JIMMA UNIVERSITY

JIMMA INSTITUTE OF TECHNOLOGY

SCHOOL OF BIOMEDICAL ENGINEERING

BIOMEDICAL IMAGING PROGRAM

M.Sc. Thesis

On

Cervix Type and Cervical Cancer Classification Using Deep Learning Technique

Thesis Submitted to School of Post Graduate Studies of Jimma University in Partial Fulfillment for the Degree of Master of Science in Biomedical Imaging

By: Lidiya Wubshet

Advisor: Gizeaddis L.Simegn (Ph.D.)

Co-advisor: Ms. Elbetel Taye (M.Sc.)

February, 2022

Jimma, Ethiopia

Declaration

I hereby declare that this research entitled **'Cervix Type and Cervical Cancer Classification Using Deep Learning Technique**' is my original work and has not been presented by any other for award of a degree in this and any other university.

Done by:

Lidiya Wubshet　　　　　　　——————————　　　　　——————————
　　　　　　　　　　　　　　　　　Signature　　　　　　　　　　　Date

Approved by:

1. Gizeaddis L.Simegn (Ph.D.)　　　　　　　　　　14/02/2022
　　　　　　　　　　　　　　　　　Signature　　　　　　　　　　　Date

2. Ms. Elbetel Taye (Msc.)　　——————————————　　　——————————
　　　　　　　　　　　　　　　　　Signature　　　　　　　　　　　Date

Examiners:

External Examiner: Dr. Million Meshesha　　　　　09/02/2022
　　　　　　　　　　　　　　　　　Signature　　　　　　　　　　　Date

Internal Examiner: Mr. Kokeb Dese　　——————————　　　——————————
　　　　　　　　　　　　　　　　　Signature　　　　　　　　　　　Date

Chair Name:　　　　　　　　——————————————　　　——————————
　　　　　　　　　　　　　　　　　Signature　　　　　　　　　　　Date

Acknowledgment

First and foremost, I would like to thank God for being with me in every situation and for blessing me much more than I deserve. I also acknowledge my advisor Dr. Gizeaddis Lamesgin and co-advisor Ms. Elbetel Taye for their efforts and time in guiding me for the development and implementation of this study.

I would like to thank clinical collaborators in JUMC Dr. Gebi (pathologist) and Dr. Habib Ibrahim (pathologist resident) for their unlimited assistance throughout the research period. Besides, I would like to thank the pathology department staff in St. Paul hospital. My heartfelt gratitude will also go to Tercha general hospital staff and clinical collaborators Dr. Wondmagegn Sisay (Gynecologist) and Mr. Wubshet H/Mariam (IESO) for their continuous assistance.

Finally, I would like to thank my family and friends for encouraging me as well as sharing ideas through this research period.

Abstract

Cervical cancer is the most common cancer among women worldwide. It is screened using visual inspection after application of acetic acid (VIA), Papanicolaou (Pap), and Human papillomavirus (HPV) test and diagnosed by histopathology test. Currently, identification of cervix type and histopathologic image classification is examined manually, thus big inter-and intra- observer variability exists. Moreover, incidence and mortality rates are consistently increasing in developing countries due to a shortage of screening facilities, skilled professionals, and lack of awareness. Therefore, there is a need for an automated system that overcomes subjectivity and inconsistency in the screening process. Different studies have proposed the classification of cervix type and cervical cancer using deep learning and machine learning techniques. However, most of the proposed techniques in the literature are limited to specific class classification. This study aims to develop an automatic cervix type and cervical cancer classification using deep learning techniques. For cervix type classification 3872 colposcopy images were collected from the Kaggle dataset. In addition, 133 colposcopy images and 915 histopathologic images were collected from local health facilities for cervix and cervical cancer classification from local health facilities. Different pre-trained models including VGGNet, ResNet, Ensemble of MobileNet, and ResNet and EfficientNet were trained and compared for cervix type classification. Prior to classification, a lightweight MobileNetv2-YOLOv3 model was created for the region of interest (ROI) extraction. The extracted cervix images were then fed to the effecientnet_b0 model for cervix type classification. A mean average precision (MAP) of 99.88% for the region of interest (ROI) extraction, an accuracy of 97%, and a Kappa score of 0.95 for the cervix type classification task were achieved using the selected models. For cervical cancer classification, an efficient_b0 pre-trained model was trained and validated using histogram matched histopathological images, and an accuracy of 95% and Kappa score of 0.92 were achieved. Even though the best performing model was constructed further improvement needs to be done using more data and by classifying sub-class of histopathology images to know disease prognosis. In addition, to make full cervical cancer screening and diagnosis method cytology image classification model construction is required.

Key words: Cervical Cancer, Cervix Type, Histopathology Image, Cervigram, Region of Interest (ROI) Extraction, Transformation Zone, Deep Learning

Table of Content

List of Figures

List of Tables

Acronym

ATZ   Atypical Transformation Zone

AUC   Area Under receiver operating Characteristics

CAD   Computer-Aided Diagnosis

CHIC   Cervical Histopathology Image Classification

CIN   Cervical Intraepithelial Neoplasia

CNN   Convolutional Neural Network

DNN   Deep Neural Network

GUI   Graphical User Interface

JUMC   Jimma University Medical Center

HCME   Hierarchical Convolutional Mixture of Experts

HSIL   High Squamous Intraepithelial Lesion

HPV   Human Papilloma Virus

LDA   Linear Discriminant Analysis

LSIL   Low Squamous Intraepithelial Lesion

OPD   Outpatient Department

ROC   Receiver Operating Characteristics

ROI   Region of Interest

SIL   Squamous Intraepithelial Lesion

SVM   Support Vector Machine

WHO   World Health Organization

YOLO   You Only Look Once

CHAPTER ONE

INTRODUCTION

1.1 Background of study

Cervical cancer is one of the leading causes of cancer-related deaths in women worldwide, with 80% of the cases occurring in developing countries [1,2]. Ethiopia is one of the nations with high cervical cancer incidence and mortality rates. According to the Human papillomavirus (HPV) information center report [3], about 6,294 new cervical cancer cases are diagnosed annually in Ethiopia. It is ranked as the 2nd leading cause of female cancer in Ethiopia and the 2nd most common cancer in women aged 15 to 44 in the country and 4,884 deaths occur annually. In remote areas, which have poor medical conditions including insufficient healthcare accessibility and qualified medical staff, cervical cancer incidence and mortality are estimated to be higher. Region-specific incidence and mortality rates for cervical cancer are shown in Figure 1.1. The rates are shown in descending order of the world (W) age-standardized rate.
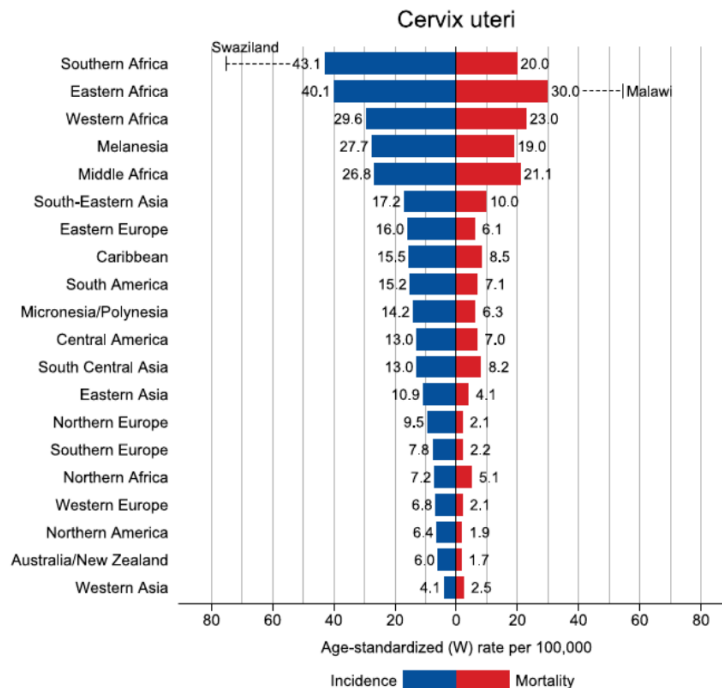
Figure 1.1: Bar chart of region-specific incidence and mortality age-standardized rates for cervical cancer in 2018 [4]

The type of cervix is identified based on transformation zone location visualizing the cervix after application of acetic acid solution or normal saline. Colposcopy is used to magnify the cervix to visualize the region of the cervix but it's not available in developing countries. However, the classification of cervix type still depends on professional skill. In developing countries, the cervix is visualized using the speculum and white light. Identifying the cervix type is important to decide treatment options and it needs skilled professionals. According to a case study [5], cervix type 3 is beyond the examiner's view and is often misclassified type of cervix. As a result, women of this cervix type are exposed to invasive cervical cancer since it may have hidden lesions. Additionally, it depends on the expert's knowledge and may lead to misdiagnosis of cervical cancer since the biopsy should be taken from this region.

On the other hand, a biopsy test is done by investigating the stained slides by pathologist visual inspection under a microscope. The process is dynamic and complex which is misidentified and misinterpreted. Visual investigation of histopathology images is highly prone to human error since the diagnostic accuracy depends on the perceptual and cognitive process of a pathologist. Due to this the interpretation of the image may lead to misdiagnosis [6].

This study focuses on classifying cervix type and cervical cancer images independently but clinically related tasks. Classifying cervix type as cervix type 1, 2, and 3 is important to make a treatment decision and to decide from which part of cervix the biopsy should be taken. Following this the biopsy is investigated under microscope whether the biopsy is normal, precancer, Adenocarcinoma, or squamous cell carcinoma.

Nowadays, technologically advanced equipment's available in the market for screening and diagnosis of cervical cancer. But due to cost, technical support and training stuff developing countries are limited to use those sophisticated equipment's. Commonly cervical cancer screening and diagnosis are done by visual inspection after application of acetic acid (VIA), pap smear, Human papillomavirus (HPV), and biopsy test. However, the mentioned screening and diagnosis methods are time taking, tedious, and prone to human error since it depends on the experience and skill of professionals.

Recently, different researches conducted to overcome the screening and diagnosis method of cervical cancer. Some works focus on cervix type classification using deep learning techniques. In

addition, for cervical cancer classification, most of the works were focused on sub-class classification and some others focused on major class classification.

1.2 Motivation of study

Recently, technological advancement in the medical field is increasing. But in developing countries, there is a lot to do. The incidence and mortality rate of cervical cancer keep increasing in a low-resource setting. This is due to low screening and the unavailability of medical equipment used for diagnosis and treatment. In addition, laboratories and the patient-professional ratio are also very low in developing countries. In Ethiopia, there are 13 histopathology laboratories available with less than one pathologist per 1 million populations. Of this, 7 laboratories (53.8%) are located in Addis Abeba and the rest 6 (46.2%) are spread across other regional states with high populations resides. In some regional states like Afar, Somalia, Gambella, Benishangul-Gumuz, Harari, and Dire Dawa city administration the histopathology laboratory service is unavailable [7]. This leads to a high burden on the pathologists working in these hospitals. Additionally, a limited number of pathologists and the unavailability of screening device allows low access to routine screening. Different studies conducted recently to improve screening and diagnosis methods of cervical cancer using computer-aided diagnosis. To our best of knowledge, they are focused on constructing models of cervix type and cervical cancer classification independently with promising classification accuracy. This study is proposed to assist professionals for both cervix type and cervical cancer classification by constructing models and deploying them to the user interface system.

1.3 Statement of the problem

According to the World Health Organization, there were about 570,000 new cases of invasive cervical cancer diagnosed in 2018, representing almost 7% of all female cancers [3]. In Ethiopia, for the year 2018, the incidence and mortality rate of cervical cancer was reported 26.4 and 18.4/100,000 women [8]. Additionally, the coverage of cervical cancer screening is only 1% [9]. Currently, cervical cancer screening and diagnosis are made by using manual methods. However, developing countries suffer a lot due to the unavailability of medical resources including medical practitioners to ensure early screening and diagnosis to decide the appropriate treatment. Visual inspection with acetic acid has become the norm for cervical cancer screening, whereas,

histopathology diagnosis is used for the final decision. The type of cervix was determined manually for treatment planning. However, the results of the manual methods are prone to human error that depends on skill and experience of professionals, tedious and time consuming.

