, 2023.

[2]   A. Ghaznavi, R. Rychtáriková, M. Saberioon, and D. Štys, "Cell segmentation from telecentric bright-field transmitted light microscopy images using a Residual Attention U-Net: A case study on HeLa line," Computers in Biology and Medicine, vol. 147, no. 105805, pp. 105805-105837, 2022.

[3]   S. Y. Chung, J. S. Chang, and Y. B. Kim, "Comprehensive clinical evaluation of deep learning-based auto-segmentation for radiotherapy in patients with cervical cancer," Frontiers in Oncology, vol. 13, no. 4, pp. 1–11, 2023.

[4]   J. R. Meyer-Arendt and D. M. Humphreys, "Quantitative morphology of cancer cells," Acta Histochemica, vol. 44, no. 1, pp. 41-48, 1972.

[5]   G. Naghdy, M. Ross, C. Todd, and E. Norachmawati, "Classification cervical cancer using histology images," in 2010 Second International Conference on Computer Engineering and Applications, vol. 1, pp. 515-519, IEEE, Mar. 2010.

[6]   F. H. Araújo, R. R. Silva, D. M. Ushizima, M. T. Rezende, C. M. Carneiro, A. G. C. Bianchi, and F. N. Medeiros, "Deep learning for cell image segmentation and ranking," Computerized Medical Imaging and Graphics, vol. 72, pp. 13-21, 2019.

[7]   L. Wen, J. Xiao, J. Zeng, C. Zu, X. Wu, J. Zhou, X. Peng, and Y. Wang,"Multi-level progressive transfer learning for cervical cancer dose prediction," Pattern Recognition, vol. 141, pp. 109606, 2023.

[8]   L. Xu, F. Cai, Y. Fu, and Q. Liu, "Cervical cell classification with deep-learning algorithms," Medical & Biological Engineering & Computing, vol. 61, no. 3, pp. 821-833, 2023.

[9]   S. J. Bowden, T. Doulgeraki, E. Bouras, G. Markozannes, A. Athanasiou, H. Grout-Smith, K. S. Kechagias, L. B. Ellis, V. Zuber, M. Chadeau-Hyam, J. M. Flanagan, K. K. Tsilidis, I. Kalliala, and M. Kyrgiou,"Risk factors for human papillomavirus infection,

cervical intraepithelial neoplasia and cervical cancer: an umbrella review and follow-up Mendelian randomisation studies, BMC Medicine" , vol. 21, no. 1, pp. 274, 2023.

[10] T. Chankong, N. Theera-Umpon, and S. Auephanwiriyakul, "Automatic cervical cell segmentation and classification in Pap smears, Computer Methods and Programs in Biomedicine", vol. 113, no. 2, pp. 539-556, 2014.

[11] G. Kumawat, S. K. Vishwakarma, P. Chakrabarti, P. Chittora, T. Chakrabarti, and J. C.-W. Lin, "Prognosis of Cervical Cancer Disease by Applying Machine Learning Techniques, Journal of Circuits, Systems and Computers", vol. 31, no. 6, pp. 2350019, 2023.

[12] N. Nazir, A. Sarwar, and B. S. Saini, "Recent developments in denoising medical images using deep learning: An overview of models, techniques, and challenges, Micron," vol. 180, pp. 103615, 2024.

[13] Y. Song, L. Zhang, S. Chen, D. Ni, B. Lei, and T. Wang, "Accurate segmentation of cervical cytoplasm and nuclei based on multiscale convolutional network and graph partitioning, IEEE Transactions on Biomedical Engineering," vol. 62, no. 10, pp. 2421-2433, 2015.

[14] K. Zhang, K. Sun, C. Zhang, K. Ren, C. Li, L. Shen, and D. Jing, "Using deep learning to predict survival outcome in non-surgical cervical cancer patients based on pathological images," Journal of Cancer Research and Clinical Oncology, vol. 149, no. 9, pp. 6075–6083, 2023.

[15] Y. F. Chen, P. C. Huang, K. C. Lin, H. H. Lin, L. E. Wang, C. C. Cheng, T. P. Chen, Y. K. Chan, and J. Y. Chiang, "Semi-automatic segmentation and classification of Pap smear cells," IEEE Journal of Biomedical and Health Informatics, vol. 18, no. 1, pp. 94–108, 2014.

