Enhancing VIA Screening for Cervical Cancer: A Comprehensive System Integrating Image Processing, Risk Factors and Follow-up Facilitation

by

Zulkar Nain
20101226
Rezwan Ahmed
20101235
Aziz Ahmed
20301395
Tasmia Rahman Mitu
20101578
Kawser Sultana
20301198

A thesis submitted to the Department of Computer Science and Engineering in partial fulfillment of the requirements for the degree of B.Sc. in Computer Science

Department of Computer Science and Engineering
Brac University
May, 2024

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Zulkar Nain

Zulkar Nain
20101226

Rezwan Ahmed

Rezwan Ahmed
20101235

Aziz Ahmed

Aziz Ahmed
20301395

Rahman

Tasmia Rahman Mitu
20101578

Kawser Sultana

Kawser Sultana
20301198

Approval

The thesis/project titled "Enhancing VIA Screening for Cervical Cancer: A Comprehensive System Integrating Image Processing, Risk Factors and Follow-up Facilitation" submitted by

1. Zulkar Nain (20101226)
2. Rezwan Ahmed (20101235)
3. Aziz Ahmed (20301395)
4. Tasmia Rahman Mitu (20101578)
5. Kawser Sultana (20301198)

Of Spring, 2024 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on May 21, 2024.

Examining Committee:

Supervisor:
(Member)

Md. Ashraful Alam, PhD
Associate Professor
Computer Science and Engineering
Brac University

Program Coordinator:
(Member)

Md. Golam Rabiul Alam, PhD
Professor
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Sadia Hamid Kazi, PhD
Chairperson and Associate Professor
Department of Computer Science and Engineering
Brac University

Abstract

Visual Inspection of cervix with Acetic Acid (VIA) is an inexpensive and effective screening test which is being conducted in many under-developped and developing regions. In medical image processing applications, the performance of Computer Vision has been promising. This research aims to establish a systematic process for VIA (Visual Inspection with Acetic Acid) screening of the cervix by incorporating Computer Vision and Machine Learning techniques. The paper analyzes the performance of VGG16, ResNet-50, YOLOv9, YOLO-NAS(Medium) on cervix images with Acetic Acid (VIA), with VGG-16 achieving 96% accuracy, ResNet-50 achieving 95%, YOLOv9 achieving 93%, and YOLO-NAS achieving 91%. Furthermore, we use feature importance scores from the Random Forest model to extract key features associated with cervical cancer from demographic, behavioral, and clinical factors. Training an ensemble model on these features yields an accuracy of 94%. The goal is to analyze the cervical images during VIA inspection and predict its outcome using Computer Vision and integrate patient risk factors so that our proposed system identifies high-risk individuals, even among VIA-negative cases and improves overall screening accuracy, surpassing the existing VIA screening method.

Keywords: Cervical Cancer, Computer Vision, CNN, VIA, YOLO, VGG, ResNet.

Acknowledgement

All praise to Allah for whom our thesis has been completed without any major interruption. We owe a tremendous debt of appreciation to Md. Ashraful Alam, PhD, our thesis supervisor, for his invaluable advice, assistance, and moral support during the duration of this study. His professionalism, insight, and commitment to our progress have been indispensable.

We would like to express our deepest gratitude to Dr. Mohammad Sayeed Hossain, Associate Professor, Department of Oncology Dhaka Medical College Hospital. His expert insights in cancer were instrumental in the completion of this study.

We would also like to express our deepest gratitude to Dr. Rasheda Begum (MBBS, DMUD, CCD, MPH, FCPS), Consultant in Gynaecology at the BAVS Maternity Hospital, Dhaka, Bangladesh. Her expert insights, suggestions and constructive feedback were instrumental in the completion of this study.

We are grateful to our beloved BRAC University and the CSE Department for facilitating this study and giving us access to the tools we needed to complete it.

And finally to our parents. Without their thorough support it may not be possible. With their kind support and prayer we are now on the verge of our graduation.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

The incidence of cervical cancer is fourth among all female cancers worldwide. There are approximately 570,000 new instances of cervical cancer every year, with an anticipated 310,000 fatalities. Moreover, it is among the most frequent cancers in 23 nations and the leading cause of cancer deaths in 36 (Allahqoli et al.,). While cervical cancer is the second most common type of cancer of women in Bangladesh. In Bangladesh, it affected 8,068 (10.6 per 100,000 women) women and caused 5.214 (7.1 per 100,00 women) deaths in 2018 (Bray et al.,) (Sung et al.,). The cells of a woman's cervix are particularly vulnerable to the development of this malignancy. It first affects the cervical area's deep tissues before spreading to the lungs, liver, and vagina. Cervical cancer is mostly caused by Human papillomavirus (HPV). It is one of the most common viral infections of the reproductive system. Human papillomavirus (HPV), a very prevalent family of viruses that is transferred through genital contact, is the main contributing factor to persistent, high-risk kinds of cervical cancer (International Agency for Research on Cancer,). One of the 15 genotypes of the cancer-causing HPV causes chronic infection of the cervical epithelium in almost all instances of cervical cancer (Hou et al.,). Although most HPV infections are self-limiting and show no symptoms. However persistent infections can lead to cervical cancer in females. Precancerous lesions can progress to cervical cancer over a long period. These lesions can be diagnosed and treated. If patients are diagnosed in precancerous form, cervical cancer can be successfully cured but it has to be detected early. But HPV infections do not cause any clinical symptoms. So, they can only be identified by cervical screening.

Cervical cancer can be detected early and affordably with Visual Inspection with Acetic Acid (VIA) screening. Its efficacy in lowering the incidence and death of cervical cancer has been demonstrated in several trials, especially in environments with limited resources. For instance, VIA screening was linked to a statistically significant decrease in the incidence and mortality of cervical cancer, according to a systematic review and meta-analysis(Lohiya et al.,). The Bangladesh government has been running a screening program recommended by the World Health Organization (WHO). As mentioned before, 90% of the new cases and deaths that occurred globally in 2020 were in low and middle-income countries (World Health Organization,). Similarly, Bangladesh has many high risk factors such as child marriage, limited social awareness about reproductive organs and sexually transmitted diseases (STD), and low socio-economic condition that can increase cervical cancer among the female population. So, the government started 417 screening centers in many levels of health care facilities in all 64 districts of Bangladesh. These screening centers adopted the visual inspection with acetic acid (VIA) method. According to the report of National Center for Cervical and Breast Cancer Screening and Training of Bangabandhu Sheikh Mujib Medical University(BSMMU), 24 hundred thousand women were screened between 2017 and 2022 with a screening coverage and positivity rate of 11.30% and 5.71% (Uddin, Sumon, Pervin, & Sharmin,).

VIA : VIA(Visual Inspection with acetic acid) plays a vital role in the early diagnosis and screening process of cervical cancer. According to the WHO (World Health Organization), generally, women aged 30-49 years old are the main target of VIA screening every 3 to 5 years. But, those who have HIV should start screening at 25 and do the screening frequently. During any phase in the menstrual cycle and also 20 weeks in pregnancy, this test can be done but in advanced pregnancy it becomes difficult to take the test (International Agency for Research on Cancer,). The process of this test is to apply 3% to 5% acetic acid to the entire cervix and after that examine by naked-eye of the uterine cervix with adequate illumination. Based on the well-defined acetowhite area that can be detected on the transformation zone of the cervix after one minute of applying acetic acid, the result of the test becomes interpreted (International Agency for Research on Cancer,). Therefore, before screening, women should be counseled about cervical cancer, the importance of this screening, test procedures, possible test results and the safe treatment options that are available etc. as lack of proper information can lead them to a negative psychological reaction if they get an abnormal screening test result. VIA test results are divided into three categories and they are Negative,Positive and suspicious of invasive cancer. According to the outcomes, necessary steps have to be taken immediately. After applying acetic acid to the entire cervix for 1 minute, if the squamous epithelium or the columnar epithelium doesn't show any change or no acetowhite area can be seen then the result is reported as VIA-negative (International Agency for Research on Cancer,).

But sometimes, after applying acetic acid to the cervix, the columnar epithelium turns white immediately but fades away after a few seconds. This matter happens due to the shrinkage of the subepithelial blood vessels which happens as the effect of the acetic acid. Even if there have been seen white areas for some moments, this situation should not be considered as VIA-positive as the blanching or shrinkage is normal (International Agency for Research on Cancer,).

Moreover, in the columnar epithelium, sometimes the indistinct transparent acetowhite patches can be visible after applying acetic acid which is suggestive as squamous metaplasia. So, this isn't considered also as VIA-positive (International Agency for Research on Cancer,).

After applying acetic acid to the entire cervix and observing 1 minute, if there are seen dense acetowhite areas in the TZ (Transformation Zone) as well as with well-defined margins, then this result is reported as VIA-positive. Here this transformation zone is the area where the inner part (endocervix) and outer part (ectocervix) around the opening of the cervix (National Cancer Institute,). When the VIA-positive, this is seen that the high-grade cervical precancerous cells appear as distinctly dense. Moreover, in the transformation zone, there are opaque acetowhite areas and the margins of the lesions are also very well defined and either flat or

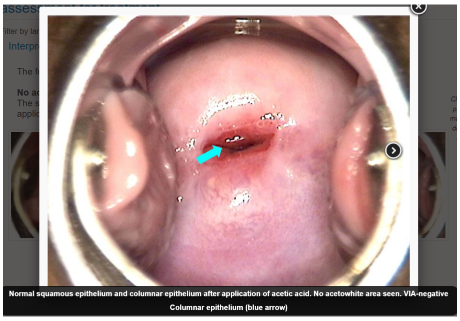

Figure 1.1: VIA-negative(Normal squamous and columnar epithelium)

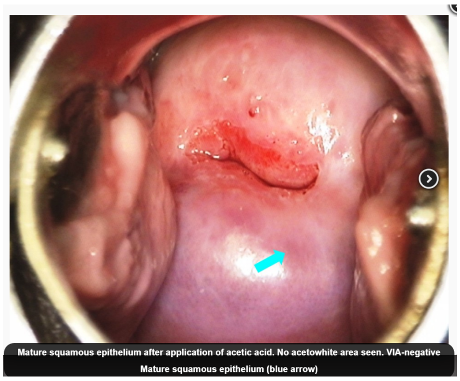

Figure 1.2: VIA-negative (Mature squamous epithelium)

raised from the surface. The diagnosis of VIA should be made based on the lesion which has the worst features as the size of the lesions may vary. Some can be a very dense area covering acetowhite whereas some can be very small. Even sometimes, if the lesion extends beyond 2mm of the endocervical canal, the inner margin of the area may not be visible anymore (International Agency for Research on Cancer,).
The diagnosis of VIA should be made based on the lesion which has the worst features as the size of the lesions may vary. Some can be a very dense area covering acetowhite whereas some can be very small (International Agency for Research on Cancer,).
Even sometimes, if the lesion extends beyond 2mm of the endocervical canal, the inner margin of the area may not be visible anymore. For the postmenopausal women, the explanation of the test is more difficult as the most severe part of the lesion can be inside the endocervical canal which is why it may not be detectable through VIA. Moreover, in high grade precancerous, there could be seen the underneath red stroma as the epithelium has the tendency to peel off. Therefore, when this underneath stroma becomes exposed, the red patches with bleeding points also become visible (International Agency for Research on Cancer,).
Moreover, it can happen that there is a large area in which acetowhite is occupied after applying acetic acid or in the acetowhite area, the surface becomes irregular or the cervix is growing frankly etc. This situation indicates suspicious invasive cancer. In this situation, a necrotic or ulcerated area can also be seen. Even sometimes the surface may not show any acetowhite area after applying acetic acid, rather the entire surface exposes red patches with bleeding points as the epithelium of the entire surface may be peeled off and the underlying growth became visible. If a woman gets the VIA test result as negative, then she may have been advised to pass through screening again after a limited time.