Recently, computer-aided diagnosis techniques proposed to automate and increase the accuracy of cervix type and cervical cancer classification. For cervix type classification some works applied transfer learning alone [34–36] without extracting ROI and achieved low accuracy compared with a two-way solution that includes ROI extraction before classification [12, 37–42]. For cervical cancer classification, most of the works previously done were focused on subclass classification [43–46, 48,49] and [47] done on major class classification.

However, most of the proposed techniques for cervical cancer classification in the literature are limited to specific class classification, magnification power-dependent, and used a limited amount of data. While, for cervix type classification the proposed methods used were imperfect, computationally complex, and did not perform real-time segmentation. Hence, the aim of this study is constructing a predictive model for classifying cervix types and diagnosing cervical cancer.

1.4 Research Questions

1. How could the region of interest be extracted from the cervix image and classified?
2. Which deep learning technique is suitable for developing a cervical histopathology image classifier model?
3. Which deep learning technique is best performing to construct an optimal model for cervix type colposcopy image classification?
4. To what extent do the proposed cervix type and cervical cancer classification model work?

1.5 Objective of Study

1.5.1 General Objective

The general objective of this study is to develop an optimal cervix type and cervical cancer classification model using the deep learning technique.

1.5.2 Specific Objectives

The specific objectives of this research are the following:

- To extract the region of interest and classify extracted image to cervix types
- To classify cervical cancer from a histopathology image
- To compare pre-trained models and select the best model for cervix type and cervical cancer classification
- To develop a user interface for classifying cervix type and cervical cancer classification
- To evaluate the performance of the proposed prototype

1.6 Significance of Study

Cervical cancer is still a burden for developing countries since the incidence and mortality rate is increasing. This is due to low cervical cancer screening coverage, the limited number of professionals, inconsistency, time-consuming, and inefficient diagnosis methods.

This study enables professionals to easily classify cervix type to decide treatment options for cervical cancer and assist them from which region the biopsy should be taken. Then the biopsy was taken from the region investigated under microscopy and the microscopic image classification was also included in this study. Therefore, this study assists professionals in screening and diagnosis cervical cancer by reducing time, subjectivity, and work burden.

It enables the future researcher to work on a full cervical cancer screening and diagnosis system using the dataset used in this research. In addition, it paves the way to work on further subclass classification of cervical cancer that is important to know the prognosis of the disease.

1.7 Scope and limitation of the study

This study is limited to classifying cervix type based on transformation zone location as type 1, type 2, and type 3 and classifying cervical histopathology images as normal, Precancer, Squamous cell carcinoma, and adenocarcinoma. Preprocessing is applied for both classification tasks, such as resizing and rescaling. In addition, for histopathology images histogram matching was applied to reduce color variation among the image. Besides, for colposcopy images ROI is extracted before

being fed into the classification model to remove irrelevant features. The input image to the model is classified independently for both tasks following their respective preprocessing step.

The cervical cancer classification was done by collecting local data from St. Paul hospital and Jimma University Medical Center (JUMC) from August 17 to October 6, 2020, to make cervical cancer classification more reliable. The dataset collected for cervix type was collected from Tercha general hospital from May 19 to July 10, 2020. However, for some classes, such as normal and precancerous classes there is a limited number of image slides since the patients did not come hospital in an early stage of the disease. In addition to that, the local data collection was limited to some hospitals only due to time and financial limitations.

CHAPTER TWO

LITERATURE REVIEW AND RELATED WORKS

2.1 Cervix

Cervix is part of the female reproductive organ structurally found at the lower fibromuscular portion of the uterus as shown in Figure 2.1. It is cylindrical or cone-like in shape and measures 3 to 4 cm long, and 2.5 cm in diameter. Which is supported by the cardinal and uterosacral ligaments that stretch between the lateral and posterior parts of the cervix and the walls of the bony pelvis. The lower half of the cervix, called the portio vaginalis, protrudes into the vagina through its anterior wall, and the upper half remains above the vagina[10].
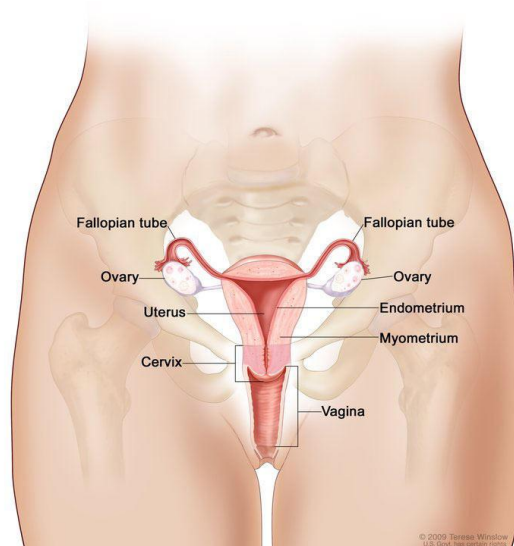
Figure 2.1: Anatomy of the female reproductive organ [11]

The cervix contains two kinds of cells: rectangular columnar cells and flat scale-like squamous cells. Transformation zone (TZ) is an area where columnar cells are constantly changing into squamous cells, and the most commonplace on the cervix for abnormal cells to develop. The location of the transformation zone varies among women. In teenage girls, the transformation zone is on the immature cervix's outer surface and is more susceptible to infection than in adult women. In older women, the transformation zone may be higher in the cervical canal [10].

The transformation zone (TZ) is termed normal when it is composed of immature and/or mature squamous metaplasia along with intervening areas or islands of columnar epithelium, with no signs of cervical carcinogenesis. While it's termed as an abnormal or Atypical Transformation Zone when proof of cervical carcinogenesis like dysplastic change is determined in the region, Identifying the transformation zone has great importance in colposcopy, as almost all manifestations of cervical carcinogenesis occur in this zone [10]. The anatomy of the cervix is shown below in Figure 2.2.
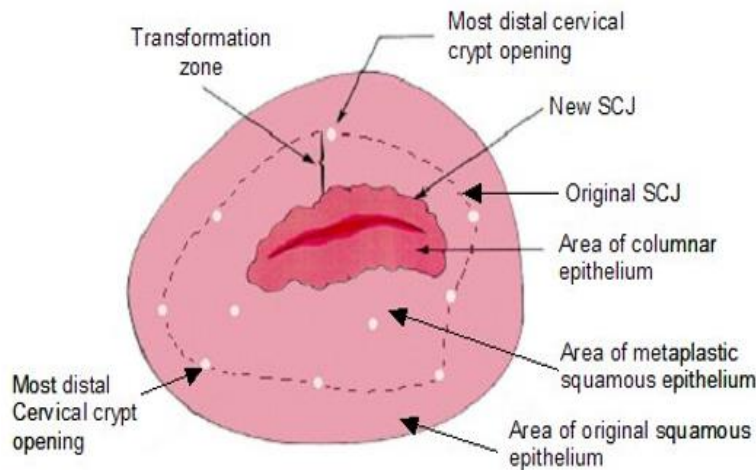
Figure 2.2: Anatomy of the cervix [12]

2.2 Cervix Type

There are three different cervix types based on the transformation zone locations [13].

- **Type 1:** is completely ectocervical, fully visible, and could be small or large.
- **Type 2:** has an endocervical component, fully visible, and may have an ectocervical component which may be small or large.
- **Type 3:** has an endocervical component, is not fully visible and may have an ectocervical component which may be small or large. The cervix type after application of acetic acid can make the landmark of TZ as shown in Figure 2.3 below.

Furthermore, most cervical cancers begin in the cells in the transformation zone. Cervix type 2 and 3 might include hidden lesions and require different treatments. Additionally, the transformation zone is directly related to excision treatment types, which determines the length of the canal

excised. Identifying the transformation zone is very important to decide treatment type since type 1 excision for type 1 TZ, type 2 excision for type 2 TZ, and type 3 excision for type 3 TZ [14].
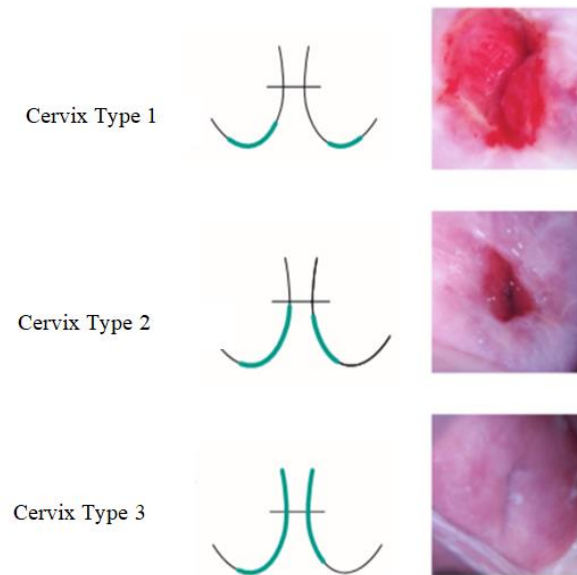
Figure 2.3: Image cervix type after application of acetic acid [13]

2.3 Cervical Cancer

Cervical cancer is the uncontrolled growth of cells in the region of the cervix which is caused by Human Papillomavirus. It initially affects the cells in the cervical region and then spreads over the entire region of the cervix [14,15]. The microscopic image of cervical cancer and other cancer have common characteristics. For example, compared with normal tissue, abnormal tissues are characterized by cell polarity disorder, uneven cell nucleus, and irregular histologic structure. However, each cancer has its organ specificity and cervical cancer has unique properties that are different from other cancers. For instance, cervical tissue is covered with squamous epithelium and its tissue and cells have characteristics of partial epithelium cells. Figure 2.4 demonstrates the gradual change of the cervical cells from normal to cancerous.
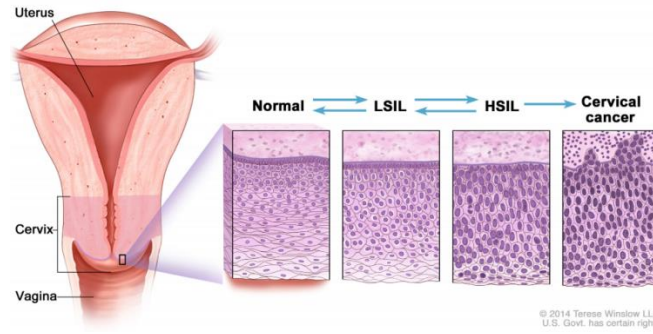
Figure 2.4: Histopathological cervical cell changes [11]

2.3.1 Pre-cancers of the cervix

The cells in the cervix at the transformation zone develop gradually to abnormal cells and develop to cervical cancer over time. The change of cell was confirmed in the lab-based on how much of the cervix tissue looks abnormal. If the case is left untreated in the precancerous stage, it will progress to cancer[9]. The pre-cancer of the cervix is a squamous cell abnormality and is currently termed a squamous intraepithelial lesion (SIL). It is classified as low-grade squamous intraepithelial lesion and high-grade squamous intraepithelial lesion.

a) Low-Grade Squamous Intraepithelial Lesion (LSIL)

The Pre-cancer of the cervix is also called Cervical Interepithelial Neoplasia (CIN1), in which much of the tissue looks normal. The histopathology examination of LSIL is shown with empty space cytoplasm in the superficial layers and part of the intermediate layer. Additionally, the undifferentiated cells are limited to the lower third of the epithelium. It is considered as least serious cervical precancer and will progress to HSIL if the lesion is not excised [16]. Figure2.5 (a) shows LISL precancer grades of cervical cancer.

b) High-Grade Squamous Intraepithelial lesion (HISL)

It is termed as CIN2/CIN3 moderate/severe dysplasia. The lesion rises in the intermediate or superficial layers of basal cells and is considered as most serious precancer. It has a greater risk of progression to invasive cancer and the histopathology image of cases are with features of enlarged nuclei and reduced cytoplasm[16]. Figure 2.5 (b) shows HISL precancer grades of cervical cancer.