[16] M. Kalbhor, S. Shinde, D. E. Popescu, and D. J. Hemanth, "Hybridization of Deep Learning Pre-Trained Models with Machine Learning Classifiers and Fuzzy Min–Max Neural Network for Cervical Cancer Diagnosis," Diagnostics, vol. 13, no. 7, pp. 1-16,2023.

[17]   N. Nazir, A. Sarwar, B. S. Saini, and R. Shams, "A Robust Deep Learning Approach for Accurate Segmentation of Cytoplasm and Nucleus in Noisy Pap Smear Images," *Computation*, vol. 11, no. 10, pp.1-20, 2023.

[18]   E. Gomez-de-Mariscal, M. Maska, A. Kotrbova, V. Pospichalova, P. Matula, and A. Munoz-Barrutia, "Deep-Learning-Based Segmentation of Small Extracellular Vesicles in Transmission Electron Microscopy Images," Scientific Reports, vol. 9, pp. 13211, 2019.

[19]   X. Li, M. Du, S. Zuo, M. Zhou, Q. Peng, Z. Chen, J. Zhou, and Q. He, "Deep convolutional neural networks using an active learning strategy for cervical cancer screening and diagnosis," Frontiers in Bioinformatics, vol. 3, pp. 1101667, Mar. 2023.

[20]   C. Li, D. Xue, X. Zhou, J. Zhang, H. Zhang, Y. Yao, F. Kong, L. Zhang, and H. Sun, "Transfer learning-based classification of cervical cancer immunohistochemistry images," 3rd International Symposium on Image Computing and Digital Medicine, pp. 102-106, 2023.

[21]   A. Haraz, H. E. D. Moustafa, A. T. Khaleel, and A. H. Eltanboly, "Deep Learning-Based Technique for The Perception of The Cervical Cancer," Mansoura Engineering Journal, vol. 48, no. 3, pp. 3, 2023.

[22]   L. W. Habtemariam, E. T. Zewde, and G. L. Simegn, "Cervix Type and Cervical Cancer Classification System Using Deep Learning Techniques," Medical Devices: Evidence and Research, vol. 15, no. June, pp. 163–176, 2022.

[23]   M. Rahimi, A. Akbari, F. Asadi, and H. Emami, "Cervical cancer survival prediction by machine learning algorithms: a systematic review," BMC Cancer, pp. 1–10, 2023, .

[24]   A. Sarwar, V. Sharma, and R. Gupta, "Hybrid ensemble learning technique for screening of cervical cancer using Papanicolaou smear image analysis," Personalized Medicine Universe, vol. 4, pp. 54–62, 2015.

[25]   O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in Lecture Notes in Computer Science, Cham: Springer International Publishing, pp. 234–241, 2015.

[26] L. Pantanowitz and M. M. Bui, "Computer-Assisted Pap Test Screening," Monographs in Clinical Cytology, vol. 25, pp. 67–74, 2020.

[27] S. Jeong, H. Yu, S. Park, D. Woo, S. Lee, G. Chong, H. Han, and J. Kim, "Comparing deep learning and handcrafted radiomics to predict chemoradiotherapy response for locally advanced cervical cancer using pretreatment MRI," Scientific Reports, vol. 24, pp. 1–11, 2024.

[28] G. Yang, J. Huang, Y. He, Y. Chen, T. Wang, C. Jin, and P. Sengphachanh, "GCP-Net: A Gating Context-Aware Pooling Network for Cervical Cell Nuclei Segmentation," Mobile Information Systems, vol. 2022, pp. 1-14, 2022.

[29] S. M. Abd-Alhalem, H. S. Marie, W. El-Shafai, T. Altameem, R. S. Rathore, and T. M. Hassan, "Cervical cancer classification based on a bilinear convolutional neural network approach and random projection," Engineering Applications of Artificial Intelligence, vol. 127, pp. 107261, 2024.

[30] D. B. Patil and T. S. Vishwanath, "Integration of Deep Learning Algorithms for Precision Cervical Cancer Analysis from Colposcopic Images," Intelligent Systems and Applications in Engineering, vol. 12, pp. 539–545, 2024.

[31] T. Aljrees, "Improving prediction of cervical cancer using KNN imputer and multi-model ensemble learning," Plos One, vol. 19, no. 1, pp.1-24, 2024.

[32] D. B. Patil and T. S. Vishwanath, "Automated Lesion Grading and Analysis of Uterine Cervix Images for Cervical Cancer Diagnosis," International Journal of Intelligent Systems and Applications in Engineering, vol. 12, no. 1, pp. 577-590, 2024.