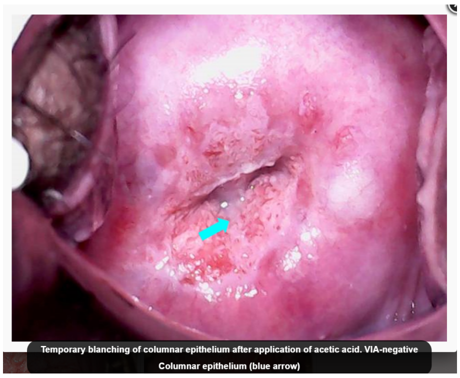

Figure 1.3: VIA-negative (Temporary blanching of columnar epithelium)

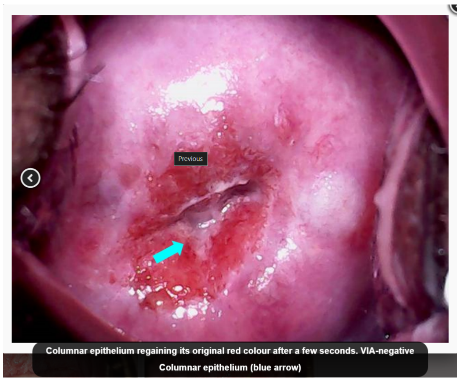

Figure 1.4: VIA-negative (Columnar epithelium regaining its original red color)

According to WHO, women should do VIA screening every 3 years. Though the priority should be given according to age. 30 to 49 years women and 25-49 years HIV positive should be given more priority. For the postmenopausal women, excisional treatment should be made available. If any women get a VIA positive result, approximate follow-up visits can help them to overcome their fear and embarrassment of gynecological check-up as well as they can know more about this and get treated. Moreover, the VIA test doesn't indicate the presence of cervical cancer, so another confirmatory test such as colposcopy is needed for this as well. However, there are limited services of colposcopy in low-resource countries. Therefore, according to WHO, the woman who is VIA-positive should get the treatment during the same visit which is called screen-and-treat approach. However, the interpreted result of the VIA test not only just depends on the screening test but also other risk factors such as that woman's age, menstrual and obstetric history, history of previous screening of cervical cancer and also if any other disease she has been suffering [9]. As, VIA test basically depends on the provider's skill mostly to properly visualize the cervix, evaluate the results and interpret them correctly. Therefore, a few errors may occur during the VIA screening test such as insufficient exposure of the cervix. If the cervix exposure is not insufficient because of the inadequate lighting or the patient's improper positioning or laxity of vaginal walls etc. then the image taken during the test may not be proper which can lead to errors in the test. In the transformation zone if the visualization is not clear or sufficient then it can also lead to error. Insufficient visualization of the margins on the acetowhite area can also lead to error in VIA screening test. Moreover, during applying acetic acid to the cervix, the faults that the providers do also lead to error. Error in recog-

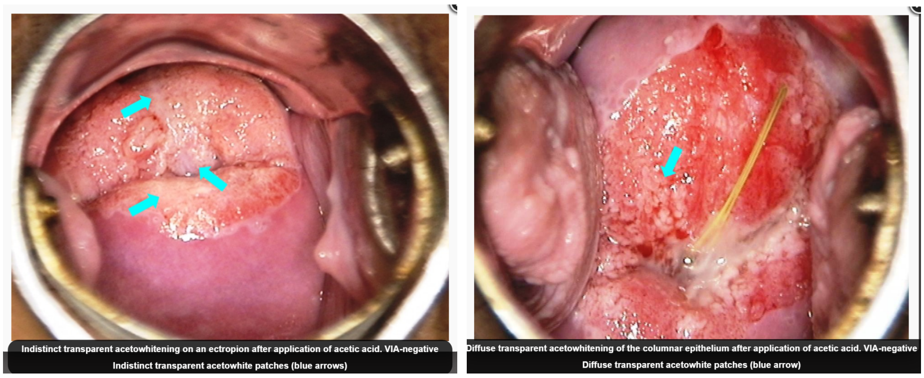

Figure 1.5: VIA-negative (Indistinct and diffuse transparent acetowhitening)

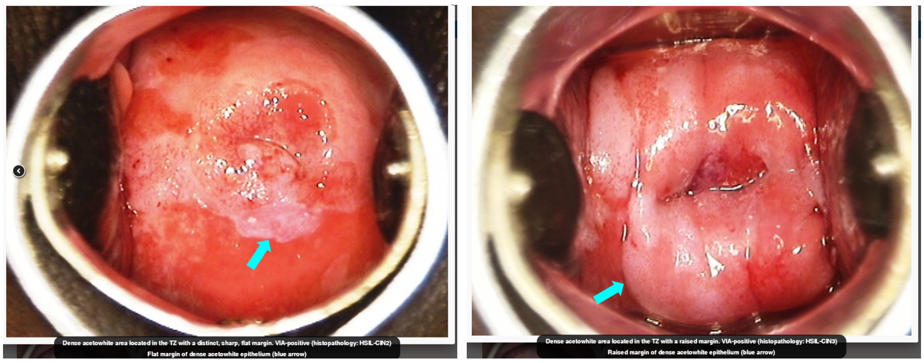

Figure 1.6: VIA-positive

nizing cervical cancer and in assessment of the postmenopausal cervix also lead to errors in the screening test [9]. Therefore, to minimize these errors, first of all, the provider should have to be skilled to do the test without any errors and examine the images. Even so, there can be some confusion and the provider may have made some error recognizing the images. To minimize this error using real-time based image processing methods or deep learning technology can give us more efficient and accurate results. Using powerful CNN models for image processing like YOLOv9, ResNet-50, VGG16 etc. to enhance the screening of VIA of cervical cancer is the most necessary aspect of our proposed system. Now-a-days, for analyzing medical images, convolutional neural networks is a great methodology (Litjens et al.,). Furthermore, it tries to combine patient risk factors into the screening process to bolster the accuracy and relevance of the results. A major component of the study is committed to analyzing the system's usefulness in real-world scenarios, with a focus on solving the challenges of accessibility and data gathering in impoverished regions like Bangladesh and least developed countries.

1.2 Motivation

Cervical cancer, a fatal disease by which women are being affected a lot worldwide. Though this is the fourth common female cancer worldwide, in 42 low-resource countries, it is the most common cancer suffered by women (Arbyn et al.,). In 2022, there were approximately 660,000 cases worldwide of cervical cancer and among them 350,000 deaths had occurred and among this 94% of these deaths occurred

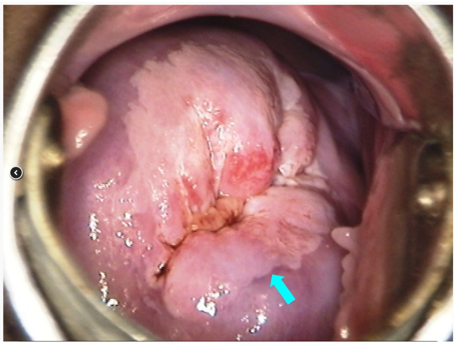

Figure 1.7: VIA-positive

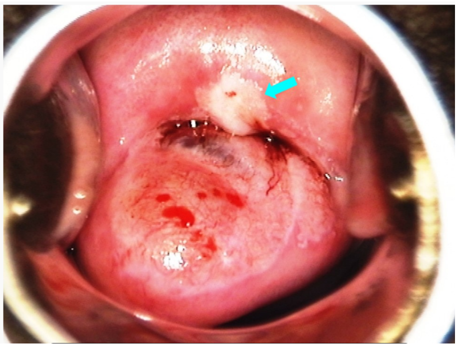

Figure 1.8: VIA-positive

in low and middle-income countries due to low-resources of vaccination system, screening, treatment (World Health Organization,). Moreover, the risk factors and the socio-economic determinants are also the cause of these deaths (World Health Organization,). There are a few methods to prevent and to cure cervical cancer in its early stage. VIA(Visual Inspection with acetic acid) is the most effective method of screening in low-resource and developing countries (*Europe PMC*,). But to get an accurate and consistent result this screening process has to be in the presence of good training and maintain quality assurance (*Europe PMC*,). But, in under-developed countries where resources and modern diagnostic machines are limited, just VIA screening can not increase screening of cervical cancer and prevent this. Therefore,

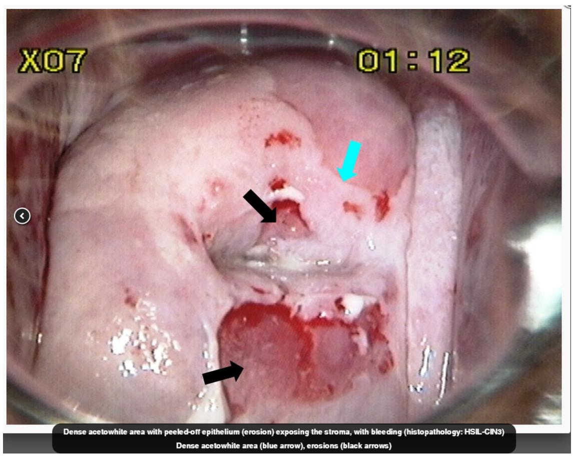

Figure 1.9: Dense acetowhite area with peeled off epithelium.

using AI (Artificial Intelligence) and deep learning technology can enhance this VIA screening. Convolutional Neural Network (CNN) , which is a part of deep learning, has become a preferred approach for examining medical images (Litjens et al.,). In the VIA screening process, if we use CNN networks like ResNet-50, VGG16, YOLOv9, YOLO-NAS we can increase the efficiency and accuracy rate of the early diagnosis of cervical cancer. In image processing and segmentation, these CNN networks can deliver more accurate and efficient results and reduce errors than only human based approaches. Therefore, a country like Bangladesh has many socio-economic factors which impact the progress of cervical cancer (Alam et al.,). Hence, this research extracts the high risk factors based on a person's history and other socio-economic and behavioral factors. The motivation for this research is to enhance VIA screening by implementing a systematic process which integrates deep learning algorithms like ResNet-50, VGG16, YOLOv9, YOLO-NAS to get a more efficient outcome of VIA screening test. Moreover, if a person is VIA-negative then the proposed system can give her an outcome if she is at high risk of having cervical cancer or not based on the extracted risk factors. As a result, Also because of this, the patients who are at high risk of getting cervical cancer can be identified and get her treatment as soon as possible. Otherwise, if a person is VIA-positive, then this system will suggest to her to get necessary treatments. Consequently, this system can enhance and motivate people for VIA screening in low-resource countries by making it easily accessible for all the people.

1.3 Problem Statement

Visual Inspection with Acetic acid is a simple, cost efficient and popular method for cervical cancer detection in low resource and underdeveloped countries. But this procedure, being solely dependent on human observation interpretation, has limitations in accuracy and consistency. And these limitations can induce higher rates of false positive and false negative detections, unfavorably affecting early detection of cervical cancer and its treatment outcomes. Moreover, current detection methods lack the incorporation of patient data, including patient lifestyle, sociodemographic information, and risk factors, which are important for more errorless detection of an individual's risk of developing cervical cancer. So a systematic model is required to combine the robust image processing techniques with patient risk factor analysis.

1.4 Research Objective

The main objective of this research is to enhance the VIA screening method using a systematic process for cervical cancer. To achieve that, we will develop a comprehensive system including higher accuracy screening results, patients' risk factors, and follow-up data. To get accurate screening results, we will use Convolutional Neural Networks such as Resnet-50, VGG16, YOLOv9, YOLO-NAS etc. The application of CNN networks in medical images gives more accurate and less error results than only human-based results. Additionally, we will integrate patients' medical history data, including socioeconomic and demographic data, and risk factors into this screening system. We will combine the patient's risk factors with this screening method to identify VIA-negative patients who are at risk of developing cervical cancer. We aim to enhance the precision and relevance of the screening results, ensuring a more extensive and inclusive approach to cervical cancer detection and early diagnosis.