(a)   (b)
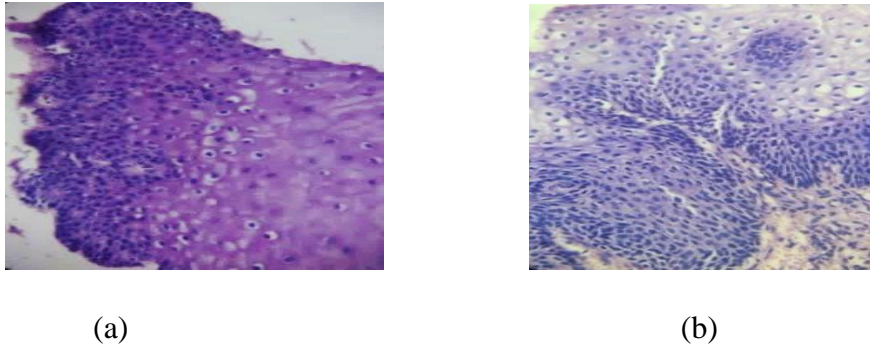
Figure 2.5: Histopathological image of squamous intraepithelial lesion (a) LSIL (b) HSIL

2.3.2 Types of Cervical Cancer Cells

There are 2 main types of cervical cancer: Squamous cell cancer and Adenocarcinoma [17].

a) Squamous cell carcinoma

Squamous cell cancer develops from the squamous cell at TZ. It is the most common cancer which accounts for 70 to 80% of cervical cancer. Microscopically, it appears as infiltrating networks of a band of neoplastic cells with intervening stroma, varying in growth pattern, cell type, and degree of differentiation. It is subdivided as keratinizing and non-keratinizing squamous cell carcinoma[17]. Figure 2.6 shows a histopathology image of squamous cell carcinoma.
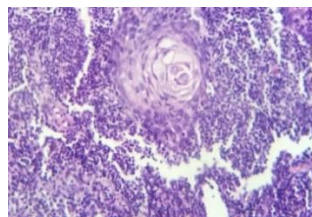
Figure 2.6: Histopathological image of squamous cell carcinoma

b) Adenocarcinoma

Adenocarcinoma is a type of cancer that starts in the gland cells where mucus is produced. The cervix has glandular cells scattered along the inside of the passage that runs from the cervix to the womb (the endocervical canal). It is less common than squamous cell cancer but has become more common in recent years. Additionally, more than 10 in every 100 cervical cancers (more than

10%) are registered. Which is treated in the same way as squamous cell cancer of the cervix [17]. Figure:2.7 below shows a histopathological image of adenocarcinoma.
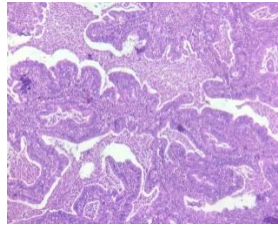
Figure 2.7: Histopathological image of adenocarcinoma

The other rare type of cervical cancer is a small cell cancer of the cervix which is diagnosed 3 in 100 women. Small cell cancers tend to grow quickly and are treated in a different way to the more common types of cervical cancer. This type happens very rarely, for example, lymphomas and sarcomas. They are treated in a different way to aforementioned cervical cancer [17].

2.4 Cervical Cancer Screening and Diagnosis Method

2.4.1 Cervical Cancer Screening

Cervical screening is used to find a change in the cells of the cervix that could lead to cancer. Clinically different procedure is undertaken to screen cervical cancer. The most common cervical cancer screening is Visual Inspection with Acetic Acid (VIA), Papanicolaou test (Pap), and Human Papillomavirus Test (HPV).

a) **Visual Inspection with Acetic Acid (VIA)**

The method involves washing the cervix with 3% to 5% acetic acid for one minute and observing the cervix with the naked eye. If white areas (acetowhite region) are seen near the transformation zone the test is considered positive for precancerous lesion or early invasive cancer. It does not require laboratory tests or intensive staff training and it is a common testing method in developing countries. In addition, the result is immediately available, allowing treatment in a single visit and thus reducing patient follow-up. The sensitivity is as good as a pap smear or better than a pap test but it's subjective and supervision is needed for quality control of the method. In addition, it does

not work in post-menopausal women because the transformation zone recedes into the cervical canal and cannot be seen with VIA[18].

b) Papanicolaou test (Pap)

It is a fast and simple screening method in which the cells are removed from the cervix and vagina examined under the microscope. This is common in developed countries since it requires infrastructure and extensively trained personnel that are lacking in low-resource settings. It lacks sensitivity and also cannot detect many cervical abnormalities. And making repeated screening is necessary due to its high rate of false-negative by single cytologic screening and it does not conform to the final stage cancer[9,16]. The procedure is shown in figure 2.8.
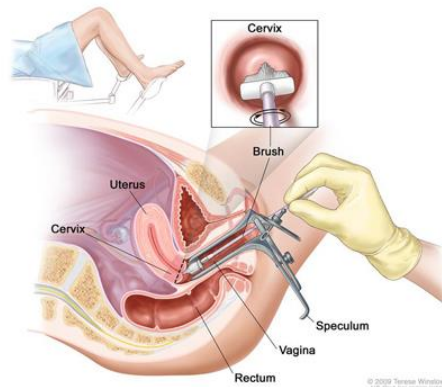
Figure 2.8: A pap test for cervical cancer[19]

c) Human Papillomavirus Test (HPV)

This test also uses cells taken from the cervix and vagina using a small brush or swab, and the specimen is sent to the laboratory for further investigation. It detects DNA from cancer-causing HPV types in a vaginal or cervical smear. It is not subjective and valuable in detecting high-grade precancerous lesions in women older than 30. Additionally, it's more sensitive than the two aforementioned methods but unaffordable for low resource areas. Additionally, it requires laboratory facilities, special equipment, and trained personnel [19].

2.4.2 Cervical Cancer Diagnosis method

a) Histopathologic test

The histopathology test can make a definite diagnosis and it is the gold standard for cervical cancer diagnosis. The aforementioned screening methods can suggest the presence of cancer. The test is performed by examining a thin section of tissue under a microscope. If the lesion is small, the doctor may remove all of it during the biopsy. The laboratory procedures are followed to prepare histology slides for examination under a microscope [20].

Step 1. Tissue fixation: this is done to prevent autolysis and putrefaction. It can be achieved by immersion of tissue in a fixative solution. The 10% neutral buffered formalin is used as a solution which allows the tissue to become adequately fixed within24-48 hours.

Step 2. Trimming: following tissue fixation transferring the specimen to cassettes using a scalpel to enable to reach a sample size that is compatible with subsequent procedures such as embedding and sectioning.

Step 3. Tissue processing: this step consists of dehydration of tissue in concentrated alcohol to remove water and formalin from tissue. Then an organic solvent such as xylene is used to remove alcohol and enable infiltration with paraffin wax.

Step 4. Embedding: after removing infiltration tissue samples from the cassette, the samples are placed in a metal base mold. The mold is then filled with molten paraffin wax and placed on a cooling surface to create a block. In this manner, solidified block provides a support matrix that allows very thin sectioning.

Step 5. Sectioning: this step includes a subsequent step in which the wax is removed from the surface to expose the tissue and the blocks are then chilled on a refrigerated plate for 10 minutes. Moreover, a microtome is aligned to tissue sections of the block. Sections are cut as a ribbon and carefully transferred to a warm water bath. Here they are allowed to float on the surface and can then be scooped up on to slide placed under the water level. Finally, slides are labeled and allowed to dry at 37°C which melt the paraffin wax and leaves the tissue section intact.

Step 6. Staining: the unstained cells are transparent and appear almost colorless. The histochemical stains such as hematoxylin and eosin are used to provide contrast to the tissue section, making tissue structure visible and easier to examine. After staining, a coverslip is mounted on the slide to help protect the specimen.
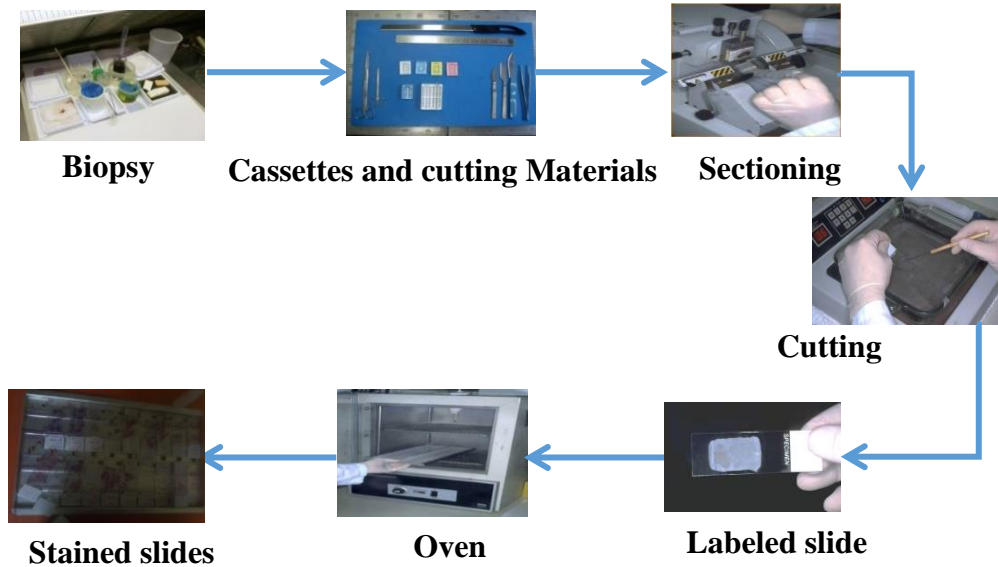
Figure 2.9: Histopathology slide preparation in the laboratory

2.4.3 Stages of Cervical Cancer treatment

If the biopsy test shows the presence of cervical cancer, the doctor will refer the woman to a gynecologic oncologist. The specialist may suggest additional tests to see if cancer has spread beyond the cervix. Cervical cancer is staged clinically with pelvic examination, under anesthesia with bladder cystoscopy and colposcopy, in combination with imaging tests including X-ray, Computed tomography (CT), Magnetic resonance image (MRI), and Positron emission tomography (PET) [21].

2.5 Computer-Aided diagnosis of cervix type and cervical cancer classification

Recently, computer-aided diagnosis (CAD) was applied on colposcopy and histopathology image classifications to automate the system. It helps medical professionals in a wide range of medical imaging tasks like disease detection, localization of cervix, cancer grading, treatment planning. Besides, researchers have been working to reduce inter-and intra- reader variability and workload

for professionals. The CAD system is implemented using machine learning, deep learning techniques, transfer learning, and hybrid techniques.

Machine learning is a subset of artificial intelligence, which allows a machine to learn meaningful patterns from data directly with handcrafted features. Often this technique is applied for histopathology images and cervicography images. It is essential to increase the objectivity of screening and diagnosis of cervical cancer. However, the strength of learning depends on the handcrafted feature. In addition, it cannot learn or handle new data due to the problem of gradient converging to zero and failing to the local maxima [22].

Deep learning is the result of the growth of machine learning that uses non-hand-crafted features to solve problems. It enables the creation of complex networks with deep layers to extract features from input images [23]. The last layer or classifier of the deep learning model may result in overfitting if less data is fed to the model. This leads to unnecessary usage of resources and complex computational power[24].

Hybrid Learning is a fusion of deep and machine learning that is done by taking advantage of both to efficiently solve the problem. This was done using CNN to extract features from unstructured data and the classical machine learning approach used as a classifier. This was experimentally proven to be a good approach to boost performance [23]. However, there is no standard approach to implement this method it depends on the researcher.