[33] P. Jiang, X. Li, H. Shen, Y. Chen, L. Wang, H. Chen, J. Feng, and J. Liu, "A systematic review of deep learning-based cervical cytology screening: from cell identification to whole slide image analysis," Artificial Intelligence Review, vol. 56, pp. 2687-2758, 2023.

[34] Duru, G., and Topatan, S., "A barrier to participation in cervical cancer screenings: fatalism," Women & Health, vol. 63, no. 6, pp. 436-444, 2023. Available: https://pubmed.ncbi.nlm.nih.gov/37303197.

[35] S. P. Oliveira, D. Montezuma, A. Moreira, D. Oliveira, P. C. Neto, A. Monteiro, J. Monteiro, L. Ribeiro, S. Goncalves, I. M. Pinto, and J. S. Cardoso, "A CAD system for automatic dysplasia grading on H&E cervical whole-slide images," Scientific Reports, vol. 13, no. 1, pp. 1–12, 2023.

[36] M. M. Rahaman, C. Li, Y. Yao, F. Kulwa, X. Wu, X. Li, and Q. Wang, "DeepCervix: A deep learning-based framework for the classification of cervical cells using hybrid deep feature fusion techniques," Computers in Biology and Medicine, vol. 136, pp. 104649, 2021.

[37] I. Pacal, "MaxCerVixT: A novel lightweight vision transformer-based approach for precise cervical cancer detection," Knowledge-Based Systems, vol. 289, pp. 111482, 2024.

[38] Y. M. Lee, B. Lee, N. H. Cho, and J. H. Park, "Beyond the Microscope: A Technological Overture for Cervical Cancer Detection," Diagnostics, vol. 13, no. 19, pp. 3079, 2023.

[39] J. Ji, W. Zhang, Y. Dong, R. Lin, Y. Geng, and L. Hong, "Automated cervical cell segmentation using deep ensemble learning," BMC Medical Imaging, vol. 23, no. 1, pp. 1–10, 2023.

[40] S. Malik, R. Sah, K. Muhammad, and Y. Waheed, "Tracking HPV infection, associated cancer development, and recent treatment efforts-A comprehensive review," Vaccines, vol. 11, no. 1, pp. 102, 2023.

[41] S. J. Bowden, T. Doulgeraki, E. Bouras, G. Markozannes, A. Athanasiou, H. Grout-Smith, K. S. Kechagias, L. B. Ellis, V. Zuber, M. Chadeau-Hyam, J. M. Flanagan, K. K. Tsilidis, I. Kalliala, and M. Kyrgiou, "Risk factors for human papillomavirus infection, cervical intraepithelial neoplasia and cervical cancer: an umbrella review and follow-up Mendelian randomisation studies," BMC Med., vol. 21, no. 1, pp. 274, 2023.

[42] H. W. Haverkos, G. Soon, S. L. Steckley, and W. Pickworth, "Cigarette smoking and cervical cancer: part I: a meta-analysis," Biomedicine & Pharmacotherapy, vol. 57, no. 2, pp. 67–77, 2003.

[43] J. P. Avila, B. M. Carvalho, and E. C. Coimbra, "A Comprehensive View of the Cancer-Immunity Cycle (CIC) in HPV-Mediated Cervical Cancer and Prospects for Emerging Therapeutic Opportunities," Cancers, vol. 15, no. 4, pp. 1333, 2023.

[44] U. D. Braumann, J. P. Kuska, J. Einenkel, L. C. Horn, M. Löffler, and M. Höckel, "Three-dimensional reconstruction and quantification of cervical carcinoma invasion fronts from histological serial sections," IEEE Transactions on Medical Imaging, vol. 24, no. 10, pp. 1286–1307, 2005.

[45] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," CA: A Cancer Journal for Clinicians, vol. 71, no. 3, pp. 209-249, May 2021.

[46] L. Bruni, L. Barrionuevo-Rosas, G. Albero, B. Serrano, M. Mena, D. Gomez, and S. De Sanjose, "Human papillomavirus and related diseases report," in ICO/IARC Information Centre on HPV and Cancer, HPV Information Centre, pp. 26-42, 2019.

[47] A. C. Bovo, P. G. Pedrao, Y. M. Guimaraes, L. R. Godoy, J. C. P. Resende, A. Longatto-Filho, and R. Dos Reis, "Combined Oral Contraceptive Use and the Risk of Cervical Cancer: Literature Review," Revista Brasileira de Ginecologia e Obstetrícia, vol. 45, no. 12, pp. 818-824, 2023.