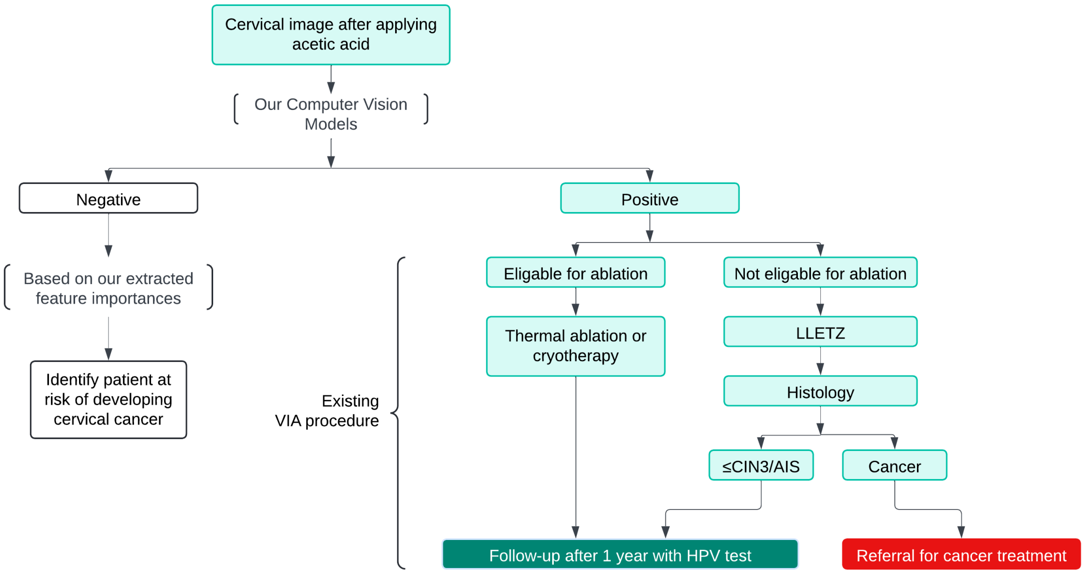

Figure 1.10: Our proposed VIA model

Our system would combine data from screening results, patient medical history, and risk factors. This combined data would be significant for evaluating future screening accuracy, allowing us to refine our system and improve predictions of the VIA results. A vital objective of our research is to design a system that would be well-suited for low-resource and underdeveloped countries like ours.

Chapter 2

Literature Review

For a low-resource country like ours, the VIA screening method is already adopted in many screening centers of Bangladesh. As cervical cancer can be detected early and it is also affordable to do a Visual inspection with Acetic Acid (VIA) screening. Therefore, by enhancing VIA screening by establishing a systematic process using image processing techniques and patient's risk factors will speed up the screening process more [2,3]. Several studies have been published based on VIA methods for preventing cervical cancer using deep learning methods or not.
Rengaswamy Sankaranarayanan, Ashrafun Nessa, Pulikattil Okkuru Esmy MD, and Jean-Marie Dangou discussed various VIA methods for cervical cancer prevention in their study "Visual inspection methods for cervical cancer prevention." According to them, though this screening method has some limitations, it has a wide accuracy variance and for low-resource countries, VIA screening methods provide a very practical approach to increase other testing services for cervical cancer in future. VIA screening methods can identify cervical intraepithelial neoplasia grade 2 or worse lesions that may cause cervical cancer in the future. Screening of visual inspection with acetic acid can reduce about 25-35% in cervical cancer incidence as well as the frequency of cervical intraepithelial neoplasia grade 2 or worse lesions in its randomized-controlled trials. They have explained various visual inspection methods like visual inspection with acetic acid, visual inspection with lugol's iodine etc. in their paper. Therefore, after describing various methods of visual inspection they have concluded that visual inspection with acetic acid is the most suitable test for low-resource countries even if it has certain limitations (Sankaranarayanan, Nessa, Esmy, & Dangou,).
In the paper "Artificial Intelligence-Based Cervical Cancer Screening on Images Taken during Visual Inspection with Acetic Acid: A Systematic Review" Roser Vinals, Magali Jonnalagedda, Patrick Petignat, Jean-Philippe Thiran and Pierre Vassilakos presented various AI based cervical cancer screening algorithms which are taken during visual inspection with acetic acid in various papers of PubMed, Google Scholar, Scopus etc. They reviewed those papers' algorithms and concluded that the images taken during VIA are very useful as AI-based algorithms can detect the precancerous and cancerous lesions very well. Moreover, according to them, for low-resource countries this process can give more useful services as they can give more accurate test results than human explanation of the screening process (Viñals, Jonnalagedda, Petignat, Thiran, & Vassilakos,). In the publication "Deep Learning for Assessing Image Focus for Automated Cervical Cancer Screening," Peng Guo,

Sanjana Singh, Zhiyun Xue, Rodney Long, and Sameer Antani described a deep learning architecture that detects in-focus smartphone cervical photos. To improve the sharpness of the photos captured in low-resource conditions utilizing the VIA technique, researchers have created three types of deep learning networks such as RetinaNet, fine-tuned deep learning models, and transfer learning models using RetinaNet, VGG16, VGG/Inception, L1 feature selection, SVM classifier, ResNet50 etc. and compared among them to get the most sharpened pictures. They have labeled them as "sharp" and "not sharp" using acetic acid on the Visual Inspection photos (Guo, Singh, Xue, Long, & Antani,).
To enhance VIA(Visual Inspection with Acetic Acid) screening for cervical cancer, different deep learning methods and image processing algorithms have been used so far. Convolutional neural networks (CNNs) and other deep learning algorithms (Resnet-101, VGG-16, VGG-19, InceptionV3) have been investigated in a number of research for the identification and classification of cervical cancer. A powerful CNN algorithm for image categorization is called VGG. In 2014, the Visual Geometry Group at Oxford invented it. The VGG16 network has 16 layers (13 convolutional, 3 fully connected). Therefore, A deep CNN architecture called ResNet was put out in 2015. It incorporates skip connections that allow the network to learn residual functions, which helps to overcome the problem of disappearing gradients in deep networks. It comes in many variations, with 50, 101, and 152 layers in ResNet-50, ResNet-101, and ResNet-152, respectively (Leonardo, Carvalho, Rezende, Zucchi, & Faria,). Tapomoy Adhikari his study "Designing a Convolutional Neural Network for Image Recognition: A Comparative Study of Different Architectures and Training Techniques" presented various convolutional neural networks for image recognition in various fields. He explained about various CNN networks like LeNet, AlexNet, VGG-16, VGG-19, ResNet-50, ResNet-101, ResNet-152 etc. and highlighted their strengths and weaknesses in the literature review of the paper (Adhikari,).
In the paper "Enhancing cervical cancer detection and robust classification through a fusion of deep learning models" Sandeep Kumar Mathivanan, Divya Francis, Saravanan Srinivasan, Vaibhav Khatavkar, Karthikeyan P and Mohd Asif Shah described about how they have used quite a few pre-trained deep learning models like Alexnet, Resnet-101, Resnet-152, InceptionV3 etc. in cervical cancer detection in detail. They have demonstrated each of their architectures and their working procedures also. In conclusion they presented a comparison among the accuracy rate among the deep learning models for the diagnosis of cervical cancer. Overall, the researchers have done these to enhance cervical cancer detection and screening (Leonardo et al.,). In the paper "Cervical Cancer Classification From Pap Smear Images Using Deep Convolutional Neural Network Models" the researchers also presented how they classified cervical cancer from images using pre-trained deep convolutional Neural Networks models. They have used various CNN models like VGG-16, VGG-19, DenseNet-121, DenseNet-169, DenseNet-201, ResNet-50, ResNet-101, ResNet-152, Inception, Xception, MobileNet, MobileNet-v2 etc. to classify the images. According to the researchers, these models have been chosen according to their strong performance in classification tasks of medical images. They also compared in various categories to evaluate which CNN model can perform the best for cervical cancer detection (Tan, Selvachandran, Ding, Paramesran, & Kotecha,). Moreover, using YOLO architectures in medical images improves diagnostic precision and efficiency

(Wijaya, Ryanto, Sidhik, Liawatimena, & Riana,). In this research, YOLOv9 has been used to detect objects in the images taken during VIA screening. To get efficient and higher accuracy for the limited dataset and being the latest architecture of the YOLO family , YOLOv9 is the better option than other YOLO models. In various papers, researchers used YOLO models for detecting objects and classifying images specially in analyzing medical images. Therefore, various medical object detection tasks have been done by YOLO algorithms and have been published in many articles as they have higher efficiency and have surpassed existing methods in performing these tasks (Qureshi et al.,). A group of researchers Md Zahid Hasan Ontor, Md Mamun Ali1, Kawsar Ahmed, Francis M. Bui, Fahad Ahmed Al-Zahrani, S. M. Hasan Mahmud and Sami Azam researched using YOLOv5 models to detect early stage cervical cancerous cells to get more efficient results than before. In their paper "Early-Stage Cervical Cancerous Cell Detection from Cervix Images Using YOLOv5 ", they have used cervical cancer screening image dataset and trained the model using YOLOv5 algorithms. They analyzed the results according to the performance and loss of each of YOLOv5's models (Ontor et al.,). In the paper "Cervical Cancer Cells Detection Using YOLOv8 Algorithm " Hendra Wijaya, Tommy Ryanto, Stephanie Catharina Sidhik, Suryadiputra Liawatimena and Dwiza Riana used YOLOv8 architecture to detect cervical cancer early. Their main motive behind this was to improve and obtain higher accuracy in real-time cell object detection than previous versions and enhancing cervical cancer screening. By using the images taken during the screening process they implemented the YOLOv8 model. They have vastly talked about the improvements of YOLO algorithms and the efficiency of YOLOv8. Finally, they made a comparison between YOLOv5 and YOLOv8 according to the result they acquired after using the model YOLOv8 and the previous studies of YOLOv5. Moreover, their dataset was limited. Though YOLOv5 performed well in previous studies but faced with limited data sets YOLOv8 obtained higher and optimal results than YOLOv5 (Wijaya et al.,).
To enhance VIA screening for cervical cancer requires combining a detailed process of integrating image processing using CNN or machine learning or Artificial Intelligence based algorithms. Moreover, the risk factors of cervical cancer also play a vital role in this. Though AI based algorithms enhance the effectiveness of cervical cancer screening, further studies and large scale testing is also required.

Chapter 3

Data Sets

3.1 Data set Sources

We collected cervix images from the cervical cancer image bank of iarc(International Agency for Research on Cancer) of WHO (International Agency for Research on Cancer,). These images are taken during the VIA procedure. The cervix images are taken after applying acetic acid. These images were classified as VIA negative and positive. We preprocessed these images with orientation and resize. We utilized picture augmentation strategies such as flipping, shearing, exposure and noise to expand the size of our dataset as our initial dataset was small.
We collected the patient history dataset from the UC Irvine machine learning repository (Fernandes, Cardoso, & Fernandes,). The dataset was collected at 'Hospital Universitario de Caracas' in Caracas, Venezuela. The dataset comprises demographic information, habits, and historic medical records.

3.2 Image Data Augmentation

Data augmentation is a highly effective strategy on images. It enhances the robustness and accuracy of the machine learning models in their ability to handle changes in real-world scenarios. In our research, data augmentation is crucial for enhancing the efficiency of models because the availability of VIA images is restricted. Here, we employed flipping, shearing, exposure and noise.