2.6 Deep learning

Deep learning is a subset of machine learning, which is a type of representation learning in which no feature selection is used. Instead, the algorithm learns which features are best for classifying the data on its own. Artificial neural networks form the basis for deep learning. Its structure is similar to the human brain with a functional group of neural networks. The neural network consists of three layers input, hidden, and output layer [24]. It was applied in screening and diagnosis of cervical cancer using three different methods including training from scratch, transfer learning, and modifying the architecture to reduce the number of parameters existing in the network. However, this technique needs enough representation of data [25]. The architecture of the neural network is shown in Figure 2.10. Besides, deep learning techniques can be used for performing classification and object detection tasks.
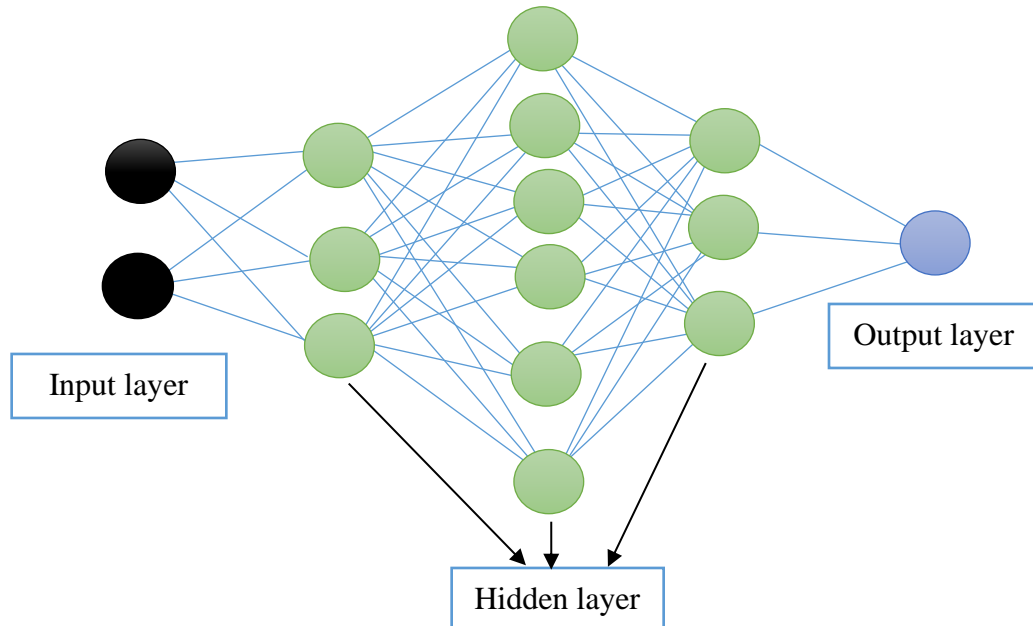
Figure 2.10: Architecture of neural network [24]

2.6.1 Object Detection

Object detection is an important component of computer vision. It consists of object localization and image classification. Object localization is used to localize where the object is found in the image. The first deep learning object localization algorithm has three modules known to be R-CNN. It detects an object accurately using rich features and segments semantically. Then the fast RCNN was introduced by Ren *et al.*[26]which merges the region proposal and CNN modules to reduce the overall computational complexity. Recently, real-time object detection models were introduced as YOLO which operate 45 frames per second in the baseline YOLO model [27]. Moreover, the object detection algorithms are applicable to detect objects in an image in lots of studies such as medical and space imaging [28].

Image classification is a key technique of a computer-aided diagnosis system that classifies images based on a labeled dataset for specific classes. Convolutional neural network (ConvNets) is a class of deep, feed-forward artificial neural networks that are applied to analyzing visual imagery. The most commonly used ConvNets in ImageNet includes VGGNet[29], MobileNet[30],ResNet[31], and EfficientNet[32] .

A convolutional neural network (CNN) is a deep artificial neural network made of neurons with learnable weights and biases. It is composed of convolution, pooling, activation, and a fully

connected layer. The convolution layer is a central part of CNN used for feature extraction from an input image and the image is convolved with a filter. The down sampling operation or reducing the size of an input image is done by the pooling layer. The activation layer is used to increase nonlinearity in output to improve the expression ability of the network. A fully connected layer is at the end of CNN used to reduce feature maps and as a classifier [33].

2.7. Related works

In this study, we present a review of works done towards automating cervix type and cervical cancer classification.

2.7.1 Related works on automatic Cervix type classification

Over the years, researchers have been working to automate cervix type classifications that are important to decide optimal treatment options. In most studies, two-stage solutions were proposed to solve the classification problem of cervix extracting the ROI and then classifying images as cervix type 1, type 2, and type 3. Moreover, the ROI extraction from raw images is used to remove irrelevant information like pubic hair, clinical instrument (speculum), hands of professionals, and Intrauterine Device (IUD) that are intended to mislead the CAD system. On the other hand, some of the studies were classified the image by fine-tuning pre-trained models. All related works used colposcopy images to accomplish cervix type classification.

Different literature has proposed cervix type classification using the transfer learning approach[34–36]. A pre-trained Inception v3 [34] achieved a better result of cervix type 2with a precision of 0.61, recall of 0.71, and f1 score of 0.65. Similarly, study [35]selected custom model 14 layer CNN and pre-trained models Inceptionv3 and ResNet 50 for comparison. The selected models yielded unpromising results. Finally, they were used ResNet 50 as a feature extractor and Inception-like residual network as a classifier. The result shows that the method performs better from selected CNN with an accuracy of 69.93%. Likewise, in[36]selected pre-trained models ResNet 34 and Inceptionv3; compared the model with and without applying augmentation techniques. The result shows that ResNet 34 performs better with applying augmentation techniques on image and yielded accuracy of 72% on the test set. The above result achieved in each study presents that the generalization ability of the model is low.

On the other hand, the two-stage solution has been proposed in [12, 37–42]which performs classification of cervix type following ROI extraction. In a study [37], a color assembly method was proposed for ROI extraction and used to learn artificial neural networks (ANN). The comparison was done using pre-trained model xception and custom model 7-layer CNN with and without extracting ROI. The CNN trained with extracted ROI yielded better accuracy of 64% and 0.79 loss. Despite the simplicity of the ROI extraction and classification model, the generalization ability of the model is low. Similarly, the simple ROI extraction method is proposed [38] using tiny ResNet (similar to ResNet 18). Then the cropped images were trained from scratch using popular models AlexNet, VGG16, Inceptionv3, and ResNet.Inceptionv3 performs best compared with selected models with 74% of accuracy.

In the study [39], 1480 images were used for classification that is base dataset only. The irrelevant features are removed using U-Net.VGG16/19 pre-trained models were used as a feature extractor from the segmented region and features are classified by three fully connected convolutional networks with 0.2 dropout. VGG16 achieved better validation accuracy of 62%. Likewise, segmentation-based preprocessing techniques using U-Net and 32 layers convolutional neural network is used as a classifier. A total of 1480 images was used to solve the classification problem. The method achieved DICE coefficient of 78% using U-Net and yielded 70% of the accuracy of the classifier[40]. Improved accuracy of 80.1% was presented in [41] using U-Net for segmentation of ROI and Caps Net neural network. The number of images used in this study was 6692. The classifier result in low accuracy and it was due to segmentation imperfection.

Another approach used Mask RCNN for ROI extraction and Hierarchical Convolutional Mixture of Experts (HCME) classification neural network. Besides, a novel loss function was designed using cross-entropy to boost classification. The method claimed, 96.77% of accuracy and a Kappa score of 0.951 [12]. This approach outperforms the aforementioned studies however the segmentation method is computationally complex and not real-time.

In the study [42], the real-time ROI extraction was implemented using the current state of art YOLO object detection model. Intel & Mobile ODT Dataset and national cancer institute (NCI) Guanacaste Project images were used to solve the cervical cancer classification problem to automate VIA screening. The result is reported in terms of the intersection of union (IOU) 0.68

and classification AUC score is 0.82. The segmentation of the cervix region is real-time but the method focused on grading CIN using colposcopy images.

Even though transfer learning is widely applied in a lot of studies with good results, the accuracy of cervix type classification was not satisfactory[34–36].To boost the performance of classification most studies focused on removing irrelevant features using different techniques[12, 37–42]. However, the methods explored were used computationally complex, imperfect, and not real-time cervix segmentation techniques. Table 2.1 shows a summary of related works for classifying cervix type.

Table 2.1 Related works for cervix type classification

Authors	ROI detection	Classification	Performance	GAP
1.Zang et al. (2018)	U-Net	CapsNet	80.1% of tests set accuracy	- Low accuracy -Computationally complex segmentation
2.Rohan G. et al. (2019)	Mask-RCNN	HCME	Accuracy: 96.77% Sensitivity: 96.82% Specificity: 98.36%	-Computationally complex segmentation
3.M.B. Bijoy.et al. (2020)	U-Net	CNN model	Accuracy: 70%	- Low accuracy -Computationally complex segmentation
4.Mamta et al. (2020)	-	ResNet& Inception v3	Accuracy: 72%	- Low accuracy
5.Daniel A. Cruz et al. (2020)	Color model	CXNN	0.79 loss value	- Low accuracy
6.Navdeep Kaur et al. (2017)	U-Net	VGG16	Validation accuracy: 62%	- Low accuracy - Computationally complex segmentation

2.7.2 Related works on automatic cervical cancer classification

Previously, research conducted on cervical cancer histopathology image classification focused on using handcraft image feature extraction methods and deep learning techniques. Effective CAD systems will improve the diagnostic efficiency of pathologists and reduce their workload in this field. All related works were used locally collected datasets.

Different literature proposed means of grading cervical interepithelial neoplasia (CIN) [43–46]. Localized fusion-based approach proposed [43] to classify squamous epithelium into Normal, CIN1,CIN2, and CIN3 grades of CIN. In this paper, 61 digitized histopathologic Hematoxylin and Eosin-stained images were used. The classification was done by dividing the epithelium region into 10 vertical segments, extracting features from each segment, classifying each segment, and fusing the CIN grade of the whole epithelium. Voting fusion of vertical segments was done using support vector machine (SVM) and linear discriminant analysis (LDA). Additionally, the left one-out validation technique was used and achieved an accuracy of 88.5%. Similarly, [44] extended the aforementioned work by dividing each segment into three parts. In this work, CNN is used for classification with 65 digitized histology images. The proposed method achieved better results compared with previous work with exact class accuracy of 77.27% while on the same data the former study achieved 75.75%. In 2019 the same author[45] proposed a hybrid method by extending the dataset to 83 images. In this paper, the handcrafted feature and deep features are computed for each segment. The method improved 15.51% and 11.66% over the aforementioned handcrafted feature classification and deep learning alone respectively. Likewise, an epithelium detection module was proposed [46] before segmentation as in the above work using 150 whole slide images (WSI). The method achieved 85% of accuracy which outperform another method with an AUC of 94.4% but it's limited to only a whole slide image, CIN grading alone, and 20x magnification power.

In the study [47], a method to classify the histopathology image as Normal, CIN lesion, and cancer using 90 histology images based on texture and lesion feature presented. The approach used handcrafted features and yielded an accuracy of 90%.

Some researches focused on classifying the most common cervical cancer called squamous cell carcinoma(SCC)[48, 49]. A sub-class classification of SCC proposed to classify histopathology

images of cervical tissue as keratinizing, non-keratinizing, and basaloid squamous cell carcinoma. The dataset used for this study is 3,012 and was augmented to 108,432 images. AlexNet is used for training the images and the system is validated with three-fold cross-validation. The result shows 93.33% of accuracy but this method is limited to SCC class classification alone[48]. Another approach [49]introduced a method to identify squamous cell carcinoma from the whole slide image. A total of 800 Hematoxylin and eosin-stained slides of squamous cell carcinoma were used in this study. The method includes extraction of a tissue region from whole slide image (WSI) and Patches of images were classified as normal and cancerous by ensembling Inception_v3, InceptionResnet-v2, and ResNet 50. Moreover, the classified patches were transformed back to WSI, and post-processing was applied on heatmaps to classify the image as SCC or Normal. The approach achieved an AUC score of 0.9784 and an FROC score of 0.5609 using the ensemble model. The classification of SCC is clinically important and also the result is acceptable but these approaches are magnification power-dependent and also limited to SCC class only.

To automate cervical cancer diagnosis the images used in the above studies were limited, due to this they were used handcrafted features for the classification of the cervical histopathologic image. Additionally, the generalization ability of models is low, classification was limited to some classes and applicable for specific magnification power. Table 2.2 represents some related works for classifying cervical cancer.

Table 2.2: Related Work for cervical cancer classification

Author	Method	Class	Accuracy	GAP
1. H. Almubaraket al., (2017)	ConvNet	Normal, CIN1, CIN2, CIN3	77.25%	- Low accuracy - Precancer grading alone
2. AlMubaraket al., (2019)	Machine and deep learning (hybrid)	Normal,CIN 1,CIN 2 and CIN3	80.72%	- Low accuracy - Precancer grading alone
3. Peng Guo, et al., (2015)	Machine learning	Normal,CIN 1,CIN 2 and CIN3	88.5%.	- Low accuracy

				- Precancer grading alone
4. Miao et al., (2018)	DCNN	keratinizing s, non-keratinizing, and basaloid squamous cell carcinoma	93.33%	-Squamous cell carcinoma
5. Sudhir et al., (2020)	DCNN	Normal, CIN1, CIN2 and CIN3	85%	- Low accuracy - Precancer grading
6. L.Weiet al., (2017)	Texture feature extraction and SVM classifier	Normal, CIN 1, CIN 2, CIN3, and cancer	90%	- Cancer class is not identified

2.7.3. Research gap

Even though the aforementioned works achieved promising results, still improvements were needed to fill the gaps on related works. These gaps include, for example, to automate cervical cancer diagnosis the images used in the above studies were limited, due to this they were used handcrafted features for the classification of the cervical histopathologic image. Additionally, the generalization ability of models is low, classification was limited to some classes and applicable for specific magnification power. To our best of knowledge, there is no work previously done to make both classification tasks in one. The two important cervical cancer screening and diagnosis systems are embedded separately which can have great importance to consume time for professionals if it is in one.