[48] D. Viveros-Carreño, A. Fernandes, and R. Pareja, "Updates on cervical cancer prevention," International Journal of Gynecological Cancer, vol. 33, no. 3, pp. 394–402, 2023.

[49] W. William, A. Ware, A. H. Basaza-Ejiri, and J. Obungoloch, "A review of image analysis and machine learning techniques for automated cervical cancer screening from pap-smear images," Computer Methods and Programs in Biomedicine, vol. 164, pp. 15–22, 2018.

[50] O. Tatar, B. Haward, P. Zhu, G. Griffin-Mathieu, S. Perez, E. McBride, A. K. Lofters, L. W. Smith, M.-H. Mayrand, E. M. Daley, J. M. L. Brotherton, G. D. Zimet, and Z. Rosberger, "Understanding the Challenges of HPV-Based Cervical Screening:

Development and Validation of HPV Testing and Self-Sampling Attitudes and Beliefs Scales," Current Oncology, vol. 30, no. 1, pp. 1206-1219, 2023.

[51]   Y. Song, L. Zhang, S. Chen, D. Ni, B. Li, Y. Zhou, B. Lei, and T. Wang, "A deep learning-based framework for accurate segmentation of cervical cytoplasm and nuclei," Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 2903-2906, 2014.

[52]   T. Xu, H. Zhang, X. Huang, S. Zhang, and D. N. Metaxas, "Multimodal deep learning for cervical dysplasia diagnosis," in Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016, Cham: Springer International Publishing, pp.115–123,2016.

[53]   T. Wang, H. Zhang, Y. Liu, and C. Zhao, "Updates in Cervical Cancer Screening Guidelines, The Bethesda System for Reporting Cervical Cytology, and Clinical Management Recommendations," Journal of Clinical and Translational Pathology, vol. 3, no. 2, pp. 75-83,2023.

[54]   J. Schlemper, O. Oktay, M. Schaap, M. Heinrich, B. Kainz, B. Glocker, and D. Rueckert, "Attention gated networks: Learning to leverage salient regions in medical images," Medical Image Analysis, vol. 53, pp. 197-207, 2019.

[55]   K. Adem, S. Kiliçarslan, and O. Cömert, "Classification and diagnosis of cervical cancer with softmax classification with stacked autoencoder," Expert Systems with Applications, vol. 115, pp. 557–564, 2019.

[56] P. Bamford and B. Lovell, "A water immersion algorithm for cytological image segmentation," in Proceedings of the 14th International Conference on Pattern Recognition (ICPR), pp. 839-841, 1998.

[57]   R. B. Perkins, N. Wentzensen, R. S. Guido, and M. Schiffman, "Cervical Cancer Screening: A Review," JAMA, vol. 330, no. 6, pp. 547-558, 2023.

[58]   A. Pirovano, L. G. Almeida, S. Ladjal, I. Bloch, and S. Berlemont, "Computer-aided diagnosis tool for cervical cancer screening with weakly supervised localization and

detection of abnormalities using adaptable and explainable classifier," Medical Image Analysis, vol. 73, pp. 102167, 2021.

[59] S. L. Tan, G. Selvachandran, W. Ding, R. Paramesran, and K. Kotecha, "Cervical Cancer Classification from Pap Smear Images Using Deep Convolutional Neural Network Models," Interdisciplinary Sciences: Computational Life Sciences, pp. 1-23, 2023.

[60] Y. Song, J. Zou, K. S. Choi, B. Lei, and J. Qin, "Cell classification with worse case boosting for intelligent cervical cancer screening," Medical Image Analysis, vol. 91, pp. 103014, 2024.

[61] D. Y. Yang and K. Bracken, "Update on the new 9-valent vaccine for human papillomavirus prevention," Canadian Family Physician, vol. 62, no. 5, pp. 399-402, 2016.

[62] M. Tomko, M. Pavliuchenko, I. Pavliuchenko, Y. Gordienko, and S. Stirenko, "Multi-label classification of cervix types with image size optimization for cervical cancer prescreening by deep learning," in Inventive Computation and Information Technologies, Singapore: Springer Nature Singapore, pp. 885-902, 2023.