Augmentation strategies	Measurements
Flip	Horizontal, Vertical
Shear	$\pm10°$ Horizontal, $\pm10°$ Vertical
Exposure	Between -10% and +10%
Noise	Up to 0.50% of pixels
Rotation	-15°, +15°

Table 3.1: Augmentation Strategies

VIA verdict	Number of Images before Augmentation	Number of Images after Augmentation
VIA Positive	192	1920
VIA Negative	187	1870

Table 3.2: Initial Image Data Set

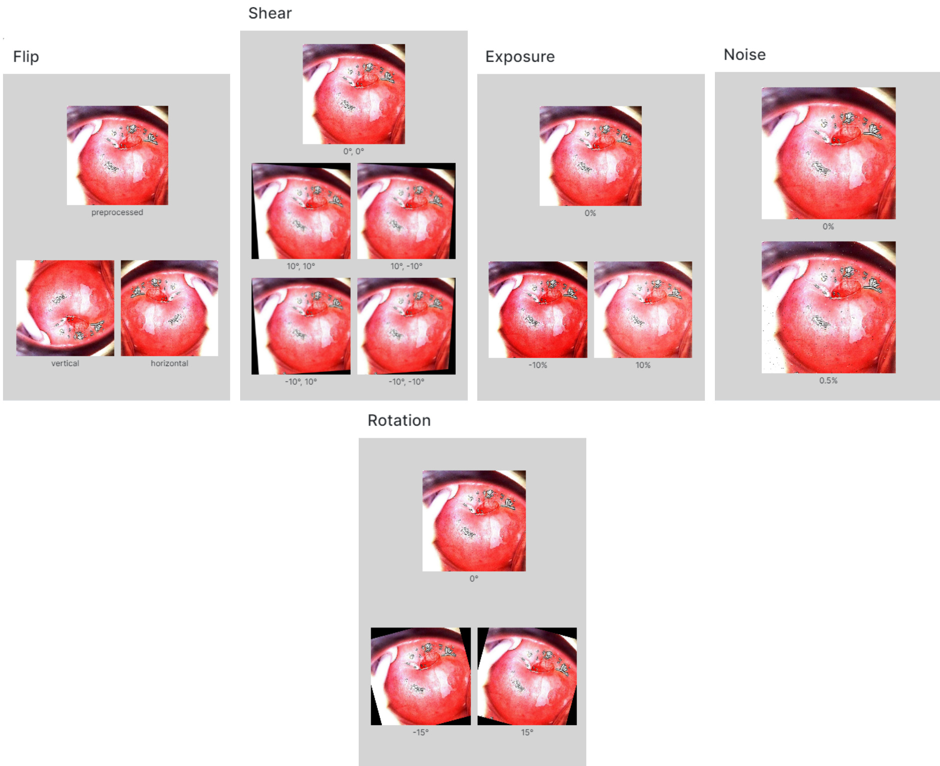

Figure 3.1: Images after Augmentation

3.3 Patient history dataset preprocessing

Missing values in the dataset were replaced by 'NaN' to facilitate data cleaning and analysis steps. Columns with numeric values were converted with appropriate numeric types for mathematical and statistical operations. Continuous variables were categorized. The ages of the patients were categorized into four groups, Below 20 years, 20–29 years, 30–44 years and 45–60 years (Figure 3.6). The 'First Sexual Intercourse' variable was categorized into five groups, Below 13 years, 13–15 years, 16–17 years, 18–19 years and Above 20 years (Figure 3.7). Initial exploratory data analysis, including correlation matrices, helped to identify which features have strong correlations with the target variable (Biopsy results). 9 most high-risk prediction Features of a positive biopsy were identified and chosen based on their relevance to cervical cancer risk, supported by literature and correlation analysis (5.7). The dataset had significant imbalance between the number of positive and negative biopsy results. This imbalance could bias the model towards predicting the majority class. To address this, oversampling of the minority class was performed using the resampling technique. Risk of overfitting was mitigated by the use of cross-validation (Table 5.2). Categorical features were converted into numeric format using one-hot encod-

Cases	Variables
858	36

Table 3.3: Patients History Dataset

ing. One–hot encoding transformed each category into a binary vector which made it suitable for ML algorithms.

The final dataset after preprocessing, contained a balanced distribution of negative and positive biopsy results and it was ready for ML model training and evaluation. Preprocessing steps ensured that the dataset was clean, numeric and free from class imbalance which facilitates effective model learning and prediction.

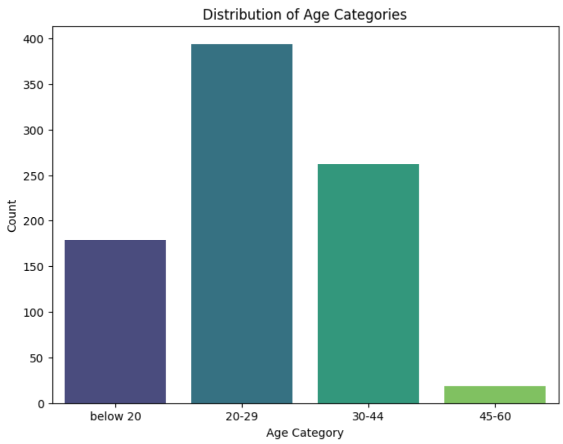

Figure 3.2: Age Distribution category

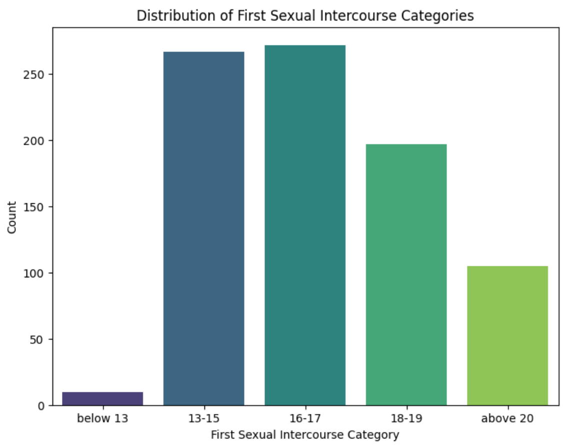

Figure 3.3: First Sexual Intercourse Distribution Category

Chapter 4

Methodology

We started by collecting raw images of cervix having VIA positive or Negative. We then processed the images for creating our database. We designed the structure of our database. Then we annotated the images of our database and did data augmentation on them. Once we made our desired database, we started training them on CNN algorithms. We used YOLOV9, VGG-16, RESNET50, YOLO-NAS. Once training was done on all four algorithms, we started evaluating and fine tuning them. Then came the second part of our thesis. We collected data for risk factor analysis. We got a database which we needed to modify. After modification, we did feature extraction using RF and then trained the extracted features on machine learning algorithms. We used SVM, RF, Linear Regression and Gradient Boosting. Once the training was done, we started evaluating them. We then used ensemble learning and combined the models for better performance. At last, we integrate the two models.

4.1 Workflow

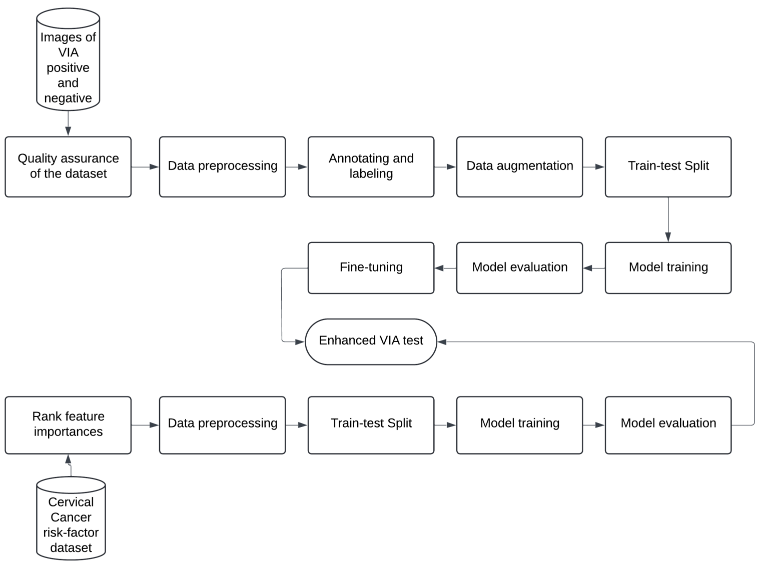

Figure 4.1: Work Flow of the research

4.2 Model Architecture

CNN

Convolutional Neural network (CNN) is the main model architecture of our paper. This is a kind of neural network and this network learns feature engineering by itself by the filters optimization. This is the extended version of Artificial Neural Network (ANN). From images, they can extract the detailed pattern of the characteristics and representations and they can automatically learn from them through the automated learning. This indicated that they use less preprocessing than other image classification algorithms. This neural network is a regularized type of feed-forward neural network (FNN) which indicates that they are fully connected networks like each neuron of one layer is connected to all neurons of another layer.CNN has been used in various kinds of tasks like image and video recognition, image classification, segmentation, medical image analysis etc. Convolutional Neural Network consists of some layers like input layer,convolutional layer, pooling layer, dense layer, output layer etc. They are very good at detecting patterns and also features in images, videos, audios etc. Moreover they can handle large amounts of data and also give higher accuracy.

4.2.1 YOLOv9

YOLOv9 is the latest version of YOLO architectures which was released on february 21, 2024 (Wang, Yeh, & Liao,). YOLO is a Convolutional Neural Network which is very efficient in detecting objects.YOLO(You Only Look Once) is an architecture family which does object detection highly efficiently and they have the capability of detecting objects in any images, video frames etc with a very remarkable speed. As their name says, detecting objects in a single pass through a neural network is their main characteristic. YOLO's main strength in object detection is that the model is very small. Moreover, its calculation is fast as it just needs to put the image into the network to get the detection output. YOLO algorithm can learn high generalized features which means it has a strong generalization ability. YOLO has 24 convolutional networks and two fully connected layers follow them. Though it is very good at detecting objects, it gives poor results if the objects are too close to each other (Jiang, Ergu, Liu, Cai, & Ma,).