CHAPTER THREE

RESEARCH METHODOLOGY

3.1 Overview

Research methodology is a systematic design of study that ensures valid and reliable results to address the objective of the study. This chapter focuses on the methodology followed in this study starting from data collection to constructed model deployment.

3.2 Research design

In this study, experimental research methodology is followed. It enables the researcher to observe the effect of an independent variable on the dependent variable. To develop cervix type and cervical cancer classification two unrelated images trained independently and come together at the end in the same user interface. The independent variables are images used and dependent variable includes the performance metrics. Figure 3.1 shows the workflow of the proposed study.

The dataset used in this study was colposcopy and colpophone images for cervix type classification and histopathologic images for cervical cancer classification. Then the datasets were split into train and test sets. Prior to classification region of interest was detected using lightweight MobileNet Training of pretrained model Vggnet16, ResNet50, ensemble of Resnet and MobileNetv2 and EffecientNetb0 done and compared to have best performing model. The dataset was split into training (used to construct model) and test set (used to evaluate the performance of the constructed model). For cervical cancer classification, histopathology images acquired locally were used.
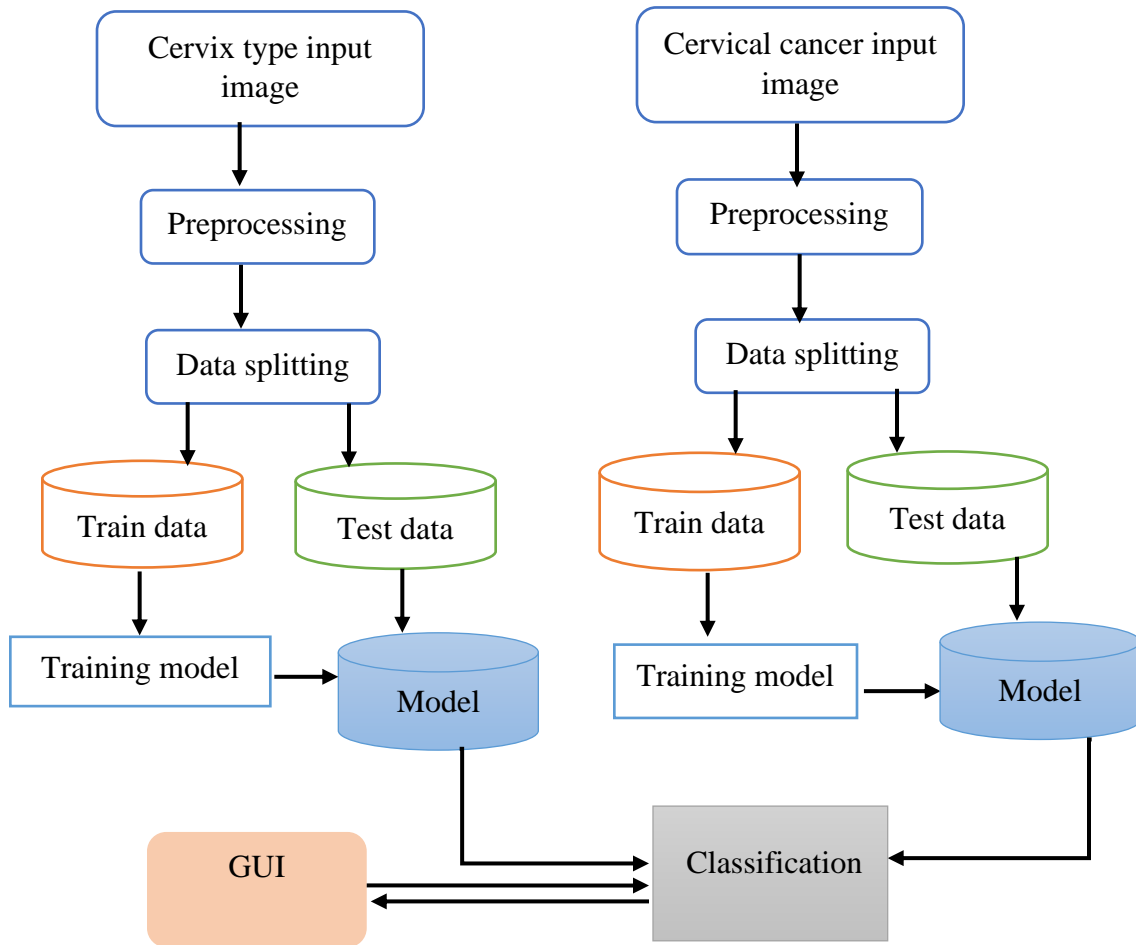
Figure 3.1: Block diagram of proposed work

3.3 Data collection and preparation

3.3.1 Study area

The histopathology dataset was collected from two referral hospitals that are Jimma University Medical Center (JUMC) and St. Paul hospital for cervical cancer classification. JUMC is well known specialized referral hospital located in the south west region of Ethiopia. The pathology department of JUMC is established in 1985 G.C. Likewise, St. Paul hospital is also a well-known specialized referral hospital located in Addis Abeba. All the dataset was locally collected for cervical cancer classification since there is no online dataset. For cervix type classification, the

local dataset was collected from Tercha General Hospital which is located in the Dawuro Zone southwest region of Ethiopia. The public dataset was collected from Kaggle repository.

3.3.2 Study population and sampling

The non-probability sampling approach was used for study during data collection. It is a type of sample selection method that is nonrandom and mostly used for exploratory study. In addition, a purposive type of non-probability sampling was used based on the purpose of the study.

The study population for cervix type non-probability sampling in which married women who visited gynecologic OPD (outpatient department) and screened for cervical cancer and the image was acquired with aid of smart phone camera. The class label of the dataset was done by a gynecologist, ISEO, and medical doctor. Likewise, the study population for cervical cancer classification was non-probability sampling through which labeled slides in the pathology laboratory were used using mobile cameras and amscope. In addition, all labeled slides in St. Paul hospitals of 2020 were acquired using a digital scanner. Finally, the class of cancer for the image was confirmed by pathologists.

3.3.3 Data collection

The dataset used for cervix type identification was collected from local health facilities and a publicly available Kaggle dataset. The image was captured for local data during screening and a public dataset was downloaded. Then, the dataset was annotated for preparing to extract ROI. While cervical cancer classification was collected from JUMC and St. Paul hospital. The task performed here was acquiring images and labeling their respective classes.

3.3.4 Preprocessing

The proposed study has two parts one is cervix type classification and cervical cancer classification. Since the colposcopy and histopathology images used for both tasks are not related, they were preprocessed separately. For colposcopy image region of interest was extracted to increase the accuracy of classification by removing irrelevant regions consisting of non-important features. Then the images were resized to have low, uniform dimensions. The whole dataset was split as 90% of train and 10% test and then for train data, 10-fold cross-validation technique was

applied with 90% 0f training and 10% validation in each fold. The histopathology images varied in color due to different cameras used to acquire images and chemicals used for slide preparation. Therefore, to reduce color variation among images histogram matching algorithm was applied to the image. Afterward, the images were resized and normalized. The dataset was split into 80% of training, 10% of validation, and 10% of test for classification.

3.4 Implementation tool selection for prototype development

Cervix type classification is important to decide from which region the biopsy should be taken since miss identification leads to misdiagnosis and treatment. The proposed method consists of the cervix-type model that is constructed by comparing different models and steps of training. It includes training pretrained models with and without extracting ROI using a lightweight Mobilenetv2yolov3 object detection model. Following this, the cervical cancer image model was constructed by training state of art model effecientnetb0 and the best performing model was selected by comparing models trained by applying histogram matched and without applying it.

The final model was selected to be the best model and used in user graphical user interface after measuring the performance of models. The graphical user interface was developed using tkinter and enable the user to classify cervix type and cervical cancer image classification. The user interface was developed through continuous guidance and feedback that has been collected from clinical collaborators and advisors. The user of a developed user interface is professionals working in gynecology and pathology departments.

3.5 Performance evaluation

Test dataset was used to measure the performance of constructed models for both classification tasks. The performance measuring metrics used in this study includes confusion matrix, accuracy, precision, recall, F1 score, kappa score, and receiver operating characteristic. Confusion matrics is a tabular classification report consisting of the actual and predicted number of the test image. It is basis for the rest of metrics through which accuracy, recall, Fl score, and Kappa score are calculated. The area under receiver operating characteristic (ROC) curve is a graphical tool to evaluate the performance of the model at the varied threshold value.

3.6 Materials used in this study

Python is a modular, open-source, and interpreted general-purpose programming language which is popular and easy to use. It is used with different packages that support overall computation which can make the system perform better. In this thesis work python, 3.8.8 is used with online free GPU in google Colab and Kaggle. In addition, amscope 3.7 software is used for the visualization of images by mounting the camera on the microscope. A digital slide scanner is also used to acquire the whole slide images. To build GUI jupyter notebook is used to edit python code with anaconda. Table 3.1 represents the materials used in the study.

Table 3.1: Materials used in the study

Software	Hardware
- Python 3.8.8 - MATLAB 2019 - Amscope 3.7 - Anaconda 4.10.3	- Personal Computer HP EliteBook 840 G3, CORE i5 - Amscope camera - Microscope OPTIKA - Digital slide scanner - Smartphone with a camera of Pixel 41 and 18

CHAPTER FOUR

RESEARCH METHODS AND ALGORITHM

4.1 Overview of cervix type classification

Figure 4.1 demonstrates steps followed to develop the final cervix type classification system. Different methods were trained using pretrained model with prepared image and train-test split validation techniques. Then looking over the result pretrained model was trained using 10-fold cross-validation technique with similar data as the former. Finally, to have good results the models were trained with ROI extracted images and 10-fold cross-validation techniques. The collected data was prepared for processing by removing non-cervix images/parts manually. Then the data prepared is fed to the object detection network for cervix localizations (ROI extraction). The ROI extracted images were finally fed to the cervix type classification model for training and validation.
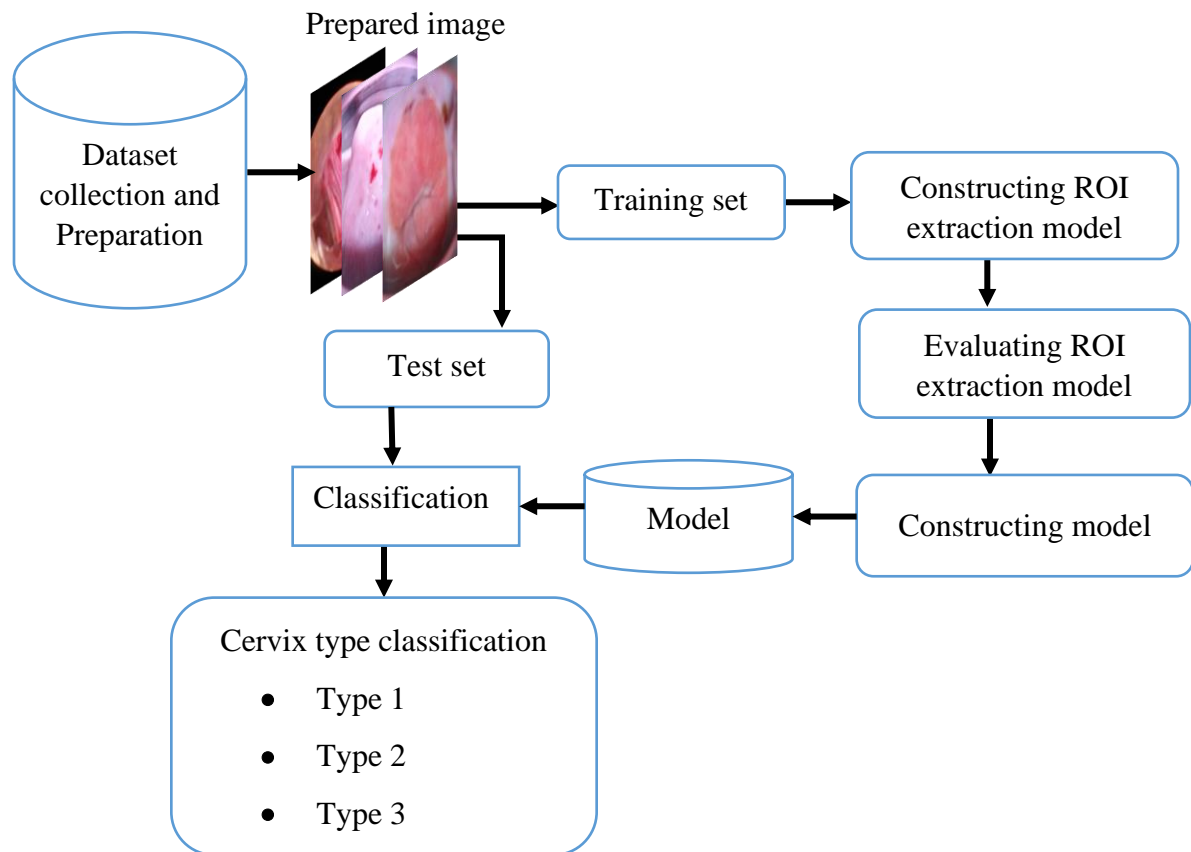

Figure 4.1: Architecture of cervix type classification

4.1.1 Datasets

The dataset used in this study was collected from a publicly available online dataset and locally acquired from health facilities. The public dataset was collected from the Intel & Mobile ODT in Kaggle challenge which contains 8,727 image-level labeled images. The size of the image varies greatly from 480x640 to as large as 3096x4128. Every image of the cervix dataset is of different sizes and captured using colposcopies. It was categorized into one of the three classes namely type 1, type 2, and type 3. All the images are color images and are in jpg format. The images present in the dataset are very high-resolution images [50].