[63] G. Kumawat, S. K. Vishwakarma, P. Chakrabarti, P. Chittora, T. Chakrabarti, and J. C. W. Lin, "Prognosis of Cervical Cancer Disease by Applying Machine Learning Techniques," Journal of Circuits, Systems and Computers, vol. 32, no. 01, pp. 2350019-2350040, 2023.

[64] E. Tanzi and M. Canuti, "HPV Infection and Cervical Cancer," in Global Health Essentials, Cham, Springer International Publishing, pp. 109-116, 2023.

[65] N. Sompawong, J. Mopan, P. Pooprasert, W. Himakhun, K. Suwannarurk, J. Ngamvirojcharoen, T. Vachiramon, and C. Tantibundhit, "Automated Pap Smear Cervical Cancer Screening Using Deep Learning," in Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 7044-7048, 2019.

[66] A. Tripathi, A. Arora, and A. Bhan, "Classification of cervical cancer using Deep Learning Algorithm," in Proceedings of the 5th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 1210–1218, 2021.

[67] B. Rigaud, B. M. Anderson, Z. H. Yu, M. Gobeli, G. Cazoulat, J. Söderberg, E. Samuelsson, D. Lidberg, C. Ward, N. Taku, C. Cardenas, D. J. Rhee, A. M. Venkatesan, C. B. Peterson, L. Court, S. Svensson, F. Löfman, A. H. Klopp, and K. K. Brock, "Automatic segmentation using deep learning to enable online dose optimization during adaptive radiation therapy of cervical cancer," International Journal of Radiation Oncology, Biology, Physics, vol. 109, no. 4, pp. 1096-1110, 2021.

[68] C. Y. Ma, J. Y. Zhou, X. T. Xu, J. Guo, M. F. Han, Y. Z. Gao, H. Du, J. N. Stahl, and J. S. Maltz, "Deep learning-based auto-segmentation of clinical target volumes for radiotherapy treatment of cervical cancer," Journal of Applied Clinical Medical Physics, vol. 23, no. 2, pp. 1–11, 2022.

[69] T. Xu, H. Zhang, X. Huang, S. Zhang, and D. N. Metaxas, "Multimodal deep learning for cervical dysplasia diagnosis," in International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, Cham, pp. 15-123, 2016.

[70] M. E. Plissiti, P. Dimitrakopoulos, G. Sfikas, C. Nikou, O. Krikoni, and A. Charchanti, "SIPAKMED: A new dataset for feature and image-based classification of normal and pathological cervical cells in Pap smear images," in 2018 25th IEEE International Conference on Image Processing (ICIP), pp. 3144-3148, IEEE, October 2018.

[71] L. Hu, D. Bell, S. Antani, Z. Xue, K. Yu, M. P. Horning, N. Gachuhi, B. Wilson, M. S. Jaiswal, B. Befano, L. R. Long, R. Herrero, M. H. Einstein, R. D. Burk, M. Demarco, J. C. Gage, A. C. Rodriguez, N. Wentzensen, and M. Schiffman, "An Observational Study of Deep Learning and Automated Evaluation of Cervical Images for Cancer Screening," Journal of the National Cancer Institute, vol. 111, no. 9, pp. 923–932, 2019.

[72] W. Mousser and S. Ouadfel, "Deep feature extraction for pap-smear image classification: A comparative study," in Proceedings of the 2019 5th International Conference on Computer and Technology Applications, pp. 6-10, 2019.

[73]   Y. Xiang, W. Sun, C. Pan, M. Yan, Z. Yin, and Y. Liang, "A novel automation-assisted cervical cancer reading method based on convolutional neural network," Biocybernetics and Biomedical Engineering, vol. 40, no. 2, pp. 611-623, 2020.

[74]   M. Xia, G. Zhang, C. Mu, B. Guan, and M. Wang, "Cervical Cancer Cell Detection Based on Deep Convolutional Neural Network," in Chinese Control Conference (CCC), vol. 2020, pp. 6527–6532, 2020.

[75]   Z. Alyafeai and L. Ghouti, "A fully-automated deep learning pipeline for cervical cancer classification," Expert Systems with Applications, vol. 141, pp. 1–40, 2020.

[76]   J.H. Lee and S. Cho, "Sensors & Transducers the Automated Diagnosis Architecture and Deep Learning Algorithm for Cervical Cancer Cell Images," *Sensors & Transducers*, vol. 249, no. 2, pp. 102–109, 2021.