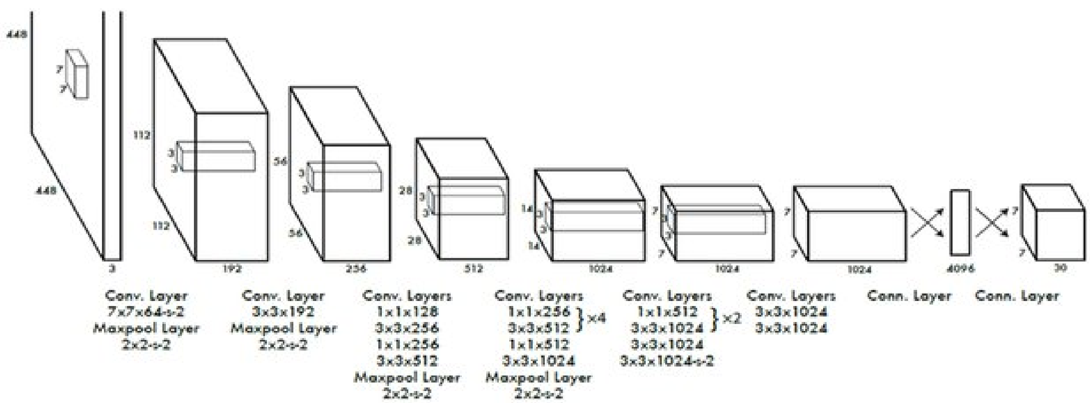

Figure 4.2: YOLO structure in detail

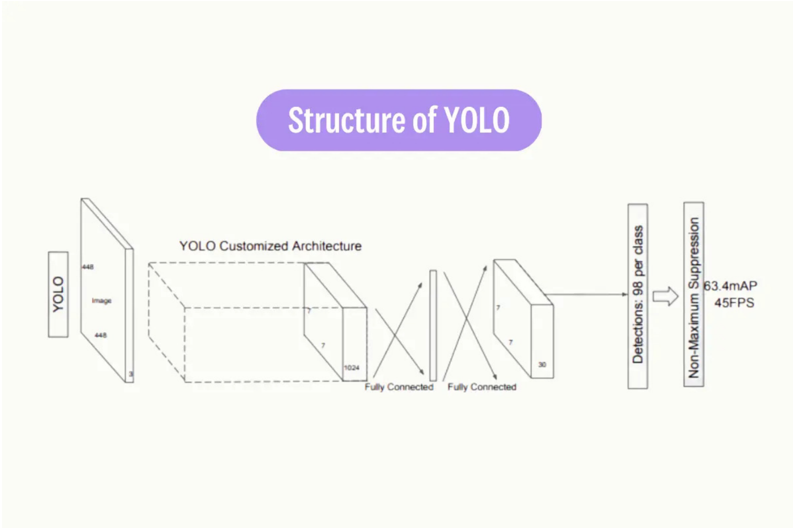

Figure 4.3: YOLO structure

There are several versions of YOLO models and each of the versions has some improvements over the previous one. Though, they have changes among them but the main architecture and principles remain the same in different versions of YOLO. Such as: YOLOv2, YOLOv3, YOLOv4, YOLOv5,YOLOv6, YOLOv7,YOLOv8 and the latest one is YOLOv9. Version to version they have improved in various aspects. YOLOv2 improves in better and faster object detection (Jiang et al.,). YOLOv3 significantly improves accuracy rate for object detection and in different sizes of objects. YOLOv4 is basically designed for optimal speed and accuracy in detecting objects which makes it suitable for real-time applications also. This model uses CSPDarknet53 as its backbone, PANet as the neck and YOLOv3 as the detection head. YOLOv5 secures an adaptive detecting mechanism based model which is more flexible in different scenarios. For the applications which demand fast response, YOLOv5 is more efficient for them as it delivers a balance in accuracy-speed tradeoff. Moreover, YOLOv5 delivers an abundance of pre-trained models which executes several detection tasks like inference, validation, training and export all together. YOLOv6 improves in accuracy and speed as well for real time applications. The main improvement is the cutting-edge object detection system. YOLOv7 improves in accuracy and speed at a higher level. It suprasses all other real-time object detectors by achieving the highest accuracy. YOLOv8 has advanced backbone and neck architectures which improves the performance of feature extraction and object detection. Moreover, it maintains a balance between accuracy and speed and also offers a variety of pre-trained models all together which makes it easier for people to use specific ones which are needed. YOLOv9 is the latest version which introduces groundbreaking techniques like Programmable Gradient Information (PGI) and the Generalized Efficient Layer Aggregation Network (GELAN) YOLOv9 achieves a significant advancement in detecting real-time objects. It has improved remarkably in efficiency, adaptability and accuracy. Moreover, this model achieves higher mAP than the other famous YOLO models like YOLOv8, YOLOv7, YOLOv5 (Wang et al.,). This model basically supports image segmentation and object detection tasks.

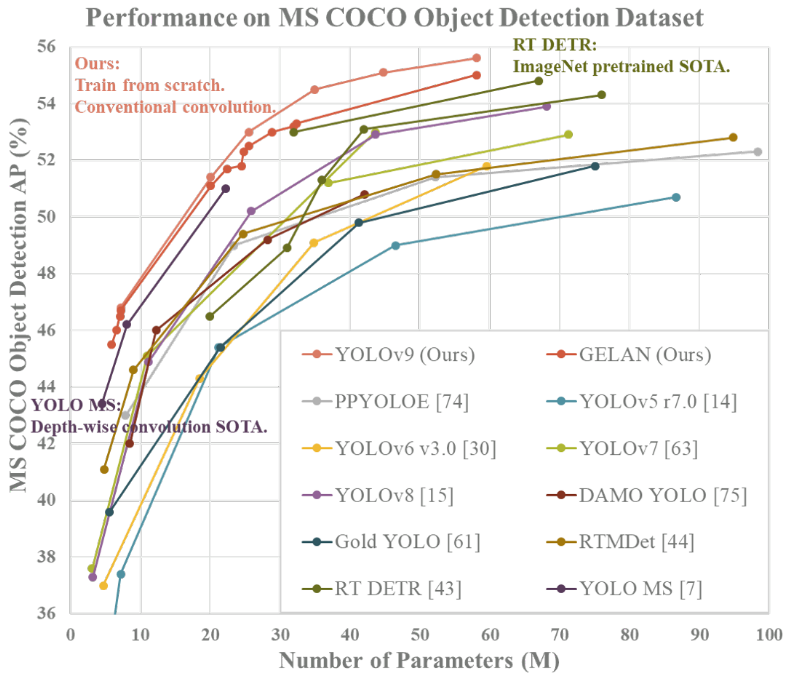

Figure 4.4: Comparison of YOLO models.

4.2.2 YOLO-NAS

YOLO-NAS (You Only Look Once - Neural Architecture Search) is another groundbreaking object detection model of the YOLO family. This is an advanced Neural Architecture Search Technology which addresses the limitations of previous YOLO models. Moreover, in accuracy-speed trade off and quantization support this model has made a great improvement which lead it to do an advancement and improve a lot in object detection. This model is more suitable for smaller datasets. This model offers three variants which are small (s), medium(m), large (l) etc according to the need of different computational and performance. Among them, for a smaller dataset YOLO-NAS-m is more suitable as it prevents overfitting of data. Moreover, this model can do faster training when the dataset is medium. If there are only a few thousands to ten thousands of data like annotated images, YOLO-NAS-m is more suitable for this type of datasets as it maintains a balance between the complexity of the model and performance.

4.2.3 VGG16

VGG16, referring to the VGG model, is also known as VGGNet. VGG (Visual Geometry Group) is a kind of Convolutional Neural Network (CNN) that was developed for enhancing the model performance of CNN by incrementing its depth. Therefore, the VGG model has deep layers such as 16 and 19 which define VGG16 and VGG19 respectively. This is one of most popular image recognition models now-a-days. The VGG architecture consists of input, convolutional layers, hidden layers and fully connected layers. The VGG16 model supports 16-layer structures. In VGG16, there are 13 convolutional layers and 3 fully connected layers. Moreover, VGG16 can achieve up to 92.7 percent test accuracy in ImageNet, a dataset containing more than 14 million training images across 1000 object classes (Simonyan

& Zisserman,). This deep learning model consists of multiple layers, incorporating convolutional, pooling, and fully connected networks. It's renowned for its capability to achieve remarkable results on various computer vision tasks, including image classification and object detection while being a simple yet effective model. The simplicity of VGG16's architecture makes it more appealing to use. Its architecture enables the model to learn intricate hierarchical presentation of visual features, leading to robust and faultless prediction. The VGG-16 model is still a preferable choice for many deep-learning implementations, despite its simplicity, due to its adaptability and effectiveness.
(Fernandes et al.,)

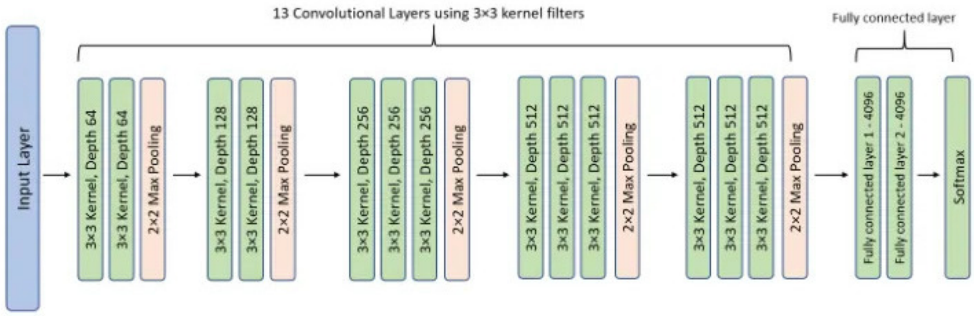

Figure 4.5: VGG16 architecture

4.2.4 ResNet-50

ResNet-50 is a kind of Convolutional neural Network which has 50 layers. This model is of the ResNet (Residual Networks) family, a deep learning model which is basically used for computer vision. They are a kind of Artificial Neural network (ANN) which forms networks by piling residual blocks. The purpose of designing this CNN model is to support hundreds or thousands of convolutional networks which couldn't be performed before by other models because of their limited performance. Moreover, the ResNet model can vanish the gradient problem also which is called "skip connections" or "shortcut connections". This feature enables direct flow of information from one layer to another. Moreover, two basic rules that the ResNet model follows. One of them is that in each layer, the number of filters are the same according to the size of the output feature map. The other rule is the number of filters will be doubled if the feature map's size becomes halved. This rule sustains each layer's time complexity. ResNet architecture has various depth models such as ResNet-18, ResNet-32, ResNet-50 and ResNet-101. Among these, ResNet-50 is notable for its proficiency and depth in tasks like image classifications. This model first came noticeable in 2015 by a paper written by a group of researchers (He, Zhang, Ren, & Sun,). Among the 50 layers of the ResNet-50 model, 48 layers are the convolutional layers, 1 layer is maxPool layer and 1 is average pool layer.

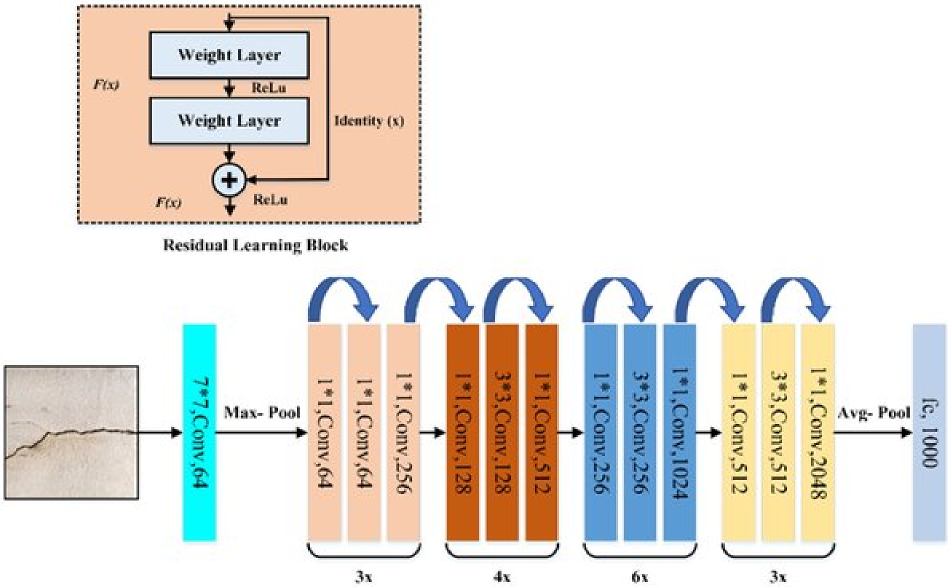

Figure 4.6: ResNet-50 architecture

Machine Learning Algorithms

4.2.5 Support Vector Machine

Support Vector Machine(SVM) , a supervised machine learning algorithm which is used mainly for classification and regression. Moreover, it can do various tasks like text and image classification, handwriting identification, face and anomaly detection etc. Finding the optimal hyperplane in an N-dimensional space which can separate the data points in different classes in the feature space is the main objective of this model. The SVM model is very effective and efficient in high-dimensional cases.As this model manages high dimensional data and non-linear relationships, this model is also considered as adaptable. Moreover, memory of this model is also proficient because of using support vectors as a subset of training points in the decision function. By this model, specifying custom kernels is also possible.