In this study, the dataset was manually cleaned since it contains some unwanted images like the image of person, hands, and also duplicates as shown in Figure 4.2. In addition, class balance and cleaning dataset was done manually using additional datasets provided in Kaggle as well as the local dataset.
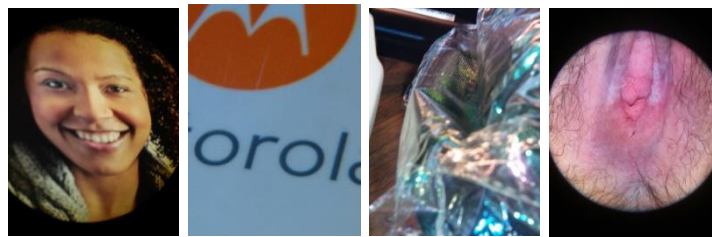
Figure 4.2: Some removed images during the data preparation step

The local dataset was collected from Tercha General Hospital, which is located in the Dawuro Zone southwest region of Ethiopia. A total of 133 Data was collected using all married women who were sent from Gyn OPD after counseling for cervical cancer screening with resolution varies from 728x945 to3120x4160. Then the clinical procedure was followed by using a speculum to access the cervix for 5% acetic acid application on the cervix, after one-minute cervix image was captured with the assistance of professionals using 13MP and 18MP Tecno mobile cameras. The collected images were then labeled in image level by three expert gynecologists, an integrated emergency surgical officer (IESO), and senior medical doctors.

The labeled local images were merged with public data for training. The total number of images in each class is presented in Table 4.1. A total of 4005 images were used for cervix type classification.

Table 4.1: Dataset for cervix type classification

No.	Dataset	Base dataset			Additional dataset			Total
		Type 1	Type2	Type3	Type1	Type2	Type3	
1	Public	338	1047	609	954	268	656	**3872**
2	Local	41	41	51	-	-	-	**133**

4.1.2 Detection of Cervical Region

To detect a region of Cervix the deep learning object detection model was used. This study used You Only Look Once (YOLO) algorithm due to its better performance like real-time detection and its speed[51]. YOLO is a popular approach for object detection, which treats object detection as a regression problem. It is real-time object detection with 75 convolutional layers with skip connections and up sampling layers. A backbone of the network is used as a feature extractor for predicting the bounding box and classes.

4.1.3 Annotation

The purpose of using annotation is to localize ROI before feeding images to the detection network. The DarkLabel2.3 annotation tool was used to annotate images. The data format of annotation was Pascal Visual Object Class (VOC) and set to only detect one object or cervix. Figure 4.3 shows the annotation of an image with annotation software and produced xml file. Public and local datasets were annotated for detecting the ROI with assistance of clinical collaborators.

(a)   (b)
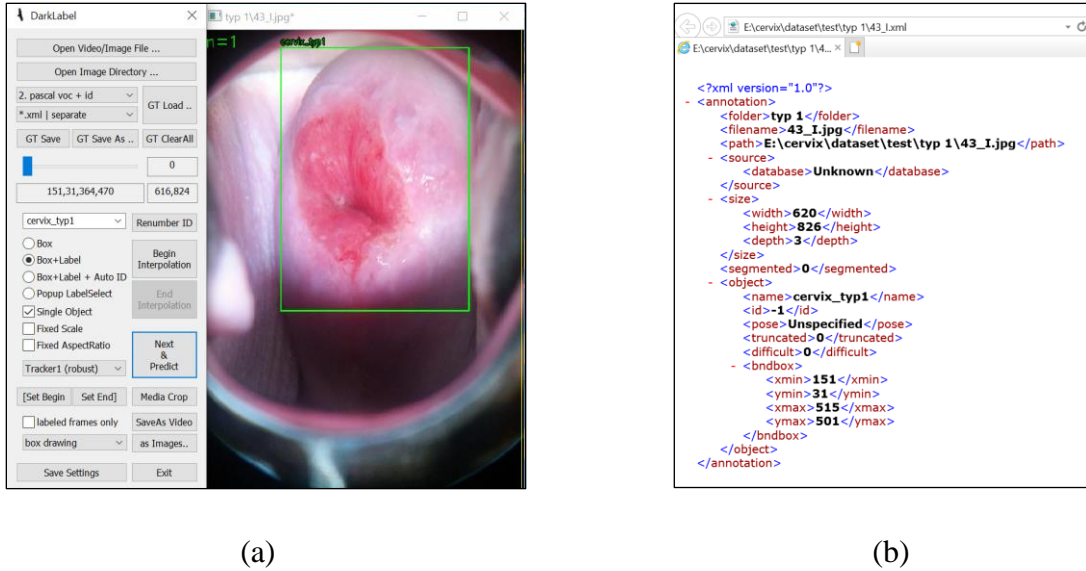
Figure 4.3: Annotation of image (a) image annotation tool (b) produced xml file

4.1.4 Localization of cervix using Mobilenetv2-Yolov3 model

4.1.4.1 YOLOv3

YOLOv3 is the third version of YOLO which predicts at three different scales. Moreover, it divides an image into grid cells, produces the cell's probabilities, and predicts boxes. The network input is any number divisible by 32 like 416x416, 608x608, and 832x832. The detection layer makes detection at feature maps of three different sizes, having strides 32,16, and 8 which are factors to down sampling image[51]. Figure 4.4 shows the architecture of YOLOv3.

Figure 4.4: Architecture of YOLOv3 [52]

4.1.4.2 MobileNet

MobileNetv2 is a lightweight 53-layer deep convolutional neural network, designed to be used in mobile applications. It is based on an inverted residual connection where the shortcut connections are between bottlenecks. It uses depthwise separable convolution followed by pointwise convolution [53].

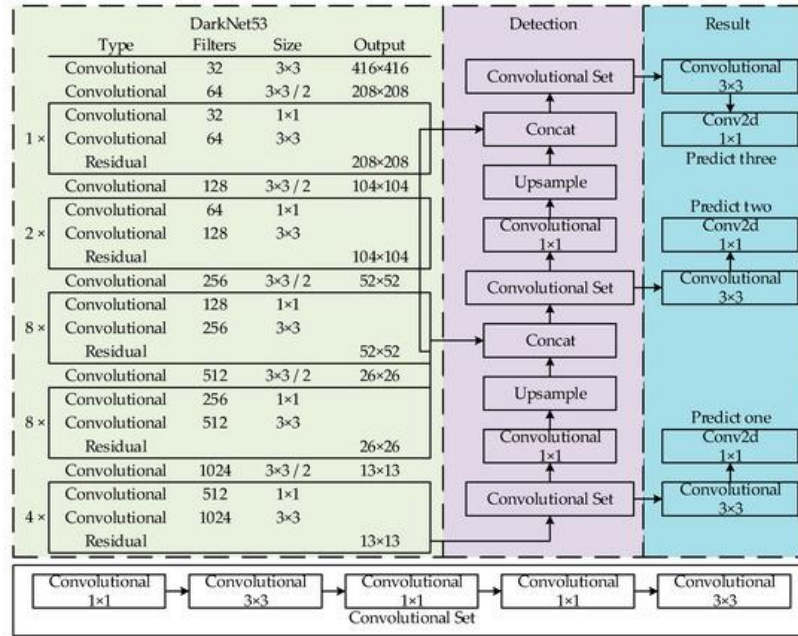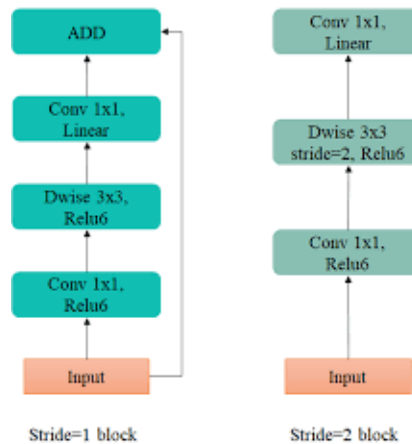Figure 4.5: The convolutional blocks in MobileNetv2 [54]

In this study, the MobileNetv2-YOLOv3 lightweight network is used for ROI extraction. In the original YOLOv3 model darknet53 is used as the backbone while in Mobilenetv2-Yolov3 mobilenetv2 is used as the backbone. The pretrained imagenet model MobileNetv2 is used to facilitate migration to the mobile terminal. Furthermore, it combines the anti-residual module with the depth-wise separable convolution. In which the depth-wise separable convolution effectively improves the computational efficiency and reduces the number of parameters. This gives the network to perform well with the lightweight character and good speed. The alpha parameter is set to be greater than 1 to control the width of the model that increases the number of filters in each layer.

A total of 4005 cervix image dataset is split into 80% of training, 10% of validation, and 10% of testing with respective annotation files. The annotation file and images were converted to TFRecord file format which stores a sequence of binary records. The input image data size was 416x416 since the region of interest is larger. Then, the input image was down sampled using strides 32,16, and 8 which resulted in 13x13, 26x26, and 52x52 of an output image. The detection network was trained with a hyperparameter adam optimizer, 0.001 as starting point of learning rate and batchsize of 8. Moreover, the function was added to crop out the region of interest using coordinate points. Then the output of detection model was generated in a respective file folder.

4.1.5 Cervix type classifications with test-train split

In this study, a pre-trained model was selected based on the number of layers, novelty, and number of parameters such as VggNet, ResNet, and EfficientNet families. The pre-trained models with a train-test split of the total dataset were used as a validation technique. Different optimization parameters were adjusted during training to check the performance of the model. Adam and stochastic gradient descent (SGD) optimizers were used with a learning rate of 0.001-0.0001. The epoch and batch size values were set to 25-100 and 32 respectively. This method paves away to the next solution by looking over the result.

4.1.5.1 VGGNet

VGGNet is a convolutional neural network architecture that focuses mainly on the effect of CNN depth on its accuracy. The input image is 224x224 RGB image and deeper than AlexNet. VGGNet

has VGG16 and VGG19 families which is 16 and 19 layers deep respectively. The architecture of VGG16 is shown in Figure 4.6 [29].