[77]   A. Khamparia, D. Gupta, J. J. P. C. Rodrigues, and V. H. C. de Albuquerque, "DCAVN: Cervical cancer prediction and classification using deep convolutional and variational autoencoder network," Multimedia Tools and Applications., vol. 80, no. 20, pp. 30399–30415, 2021.

[78]   D. Jia, Z. He, C. Zhang, W. Yin, N. Wu, and Z. Li, "Detection of cervical cancer cells in complex situation based on improved YOLOv3 network," Multimedia Tools Application, vol. 81, no. 6, pp. 8939–8961, 2022.

[79]   O. Attallah, "CerCan· Net: Cervical Cancer Classification Model via Multi-layer Feature Ensembles of Lightweight CNNs and Transfer Learning," Expert Systems with Applications, vol. 120624, 2023.

[80]   M. M. Kalbhor and S. V. Shinde, "Cervical cancer diagnosis using convolution neural network: feature learning and transfer learning approaches," Soft Computing, vol. 1, pp. 0123456789, 2023.

[81]   S. Y. Chung, J. S. Chang, and Y. B. Kim, "Comprehensive clinical evaluation of deep learning-based auto-segmentation for radiotherapy in patients with cervical cancer," Frontiers in Oncology, vol. 13, no. 4, pp. 1–11, 2023.

[82] L. Xu, F. Cai, Y. Fu, and Q. Liu, "Cervical cell classification with deep-learning algorithms," Medical & Biological Engineering & Computing, vol. 61, no. 3, pp. 821-833, 2023.

[83] K. Bora, M. Chowdhury, L. B. Mahanta, M. K. Kundu, and A. K. Das, "Pap smear image classification using convolutional neural network," in ACM International Conference Proceeding Series, 2016.

[84] X. Q. Zhang and S. G. Zhao, "Cervical image classification based on image segmentation preprocessing and a CapsNet network model," International Journal of Imaging System and Technology, vol. 29, no. 1, pp. 19–28, 2019.

[85] M. Rohmatillah, S. H. Pramono, Rahmadwati, H. Suyono, and S. A. Sena, "Automatic Cervical Cell Classification Using Features Extracted by Convolutional Neural Network," 2018 Electrical Power, Electronics, Communications, Control, and Informatics Seminar (EECCIS 2018), no. 1, pp. 382–386, 2018.

[86] K. K. Harinarayanan and N. J. OU, "Deep neural network based assistive screening for cervical cancer," in Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, 2018, pp. 1473-1483.

[87] Y. Promworn, S. Pattanasak, C. Pintavirooj, and W. Piyawattanametha, "Comparisons of pap smear classification with deep learning models," in Proceedings of the 2019 IEEE 14th International Conference on Nano/Micro Engineered and Molecular Systems (NEMS), Bangkok, Thailand, pp. 282-285, 2019.

[88] P. Guo, S. Singh, Z. Xue, R. Long, and S. Antani, "Deep learning for assessing image focus for automated cervical cancer screening," in 2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI), pp. 1-4, IEEE, 2019.

[89] E. Hussain, L. B. Mahanta, C. R. Das, and R. K. Talukdar, "A comprehensive study on the multi-class cervical cancer diagnostic prediction on pap smear images using a fusion-based decision from ensemble deep convolutional neural network," Tissue Cell, vol. 65, no. February, pp. 101347, 2020.

[90]   A. Yilmaz, A. A. Demircali, S. Kocaman, and H. Uvet, "Comparison of Deep Learning and Traditional Machine Learning Techniques for Classification of Pap Smear Images," 2020.

[91]   J. Martinez-Mas, A. Bueno-Crespo, R. Martinez-Espana, M. Remezal-Solano, A. Ortiz-Gonzalez, S. Ortiz-Reina, and J.-P. Martinez-Cendan, "Classifying Papanicolaou cervical smears through a cell merger approach by deep learning technique," Expert Systems with Applications, vol. 160, pp. 113707, 2020.

[92]   M. A. Mohammed, F. Abdurahman, and Y. A. Ayalew, "Single-cell conventional pap smear image classification using pre-trained deep neural network architectures," BMC Biomedical Engineering, vol. 3, no. 1, pp. 1–8, 2021.

[93]   X. Tan, K. Li, J. Zhang, W. Wang, B. Wu, J. Wu, X. Li, and X. Huang, "Automatic model for cervical cancer screening based on convolutional neural network: a retrospective, multicohort, multicenter study," Cancer Cell International, vol. 21, no. 1, pp. 1–10, 2021.