4.2.6 Random Forest

Random Forest, a strong tree learning technique in machine learning algorithms which is broadly used for classification and regression functions. During its training phase, this algorithm works by creating a number of decision trees where every tree is built using a random subset of the dataset. These trees are created to measure a random subset of features in each partition. This feature of the algorithm reduces risks of overfitting and improves overall prediction of the performance. Therefore, it delivers high predictive accuracy. Moreover, this algorithm handles complex data and missing values as well. The Random Forest algorithm also handles large datasets and makes the whole process faster and more efficient.

4.2.7 Logistic Regression

Logistic regression is a supervised machine learning algorithm that models the probability of a discrete outcome given an input variable. It is a fundamental statistical algorithm that is used for binary classification problems. Logistic regression is widely used for predictive and classification problems. It can perform well with smaller datasets. In this data analysis method, mathematical models are used to define the relationships between two data factors. It is simple to implement, interpret, and very effective to train. It takes independent variables as input and outputs a probability value between 0 and 1. However, the output of this algorithm must be a categorical or discrete value, as logistic regression estimates the output of a categorical dependent variable. Moreover, it classifies unknown variables very efficiently. It performs well while the data set is linearly separable. Additionally, training a logistic regression model is comparatively faster and more computationally efficient than some other machine learning models.

4.2.8 Gradient Boosting

Gradient boosting is a widely used machine learning algorithm which is used for both regression and classification problems. This boosting algorithm is one kind of ensemble learning algorithm that trains the models sequentially with each new model trying to precise the previous model. Several weak learning algorithms are combined to form a strong learning algorithm in this machine learning ensemble approach.

In this powerful boosting algorithm each latest model is trained to minimize the loss function- such as cross-entropy or mean squared error of the previously trained model using gradient descent. It iteratively combines multiple predictions of weak learners, typically decision trees, sequentially to form strong learners. Besides, this algorithm stands out for its prediction speed and precision, particularly on huge and complex datasets. In each iteration this algorithm evaluates the gradient of the loss function of the current ensemble model's prediction and then trains a new weak model to minimize this gradient. This method refines a model's accuracy by optimizing the model's weights based on the errors of previous iterations, leading to increased accuracy and decreasing prediction errors gradually.

Chapter 5

Results And Discussion

5.1 Results

Image dataset

5.1.1 YOLOv9

For the model YOLOv9, we get 93.8% precision and 95.1% recall with 99.1% mAP50 in 100 epochs. It can be observed from Figure 5.1 that the model is becoming more accurate in identifying only the correct objects and improvement in recall suggests the model is better at detecting relevant objects. Further, mean Average Precision at IoU threshold of 0.5(mAP50) shows overall improvement detection accuracy and mean Average Precision over multiple IoU thresholds(mAP50-95) indicate robust performance across different IoU levels. Overall, the model exhibits strong and consistent performance in classification, precision, recall, and detection accuracy.

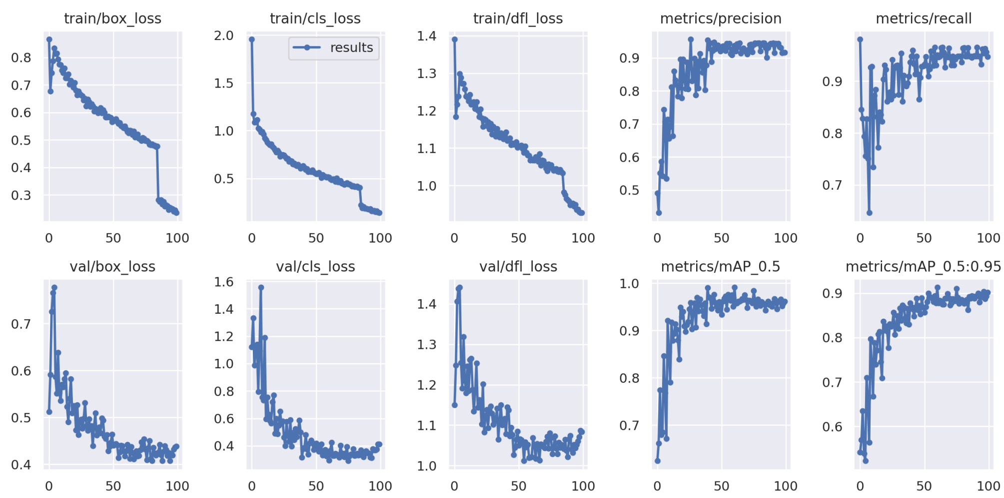

Figure 5.1: YOLOv9 metrics

5.1.2 YOLO-NAS

For model YOLO-NAS, we get 91.7% precision and 96.6% recall with 99.7% mAP50 in 100 epochs. It can be observed from Figure 5.1 that the model is becoming more accurate in identifying only the correct objects and improvement in recall suggests the model is better at detecting relevant objects. Further, mean Average Precision at IoU threshold of 0.5(mAP50) shows overall improvement detection accuracy and mean Average Precision over multiple IoU thresholds(mAP50-95) indicate robust performance across different IoU levels. Overall, the model exhibits strong and consistent performance in classification, precision, recall, and detection accuracy.

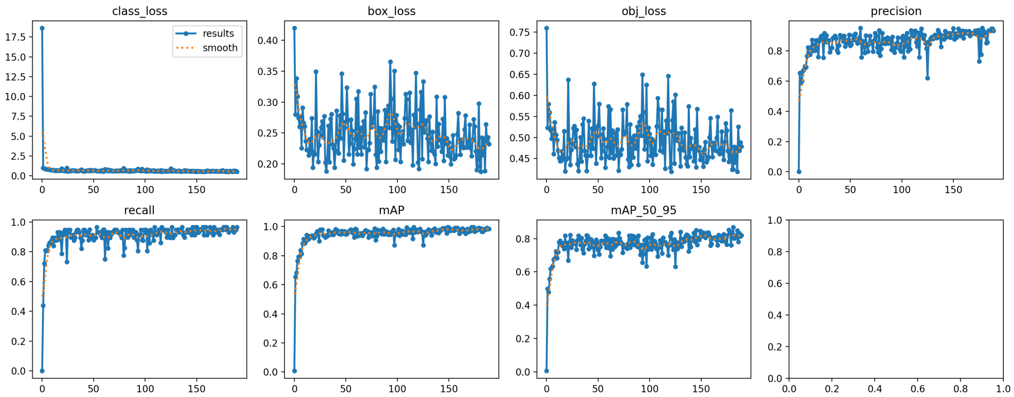

Figure 5.2: YOLO-NAS metrics

5.1.3 VGG16

Model Accuracy
For model VGG16, we can see in figure 5.5, Initially the train accuracy was approximately 0.73. When the epoch reaches 50, the training accuracy reaches its optimal which is approximately 0.99. Now, if we look at the validation accuracy, initially it was 0.56. On epoch 10, the validation accuracy is approximately 0.86. After 50 epochs, the validation accuracy comes to approximately 0.95.

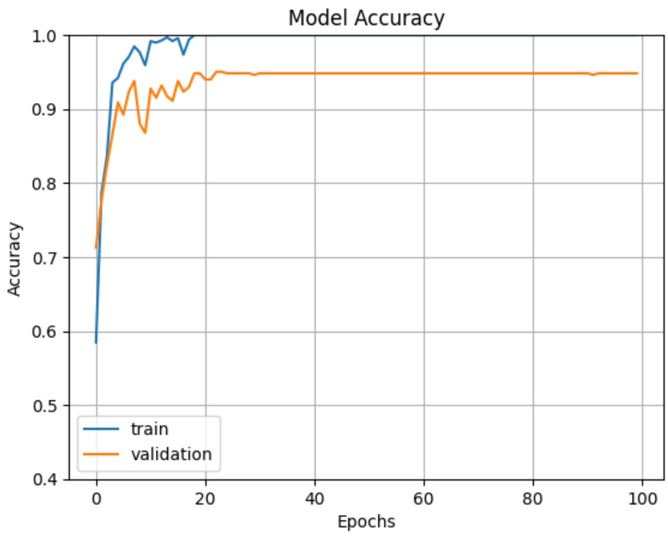

Figure 5.3: VGG16 Model Accuracy curve

Model Loss For the model VGG16, we can see in the figure 5.7 initially the train loss shows approximately 13. The train loss decreases significantly through 5-20 epochs. From 20-50 epoch there are some spikes. At 50 epochs the train loss is approximately 0.1. Now if we look at the validation loss, at the initial stage, it is 18. At 50 epochs it reaches to near 0.33 approximately.

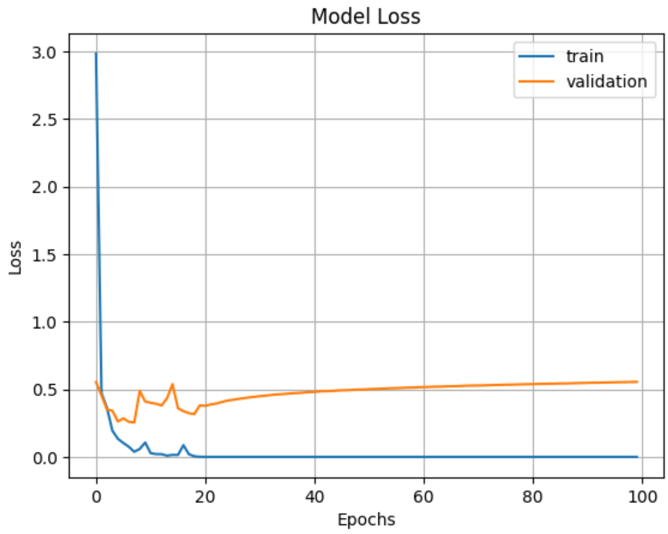

Figure 5.4: VGG16 model loss

5.1.4 ResNet-50

Model Accuracy

For the model ResNet-50, we can see in the figure 5.6, initially the training accuracy was approximately 0.63. When the graph reaches 10 epochs, the training accuracy reaches approximately 0.92. Now if we look at the validation accuracy, initially it is approximately 0.69. After 10 epochs it reaches to approximately 0.88. After 50 epochs it reaches 0.95.

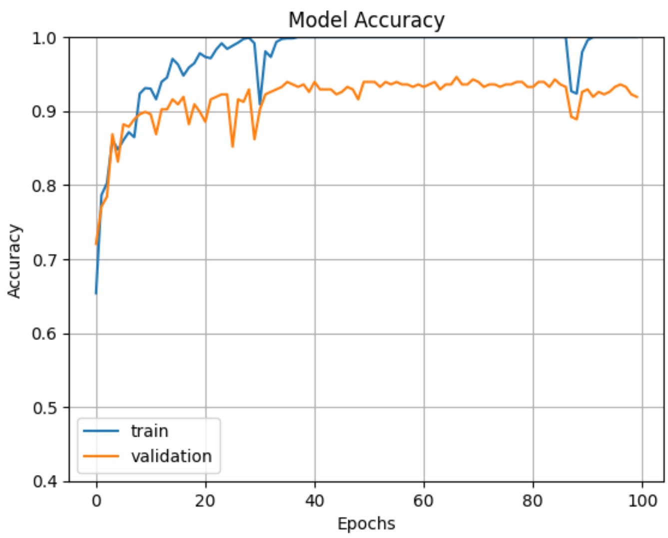

Figure 5.5: ResNet-50 Model Accuracy curve

Model Loss

For the model ResNet-50, we can see in the figure 5.7 initially the train loss shows approximately 0.78. The train loss decreases significantly through 5-20 epochs. From 20-50 epoch there are some spikes. At 50 epochs the train loss is approximately 0.1. Now if we look at the validation loss, at the initial stage, it is 0.55. At 50 epochs it reaches to near 0 approximately.