Figure 4.6: VGGNet Architecture [29]
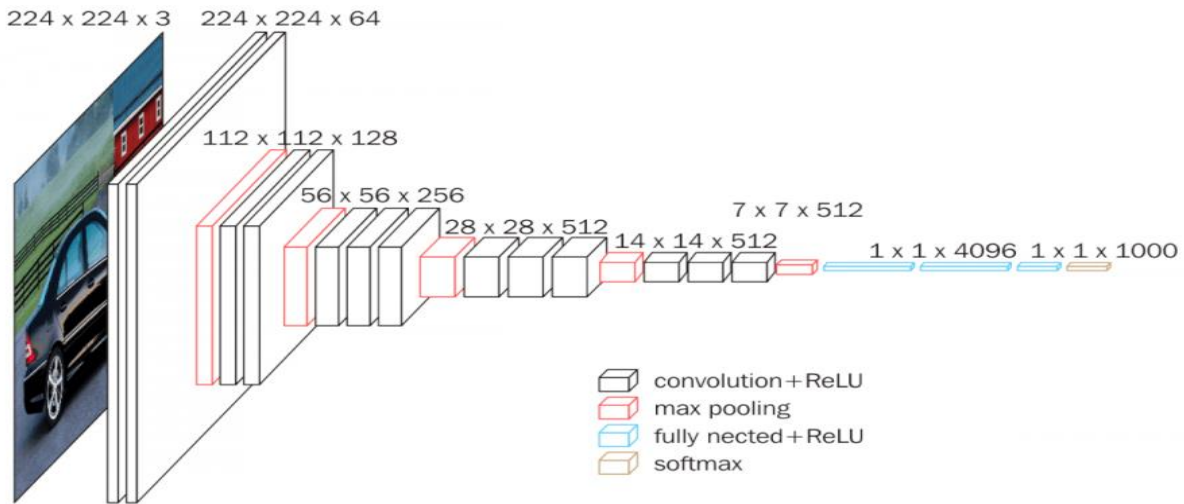
4.1.5.2 ResNet

ResNet is one of the most powerful ConvNet which achieved layers deep. It is used in many computer vision tasks. The novelty with this architecture is the concept of skip connection which overcomes the vanishing gradient problem[31]. The building block of the residual network is shown below in Figure 4.7.
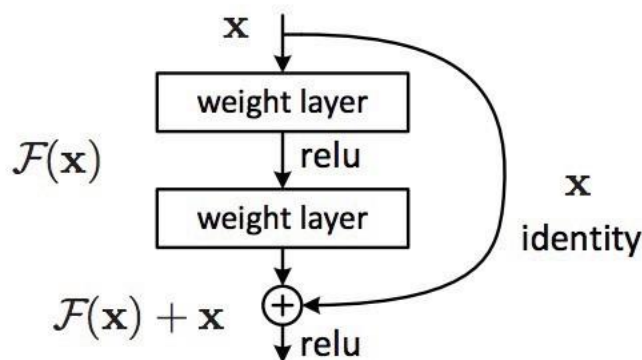

Figure 4.7: Residual learning: building block[31]

4.1.5.3 EfficientNet

EfficientNet is a state of art classification convolutional neural network due to its principled method of model scaling. This made it have better accuracy and efficiency. Before this scaling method was introduced there was no other known model scaling method. Scaling the model consists of the width, depth, and resolution wise. The method uniformly scales network depth, width, and resolution with a set of fixed scaling coefficients. Additionally, it transfers well in other datasets and achieves state-of-the-art accuracy on 5 out of 8 widely used datasets such as CIFAR-100 (91.7%) and Flowers (98.8%), while reducing parameters by up to 21x than existing ConvNets. Figure 4.8 indicates the architecture of EfficientNetB0 which is a baseline for other EfficientNet families [32].
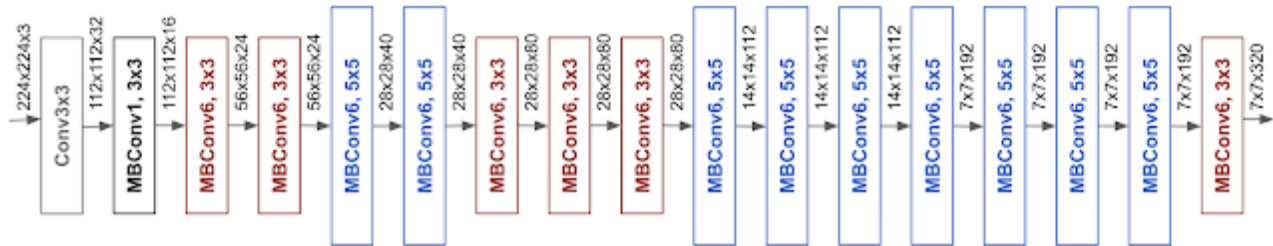

Figure 4.8: Architecture of baseline network EffecientNetB0 [32]

The compound scaling method uses the compound coefficient and consistently improves model accuracy and efficiency for scaling up existing models. Experiments were done for existing models MobileNet (+1.4 % imagenet accuracy), and ResNet (+0.7%) compared to the conventional scaling method. Taking these experiments as an advantage the base model was developed by using a multi-objective neural architecture search that optimizes both accuracy and floating-point operation. The developed families of efficientnet are B0, B1, B2, B3, B4, B5, B6, and B7 from which B0 is base model B1-B7 are scaled models from the base model. Additionally, this model can reduce the parameter size and FLOPS by an order of magnitude. EfficientNet-B0 architecture is a mobile-sized architecture having 11M trainable parameters. Unlike the conventional model scaling which scales the model in one dimension only as shown in Figure 4.9, EfficientNet did compound scaling of the model in a well-structured manner. The compound coefficient Φ is used for compound scaling with alpha, gamma, and beta as scaling multiplier[32]. The compound coefficient is user-defined and scaling parameters are shown in equations 4.1-4.3.

$$depth: d = \alpha^{\Phi}\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(4.1)}$$

$$width: \omega = \beta^{\Phi}\ldots\ldots\ldots\ldots\ldots\ldots \text{(4.2)}$$

$$resolution: r = \gamma^{\Phi}\ldots\ldots\ldots\ldots\ldots \text{(4.3)}$$

Such that $\alpha. \beta^2. \gamma^2 \approx 2$

$\alpha \geq 1, \beta \geq 1, \gamma \geq 1$ Where: α, β, γ are scaling multiplier for depth, width, and resolution
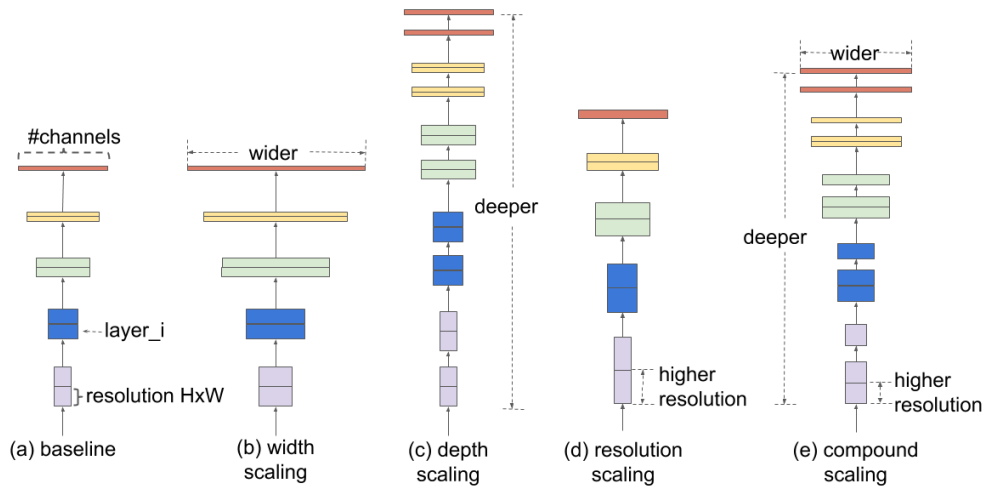
Figure 4.9: Model scaling method[32]

4.1.6 Cervix type classification with 10 fold cross-validation

The pretrained model explained above was used and compared by training it with and without ROI extracted image. The three models were compared using the same optimization parameters and methods. The k fold cross-validation was used to partition data. Moreover, the L2 regularization technique was applied in each layer to reduce overfitting. The learning rate of 0.00001 was used as a starting point. The epoch and batch size were 25 and 32 in each fold respectively. Additionally, ResNet 50 and Mobilenetv2 were ensembled and trained using a learning rate of 0.0001,32 batchsize and 50 epoch.

4.1.6.1.Cross-validation

Cross-validation is one of the satistical methods used to estimate machine learning skills. K-fold cross-validation is a method used to estimate the true performance of the machine learning model,

select the best performing model and allow parameter tuning. In this study, the k-fold cross-validation technique was used. It is selected due to its less bias compared to the hold-out method. The value of k should not be too high or too low and is mostly recommended to be 5 to 10. For instance, if k=2 it results in the highly biased model like train test split. The k parameter can give a chance to the dataset to be held out once in each fold as the validation set and k-1 as training set [55]. Additionally, it ensures that every dataset is observed during training and, enables the model to learn well. Figure 4.10shows steps followed during k fold cross-validation.
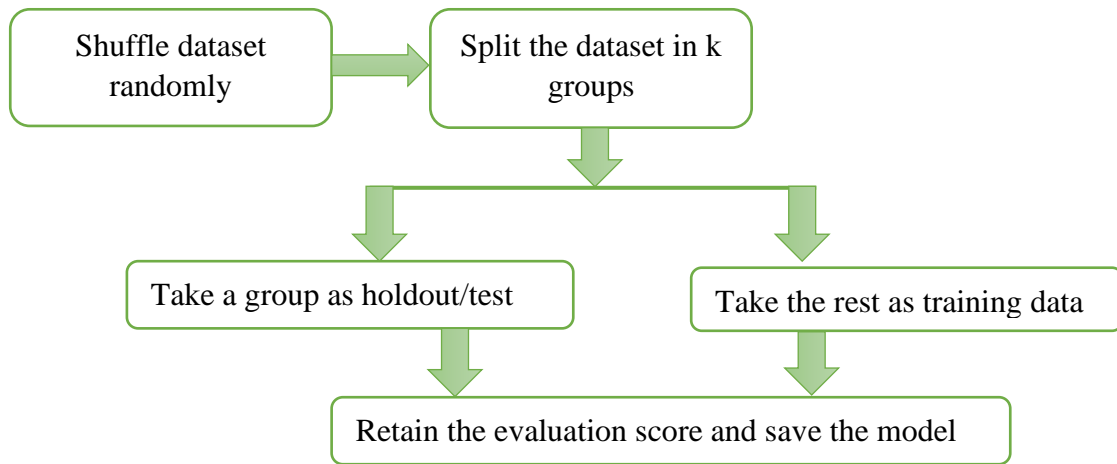
Figure 4.10: Steps followed during k fold cross validation

In this study, the k value was selected to be 10 and all dataset was split into folds. As demonstrated in Figure 4.11,the whole dataset was split into 10 foldsof training and validation block. Additionally, in each fold there exist one validation block with a different image sequence in the next fold. The dataset split for 10 fold cross-validation was set to be 90% training and 10% validation. Moreover, the model was tested by unseen images of the cervix.

Training
validation
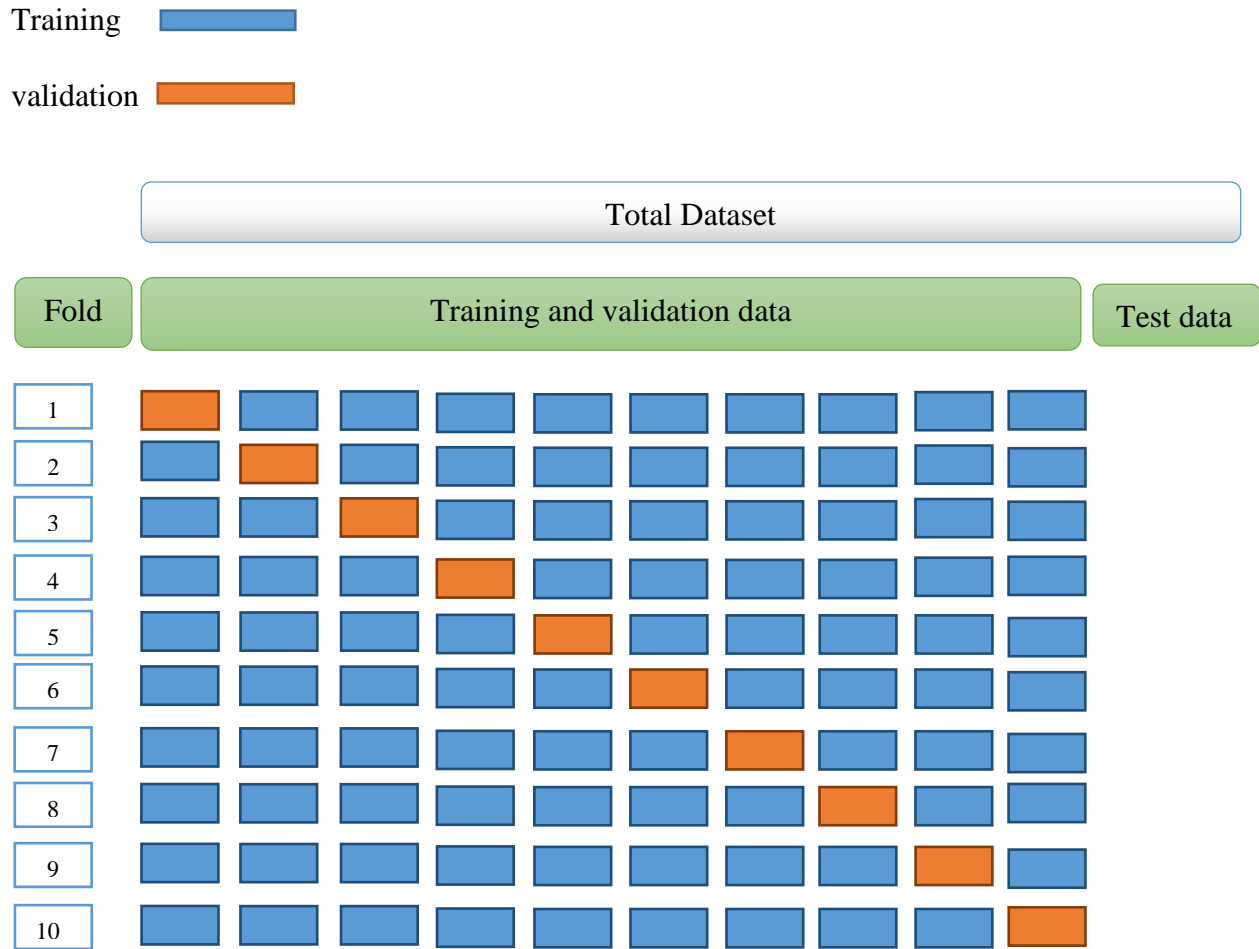
Figure 4.11: 10-fold cross-validation