[94]   S. Dhawan, K. Singh, and M. Arora, "Cervix image classification for prognosis of cervical cancer using deep neural network with transfer learning," EAI Endorsed Transactions on Pervasive Health and Technology, vol. 7, no. 27, pp. 1–9, 2021.

[95]   A. Dongyao Jia, B. Zhengyi Li, and C. Chuanwang Zhang, "Detection of cervical cancer cells based on strong feature CNN-SVM network," Neurocomputing, vol. 411, pp. 112–127, 2020.

[96]   Y. R. Park, Y. J. Kim, W. Ju, K. Nam, S. Kim, and K. G. Kim, "Comparison of machine and deep learning for the classification of cervical cancer based on cervicography images," Scientific. Reports., vol. 11, no. 1, pp. 1–11, 2021.

[97]   L. Cao, J. Yang, Z. Rong, L. Li, B. Xia, C. You, G. Lou, L. Jiang, C. Du, H. Meng, W. Wang, M. Wang, K. Li, and Y. Hou, "A novel attention-guided convolutional network for the detection of abnormal cervical cells in cervical cancer screening," Medical Image Analysis, vol. 73, p. 102197, 2021.

[98]   K. M. A. Adweb, N. Cavus, and B. Sekeroglu, "Cervical Cancer Diagnosis Using Very Deep Networks over Different Activation Functions," IEEE Access, vol. 9, pp. 46612–46625, 2021.

[99]   R. Elakkiya, K. S. S. Teja, L. Jegatha Deborah, C. Bisogni, and C. Medaglia, "Imaging based cervical cancer diagnostics using small object detection - generative adversarial networks," Multimedia Tools and Applications, vol. 81, no. 1, pp. 191–207, 2022.

[100] B. Harangi, J. Toth, G. Bogacsovics, D. Kupas, L. Kovacs, and A. Hajdu, "Cell detection on digitized pap smear images using ensemble of conventional image processing and deep learning techniques," in 2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA), pp. 38-42, September 2019.

[101] H. Lin, Y. Hu, S. Chen, J. Yao, and L. Zhang, "Fine-grained classification of cervical cells using morphological and appearance based convolutional neural networks." IEEE Access, 7, pp. 71541-71549, 2019.

[102] R. Gorantla, R. K. Singh, R. Pandey, and M. Jain, "Cervical cancer diagnosis using CervixNet - A deep learning approach," in 2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE), pp. 397-404, October 2019.

[103] K. H. S. Allehaibi, L. E. Nugroho, L. Lazuardi, A. S. Prabuwono, and T. Mantoro, "Segmentation and classification of cervical cells using deep learning," IEEE Access, vol. 7, pp. 116925-116941, 2019.

[104] M. M. Rahaman, C. Li, X. Wu, Y. Yao, Z. Hu, T. Jiang, X. Li, and S. Qi, "A survey for cervical cytopathology image analysis using deep learning," IEEE Access, vol. 8, pp. 61687–61710, 2020.

[105] C. Dharani, S. Kaviya, S. Maheshwari, K. Monisha, and P. Elayaraja, "Visualization of Cervical Cancer Classification using Deep Convolutional Neural Network," International Journal of New Innovations in Engineerings and Technology, Special Issue, no. 2005, pp. 196-206, 2020.

[106] H. Chen, J. Liu, Q. M. Wen, Z. Q. Zuo, J. S. Liu, J. Feng, B. C. Pang, and D. Xiao, "CytoBrain: Cervical Cancer Screening System Based on Deep Learning Technology," Journal of Computer Science and Technology, vol. 36, no. 2, pp. 347–360, 2021.

[107] Y. Kano, H. Ikushima, M. Sasaki, and A. Haga, "Automatic contour segmentation of cervical cancer using artificial intelligence," Journal of Radiation Research, vol. 62, no. 5, pp. 934–944,2021.

[108] A. Desiani, M. Erwin, B. Suprihatin, S. Yahdin, A. I. Putri, and F. R. Husein, "Bi-path Architecture of CNN Segmentation and Classification Method for Cervical Cancer Disorders Based on Pap-smear Images," IAENG International Journal of Computer Science, vol. 48, no. 3, pp. 1–9, 2021.

[109] M. Kuko and M. Pourhomayoun, "Single and Clustered Cervical Cell Classification with Ensemble and Deep Learning Methods," Information Systems Frontiers, vol. 22, no. 5, pp. 1039–1051, 2020.