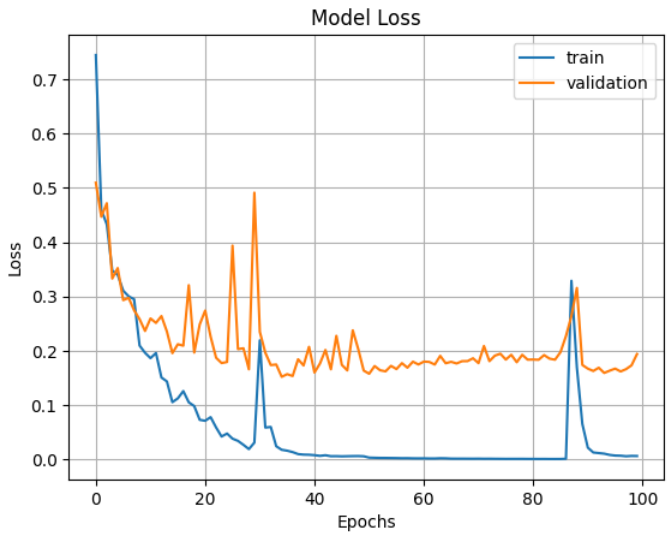

Figure 5.6: ResNet-50 Model Loss curve

Patients History Dataset

5.1.5 Random Forest

Feature Importance

Random Forest was used to determine the feature importances. Random Forest is robust to overfitting due to the averaging of multiple decision trees. RF handles Non-linear Relationships. It can capture complex interactions between features, making it effective for datasets with non-linear relationships (Breiman,). Top features are hormonal contraceptives (years), number of pregnancies and number of sexual partners. Hormonal contraceptives (years) is the most important predictor. It indicates the duration of hormonal contraceptive use has a significant impact on cervical cancer. Moderately important Features are duration of smoking and use of IUD (Intrauterine Device). Other contributing Features are categories of first sexual intercourse, certain categories of age and status of hormonal contraceptives and smoking. These findings align with known medical knowledge about the risk factors for cervical cancer.

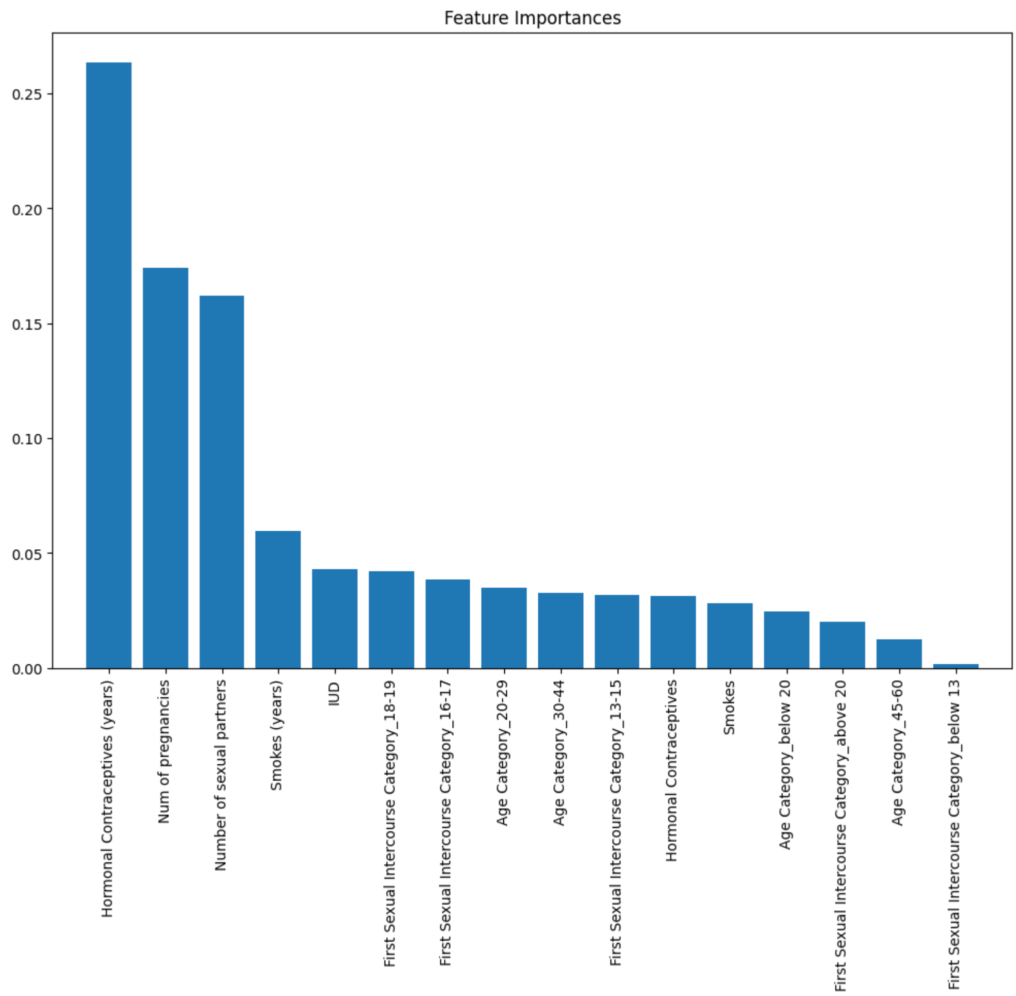

Figure 5.7: Feature Importances

Classification Report

The Random Forest classifier shows excellent performance in predicting cervical cancer risk based on our patient history dataset. In table 5.1, we can see it demonstrates high accuracy and reliability. For negative biopsy, of all instances predicted as negative, 99% were correct. And of all actual negative instances, 95% were correctly predicted. F1 score of 0.97 indicates a balanced performance. On the other hand, for positive biopsy, of all instances predicted as positive, 94% were correct. And of all actual positive instances, 99% were correctly predicted. F1 score of 0.97 for positive biopsy also indicates a balanced performance. In table 5.2, we can see the Mean cross-validation is 97.5% indicating consistent performance across all folds.

Metric	Class 0 (Negative Biopsy)	Class 1 (Positive Biopsy)	Overall
Precision	0.99	0.94	0.97
Recall	0.95	0.99	0.97
F1-Score	0.97	0.97	0.97
Support	129	120	249
Accuracy			0.97
Macro Average	0.97	0.97	0.97
Weighted Average	0.97	0.97	0.97

Table 5.1: RF Performance Table

Cross-Validation Scores	Mean Cross-Validation Score	Standard Deviation
[0.98393574, 0.97590361, 0.97590361, 0.97188755, 0.96774194]	0.9751	0.0054

Table 5.2: Cross-Validation Summary

Confusion Matrix

In figure 5.8, we can see the model correctly predicted 119 positive biopsy results (True Positive) and incorrectly predicted 7 negative biopsy results as positive (False Positive). On the other hand, the model correctly predicted 122 negative biopsy results (True Negative) and incorrectly predicted 1 positive biopsy result as negative (False Negative).

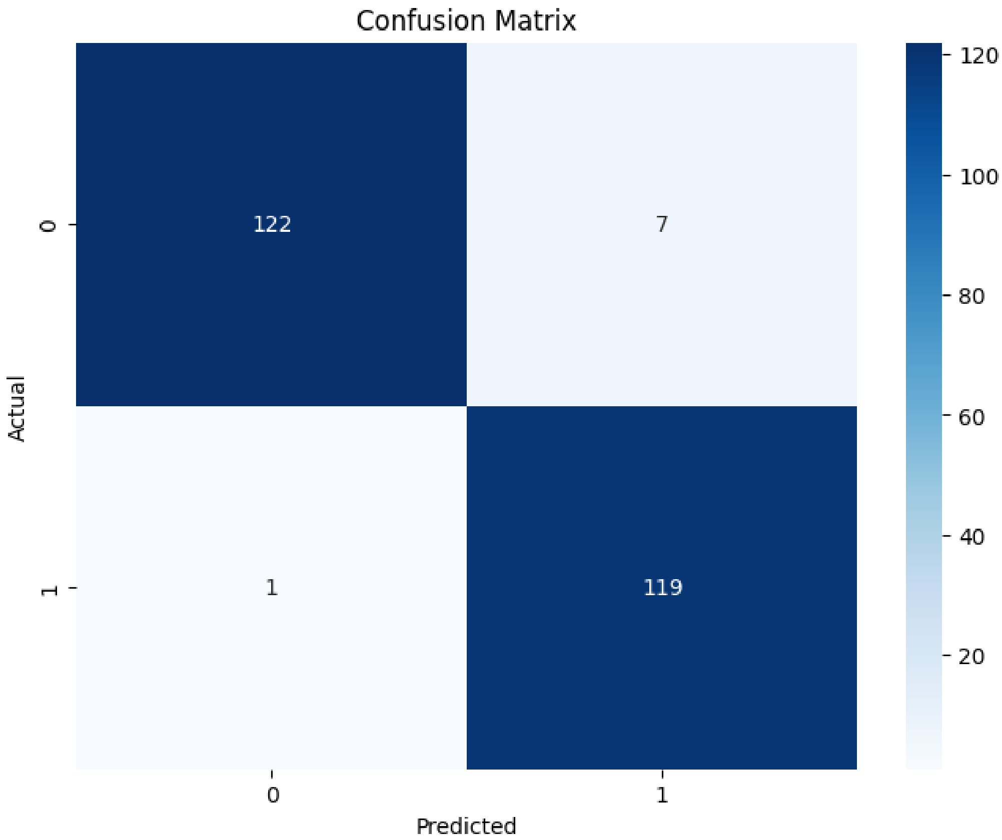

Figure 5.8: RF Confusion Matrix

5.2 Discussion

5.2.1 Computer Vision Models

For the image dataset, the analysis tried to make predictions using the models covered below as shown in the table 5.1 and 5.2. The study found that the chosen algorithms are reasonably effective at detecting cervical abnormalities and providing a verdict of VIA positive or negative. YOLOv9 obtains slightly better precision and a more consistent performance across different IoU thresholds which makes it the preferred model for high accuracy. Both classification models exhibit strong performance. Although VGG16 has a slight edge in both training and validation accuracy. High accuracy metrics for both models shows their effectiveness in feature extraction and classification tasks. Although a marginal difference in validation accuracy suggests that VGG16 might be slightly better suited for our specific classification problem.

Object detection models	Precision	Recall	mAP50	mAP50-100
YOLOv9	0.93	0.95	0.99	0.91
YOLO-NAS	0.91	0.96	0.99	0.82

Table 5.3: Object detection models

Classification models	Precision	Recall	F1 Score	Training Accuracy	Validation Accuracy	AUC
VGG16	0.96	0.93	0.94	0.99	0.94	0.98
ResNet-50	0.95	0.91	0.93	0.99	0.93	0.99

Table 5.4: Classification models

5.2.2 Machine Learning Models

To identify patients at risk of developing cervical cancer, In table 5.3 we can see Random Forest outperforms all the other models. RF achieved the highest mean cross-validation score of 97.51% and the lowest deviation of 0.0054. This indicates not only high accuracy but also consistent performance across different subsets of the data. On the other hand Logistic Regression has the lowest performance among all the models. LR achieved the lowest mean score of 54.90% and a standard deviation of 0.0257. This means that linear models may not capture non-linear relationships in the data. The Ensemble model also performs well with a mean score of 93.97% and a standard deviation of 0.0144. Gradient Boosting also shows good performance with a mean score of 87.14% but compared to Random Forest and the Ensemble model, it has a higher standard deviation of 0.0271. The SVM model performs moderately with a mean score of 70.10% and a standard deviation of 0.0244. Figure 5.9 shows the model performance comparison.