[110] E. L. P. da Silva, "Combining machine learning and deep learning approaches to detect cervical cancer in cytology images," pp.1-76, 2021

[111] I. Pacal and S. Kilicarslan, "Deep learning-based approaches for robust classification of cervical cancer," Neural Computing and Applications, vol. 35, no. 25, pp. 1-16, 2023.

[112] C. A. Meza Ramirez, M. Greenop, Y. A. Almoshawah, P. L. Martin Hirsch, and I. U. Rehman, "Advancing cervical cancer diagnosis and screening with spectroscopy and machine learning," Expert Review of Molecular Diagnostics, vol. 23, no. 5, pp. 375-390, 2023.

[113] S. Fekri-Ershad and M. F. Alsaffar, "Developing a Tuned Three-Layer Perceptron Fed with Trained Deep Convolutional Neural Networks for Cervical Cancer Diagnosis," Diagnostics, vol. 13, no. 4, pp. 1-18, 2023.

[114] N. Youneszade, M. Marjani, and C. P. Pei, "Deep learning in cervical cancer diagnosis: architecture, opportunities, and open research challenges," IEEE Access, vol. 11, pp. 6133-6149, 2023.

[115] C. Chen, Y. Cao, W. Li, Z. Liu, P. Liu, X. Tian, C. Sun, W. Wang, H. Gao, S. Kang, S. Wang, J. Jiang, C. Chen, and J. Tian," The pathological risk score: A new deep learning-based signature for predicting survival in cervical cancer," Cancer Medicine, vol. 12, no. 2, pp. 1051-1063, 2023.

[116] Y. Morankar, V. Pange, P. Nagare, O. Randhve, M. Kalbhor, and S. Shinde, "A Deep Learning Pipeline for Cervical Cancer Detection," in Intelligent Systems for Smart Cities (ICISA 2023), pp. 335-354, 2024.

[117] X. Li, M. Du, S. Zuo, M. Zhou, Q. Peng, Z. Chen, J. Zhou, and Q. He, "Deep convolutional neural networks using an active learning strategy for cervical cancer screening and diagnosis," Frontiers in Bioinformatics, vol. 3, pp. 1101667, 2023.

[118] Z. Kang, Y. Li, J. Liu, C. Chen, W. Wu, C. Chen, X. Lv, and F. Liang, "H-CNN combined with tissue Raman spectroscopy for cervical cancer detection," Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, vol. 291, pp. 1-17, 2023.

[119] S. S. T. Sheela Shiney, and R. J. Rose, "Deep auto encoder based extreme learning system for automatic segmentation of cervical cells," IETE Journal of Research, vol. 69, no. 7, pp. 4066-4086, 2023.

[120] O. Attallah, "CerCan· Net: Cervical cancer classification model via multi-layer feature ensembles of lightweight CNNs and transfer learning," Expert Systems with Applications, vol. 229, pp. 1-40, 2023.

[121] M. J. Del Moral-Argumedo, C. A. Ochoa-Zezzati, R. Posada-Gómez, and A. A. Aguilar-Lasserre, "A Deep Learning approach for automated Cytoplasm and Nuclei cervical segmentation," Biomedical Signal Processing and Control, vol. 81, pp. 104483, 2023.

[122] S. Yu, X. Feng, B. Wang, H. Dun, S. Zhang, R. Zhang, and X. Huang, "Automatic classification of cervical cells using deep learning method," IEEE Access, vol. 9, pp. 32559–32568, 2021.

[123] T. Mostafid, "Overview of VGG16, ResNet50, Xception and MobileNet neural networks,"Medium,11-Dec2023.Available:https://medium.com/@t.mostafid/overview-of-vgg16-xception-mobilenet-and-resnet50-neural-networks-c678e0c0ee85

[124] P. Wang, J. Wang, Y. Li, L. Li, and H. Zhang, "Adaptive Pruning of Transfer Learned Deep Convolutional Neural Network for Classification of Cervical Pap Smear Images," IEEE Access, vol. 8, pp. 50674–50683, 2020.

[125] A. E. Carpenter, T. R. Jones, M. R. Lamprecht, C. Clarke, I. H. Kang, O. Friman, D. A. Guertin, J. H. Chang, R. A. Lindquist, J. Moffat, P. Golland, and D. M. Sabatini, "CellProfiler: image analysis software for identifying and quantifying cell phenotypes," Genome Biology, vol. 7, pp. 1-11, 2006.