Model	Cross-Validation Scores	Mean Cross-Validation Score	Standard Deviation of Cross-Validation Scores
Random Forest	[0.98393574, 0.97590361, 0.97590361, 0.97188755, 0.96774194]	0.9751	0.0054
SVM	[0.73092369, 0.69076305, 0.66666667, 0.72690763, 0.68951613]	0.7010	0.0244
Logistic Regression	[0.56626506, 0.53815261, 0.53012048, 0.59036145, 0.52016129]	0.5490	0.0257
Gradient Boosting	[0.89156627, 0.87148594, 0.81927711, 0.88353414, 0.89112903]	0.8714	0.0271
Ensemble	[0.96385542, 0.92369478, 0.93172691, 0.94779116, 0.93145161]	0.9397	0.0144

Table 5.5: Performance of ML models based on Cross-Validation

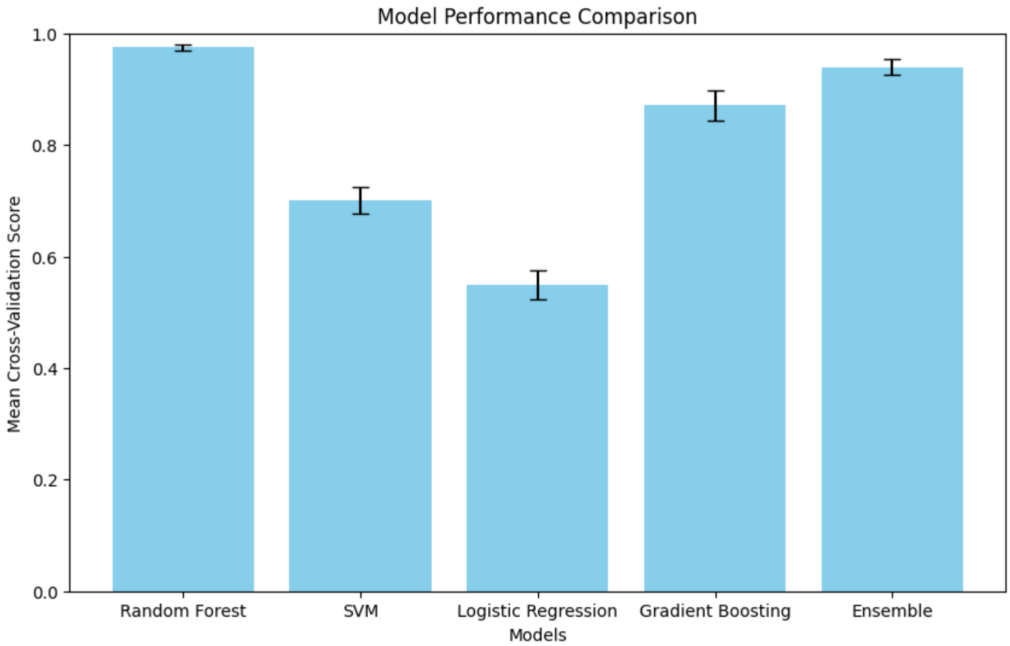

Figure 5.9: Model Performance Comparison

Chapter 6

Conclusion and Future Works

6.1 Conclusion

In this research, we tried to investigate the capability of machine learning and deep learning algorithms to strengthen VIA screening for cervical cancer. CNN models like YOLOv9, YOLO-NAS, VGG16, and ResNet50 have been implemented to examine cervix images from the VIA procedure. Whereas, logistic regression, SVM, random forest, and gradient boosting were utilized to analyze patient data to compare with risk factors and predict the future risk of developing cervical cancer because of a risky lifestyle. Visual Inspection with Acetic Acid is undoubtedly a useful tool for screening cervical cancer in underdeveloped and low-resource countries. Through our research, we tried to address the VIA screening method subjectivity and human error connected with the naked-eye examination, by leveraging deep learning models. While our research evidences the promising possibility of deep learning and machine learning favoring VIA screening for cervical cancer, ensuring data integrity remains predominant. Given that the VIA screening procedure is highly sensitive, prioritizing the integration of robust security-ensuring procedures has been a great concern for us. Data privacy and security have been significant issues in every healthcare setting worldwide. Therefore, data privacy and security policy for medical services should explicitly include regulations for data ownership, sovereignty, access and control over data usage and storage (Zahid, Sharma, Wingreen, & Inthiran,).

6.2 Future Works

Visual Inspection of cervix with Acetic Acid (VIA) is a test that can be used to extract a lot more information from the information that we are currently extracting. Using our model, a robust system can be built connecting the risk factor information with the images of the cervix. A larger dataset of demographic, Behavioral and clinical factors with correlated VIA test results can be used to train machine learning models to predict VIA results. It can also help make a comprehensive database of patients at high risk of being cervical cancer positive and encourage them to be interested in regular healthcare checkups. This can result in personalized medicine which is proven to work better than the generalized medicine. Further, the Computer Vision models can be trained to predict ablation eligibility to completely automate the VIA test procedure.

Bibliography

Adhikari, T. (2023, January). Designing a convolutional neural network for image recognition: A comparative study of different architectures and training techniques. *Social Science Research Network*. doi: 10.2139/ssrn.4366645

Alam, N. E., Shariful, I., Rayyan, F., Ifa, H. N., Khabir, M. I. U., Chowdhury, K., Mohiuddin, A. K. M. (2022, January). Lack of knowledge is the leading key for the growing cervical cancer incidents in bangladesh: A population based, cross-sectional study. *PLOS Global Public Health*, *1*.

Allahqoli, L., Laganà, A. S., Mazidimoradi, A., Salehiniya, H., Günther, V., Chiantera, V., ... Alkatout, I. (2022, November). Diagnosis of cervical cancer and pre-cancerous lesions by artificial intelligence: A systematic review. *Diagnostics*, *12*(11), 2771.

Arbyn, M., Weiderpass, E., Bruni, L., De Sanjosé, S., Saraiya, M., Ferlay, J., Bray, F. (2020, February). Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis. *The Lancet Global Health*, *8*(2), e191-e203. doi: 10.1016/s2214-109x(19)30482-6

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R., Torre, L., Jemal, A. (2018). Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, *68*(6), 394-424. doi: 10.3322/caac.21492

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. doi: 10.1023/A:1010933404324

Europe pmc. (n.d.). Retrieved from `https://europepmc.org/article/med/17679017`

Fernandes, K., Cardoso, J., Fernandes, J. (2017). *Cervical cancer (risk factors).* UCI Machine Learning Repository. doi: 10.24432/C5Z310

Guo, P., Singh, S., Xue, Z., Long, R., Antani, S. (2019). Deep learning for assessing image focus for automated cervical cancer screening. In *2019 ieee embs international conference on biomedical health informatics (bhi)* (p. 1-4). doi: 10.1109/BHI.2019.8834495

He, K., Zhang, X., Ren, S., Sun, J. (2016, June). Deep residual learning for image recognition. In *Ieee conference publication.* IEEE. Retrieved from `https://ieeexplore.ieee.org/document/7780459`

Hou, X., Shen, G., Zhou, L., Li, Y., Wang, T., Ma, X. (2022, March). Artificial intelligence in cervical cancer screening and diagnosis. *Frontiers in Oncology*, *12*.

International Agency for Research on Cancer. (n.d.). *Atlas of visual inspection of the cervix with acetic acid for screening, triage, and assessment for treatment.* Retrieved from https://screening.iarc.fr/atlasviadetail.php?Index=29&e=#0

International Agency for Research on Cancer. (2022). *Cervical cancer awareness month 2022.* Retrieved from https://www.iarc.who.int/infographics/cervical-cancer-awareness-month-2022/

Jiang, P., Ergu, D., Liu, F., Cai, Y., Ma, B. (2022). A review of yolo algorithm developments. *Procedia Computer Science*, *199*, 1066-1073. doi: 10.1016/j.procs.2022.01.135

Leonardo, M. M., Carvalho, T. J., Rezende, E., Zucchi, R., Faria, F. A. (2018). Deep feature-based classifiers for fruit fly identification (diptera: Tephritidae). In *2018 31st sibgrapi conference on graphics, patterns and images (sibgrapi)* (p. 41-47). doi: 10.1109/SIBGRAPI.2018.00012

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... Sánchez, C. I. (2017, December). A survey on deep learning in medical image analysis. *Medical Image Analysis*, *42*, 60-88.

Lohiya, A., Daniel, R., Kumar, D., Varghese, C., Rath, R., S.A, R., Nongkynrih, B. (2022, February). Effectiveness of visual inspection with acetic acid (via) screening on cervical cancer mortality and incidence - a systematic review and meta-analysis. *Asian Pacific Journal of Cancer Prevention*, *23*(2), 399-407. doi: 10.31557/apjcp.2022.23.2.399

National Cancer Institute. (n.d.). *Nci dictionary of cancer terms.* Retrieved from https://www.cancer.gov/publications/dictionaries/cancer-terms/def/transformation-zone

Ontor, M. Z. H., Ali, M. M., Ahmed, K., Bui, F. M., Al-Zahrani, F. A., Mahmud, S. M. H., Azam, S. (2023, January). Early-stage cervical cancerous cell detection from cervix images using yolov5. *Computers, Materials & Continua*, *77*. doi: 10.32604/cmc.2023.032794

Qureshi, R., Ragab, M. G., Abdulkader, S. J., Muneer, A., Alqushaib, A., Sumiea, E. H., Alhussian, H. (2023, July). *A comprehensive systematic review of yolo for medical object detection (2018 to 2023).* doi: 10.36227/techrxiv.23681679.v1

Sankaranarayanan, R., Nessa, A., Esmy, P. O., Dangou, J. M. (2012, April). Visual inspection methods for cervical cancer prevention. *Baillière's Best Practice & Research. Clinical Obstetrics & Gynaecology/Baillière's Best Practice and Research in Clinical Obstetrics and Gynaecology*, *26*(2), 221-232. doi: 10.1016/j.bpobgyn.2011.08.003

Simonyan, K., Zisserman, A. (2014, September). *Very deep convolutional networks for large-scale image recognition.*

Sung, H., Ferlay, J., Siegel, R., Laversanne, M., Soerjomataram, I., Jemal, A., Bray, F. (2021). Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, *71*(3), 209-249. doi: 10.3322/caac.21660

Tan, S. L., Selvachandran, G., Ding, W., Paramesran, R., Kotecha, K. (2023, November). Cervical cancer classification from pap smear images using deep convolutional neural network models. *Interdisciplinary Sciences: Computational Life Sciences*. doi: 10.1007/s12539-023-00589-5

Uddin, A. F. M. K., Sumon, M. A., Pervin, S., Sharmin, F. (2023). Cervical cancer in bangladesh. *South Asian Journal of Cancer*, *12*(1), 36-38. doi: 10.1055/s-0043-1764202

Viñals, R., Jonnalagedda, M., Petignat, P., Thiran, J. P., Vassilakos, P. (2023, February). Artificial intelligence-based cervical cancer screening on images taken during visual inspection with acetic acid: A systematic review. *Diagnostics*, *13*(5). doi: 10.3390/diagnostics13050836

Wang, C. Y., Yeh, I. H., Liao, H. Y. M. (2024, February). *Yolov9: Learning what you want to learn using programmable gradient information.*

Wijaya, H., Ryanto, T. R. T., Sidhik, S. C. S. S. C., Liawatimena, S. L. S., Riana, D. R. D. (2024, January). *Cervical cancer cells detection using yolov8 algorithm.* doi: 10.2139/ssrn.4779266

World Health Organization. (2020). *National strategy for cervical cancer prevention and control in bangladesh, 2017–2022.*

World Health Organization. (2024, March). *Cervical cancer.* Retrieved from `https://www.who.int/news-room/fact-sheets/detail/cervical-cancer?gclid=CjwKCAiAzJOtBhALEiwAtwj8tlI4nrX8KFgaMmg_GSoC1mrV_GGdsvzgCKjfsby3dK5fRDILz7lSvRoCWZsQAvD_BwE`

Zahid, A., Sharma, R., Wingreen, S. C., Inthiran, A. (2022). Soft systems modelling of design artefacts for blockchain-enabled precision healthcare as a service. In *Iceb 2022 proceedings (bangkok, thailand).* Retrieved from `https://aisel.aisnet.org/iceb2022/42`