4.1.7 Regularization techniques

Overfitting is a common problem in neural networks which affects the generalization ability of a model. This happens when the model trains well but losses the ability to predict in new data. It is handled in many ways including early stopping, L1, L2, and dropout techniques. In this study, an L2 layer weight decay regularizer was used to smooth the parameter distribution and reduce the magnitude of a parameter. This results in reduced overfitting and increased accuracy of the model [56].

4.2 Method overview of cervical cancer classification

For cervical histopathology image classification data was collected from Jimma University Medical center (JUMC) and St. Paul Hospital and augmented to have more data and to bring class

balance. Resizing, normalizing, and histogram matching were done after data augmentation. Finally, the preprocessed images were fed to a classification model to classify images as Normal, Precancer, Squamous cell carcinoma, and Adenocarcinoma. The cervical cancer classification architecture is shown in Figure 4.12.

Figure 4.12: Overall working scheme of cervical cancer classification
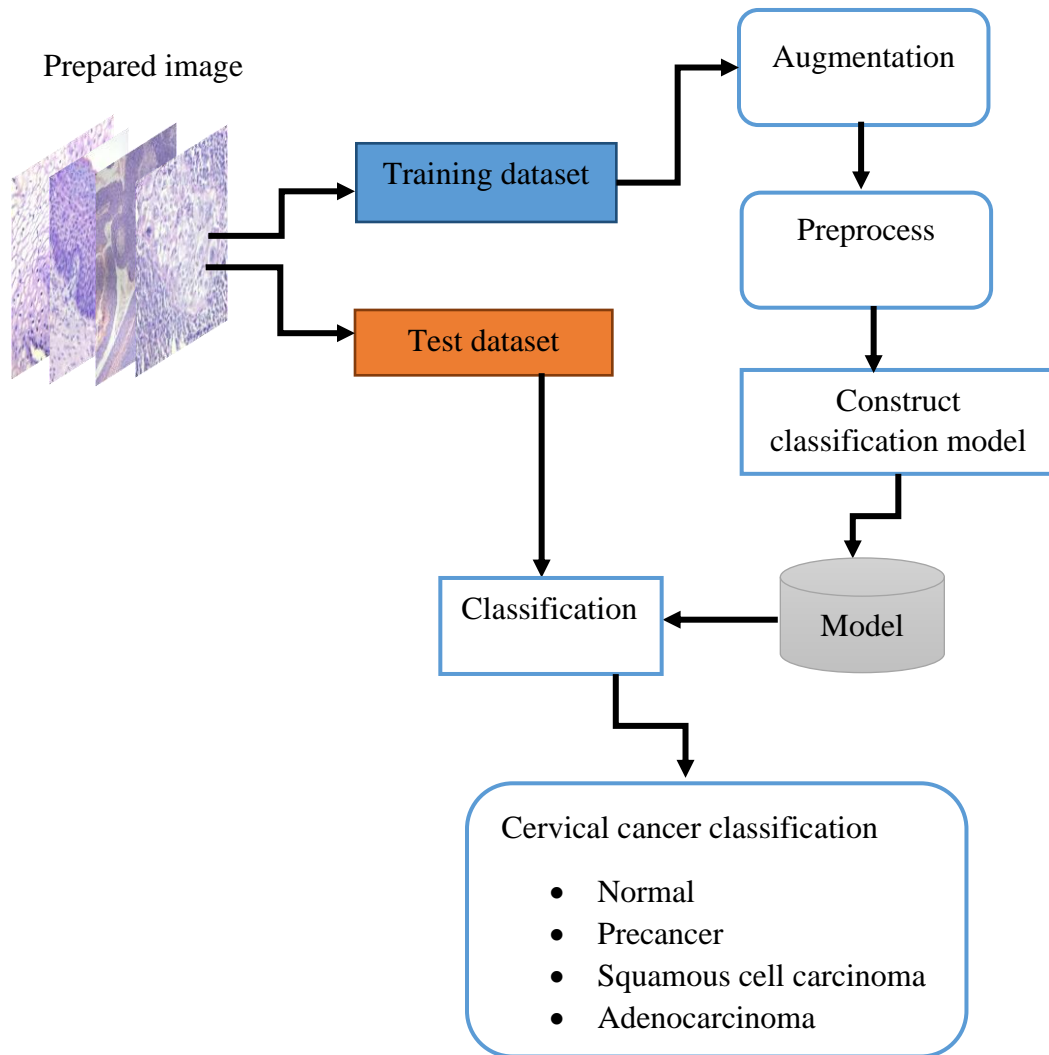
4.2.1 Dataset

For cervical cancer classification, histopathological images were collected. Hundred forty-one slides were collected; 101 from JUMC and 40 from St. Paul Hospital. Table 4.2 shows the number of images collected from each hospital. The resolution of the images was as small as 419x407 and

as large as 2048x1536. Mobile cameras, Amscope camera, and digital slide scanning were used for image acquisition.

a) Mobile camera

The dataset was collected using two microscopes with the assistance of a pathologist. All cervical cancer biopsy processed slides were investigated under a microscope. The image was taken by using 42MP Nokia mobile, 18 MP, and 13 MP Tecno mobile. Then, the image was manually cropped with the assistance of a pathologist without affecting descriptive features.

b) Amscope camera

Amscope camera is used to take the image using prepared slides in the JUMC pathology department. The calibration was done first for 4x, 10x, and 100x magnification power in scope software. Then the camera is mounted on an OPTIKA microscope and the image is snapped accordingly.

c) Digital slide scanning

The whole slide image is collected from St. Paul hospital using a digital slide scanning system. Then the image was cropped to have the same image as images acquired through smartphone and digital cameras. Then the images labeling was confirmed by pathologists in JUMC.

Table 4.2 Dataset for cervical cancer classification

From JUMCH		
Class	**Magnification**	**Number of images**
Normal	4X,10X,100X	40
Precancerous	4X,10X,100X	147
Squamous cell carcinoma	4X,10X,100X	230
Adenocarcinoma	4X,10X,100X	161
From St.Poul		
Normal	40X	77
Squamous cell carcinoma	40X	200

Adenocarcinoma	40X	60
Total		915

4.2.1.1 Data split

In this study, the dataset was split into 80% of train, 10% of validation, and 10% of test set. The training dataset was augmented before applying preprocessing method. Figure 4.13 show the proportion of image in each set with the respective class.
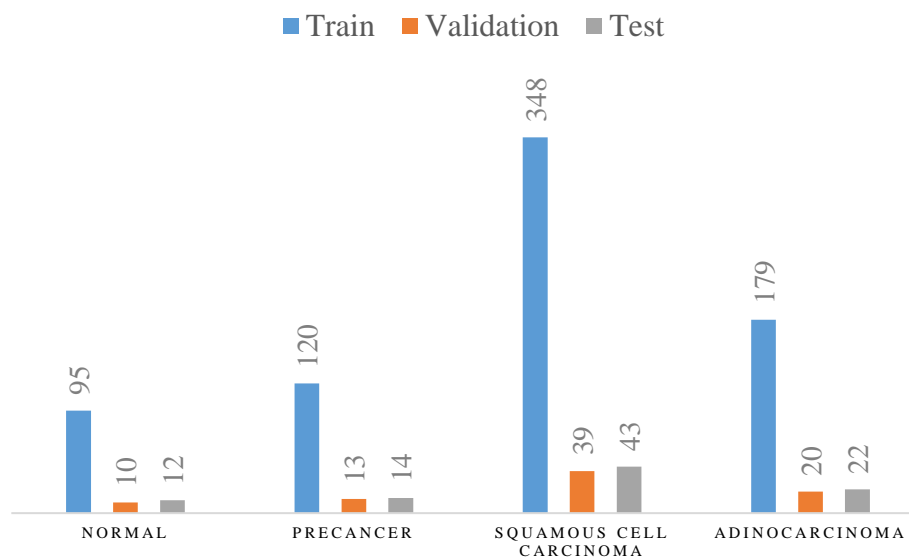
Figure 4.13: Biopsy image dataset

4.2.2 Preprocessing Method

4.2.2.1 Histogram matching

Hematoxylin and eosin (H&E) stain is one of the primary tissue stains for histology. This is because H stain makes nuclei easily visible in blue against a pink background of cytoplasm and other tissue regions. This enables a pathologist to easily identify and evaluate the tissue; a highly manual process. For automated image analysis, these (H&E) stained images need to be normalized. This is because of the significant variation in image color arising from both sample preparation and imaging conditions. Therefore normalizing H&E images is important for CNN classification[57].

Histogram matching was applied to images to reduce the color variation among images. Moreover, it is used as lightweight color normalization for image processing and is done by separating color and intensity information in space. In this algorithm, the histogram of the source image will be very much like the target image in a way such that the brightness and color of the source image will be very much like the target image [57]. Figure 4.14 shows the template image which best represents the stain of the histopathology image.
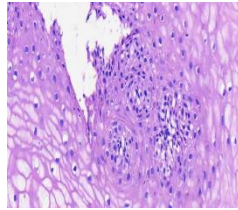
Figure 4.14: Template image

Histogram matching is done using the steps listed below [58]:

1. Equalizing the original image's histogram using cdf as a transformation function using the equation:

$$s_k \; = \; T(r_k) = \textstyle\sum_{j=0}^{k} \frac{n_j}{n} = \sum_{j=0}^{k} p(r_j) \dots\dots\dots\dots \text{(4.1)}$$

Where:

- $T(r_k)$: transformation function
- s_k : new (mapped) pixel level
- r_k : original pixel level
- n_j : number of pixels having the intensity r_k
- n : total number of pixels in the image

2. Equalizing the desired probability mass function using equation 4.2

$$V_k \; = \; G(z_k) = \textstyle\sum_{j=0}^{k} p(z_j) \dots\dots\dots\dots \text{(4.2)}$$

Where:

- $G(z_k)$: transformation function
- V_k: new (mapped) pixel level of the reference image

- z_k: original pixel level of the reference image

3. Applying the inverse transformation function to the value obtained in step 1

$$z = G^{-1}(s)\dots\dots\dots (4.3)$$

where:

- z: new pixel value after matching

The algorithm for obtaining the inverse of transformation function for a given value of s_k would search through the values of v_k and store the index of the most recent value whose absolute difference $|v_k - s_k|$ is the smallest[58]. The steps above illustrate the way of mapping template histogram to the original image histogram. This step is crucial to increase the generalization ability of the model by reducing color variation among images.

4.3 Cervical cancer classification

For cervical cancer classification, the current state of art classification model efficientnetb0 was used to classify multi-class histopathology images. The SoftMax layer of the model was trained with a raw image (histogram matching not applied) and preprocessed image using histogram matching for comparison. Adam optimizer with a learning rate was set to be 0.0001. Epoch and batch size was set to 50 and 32 respectively. The model was validated using the train-test split technique. The comparison was done to see the generalization ability of model with color normalized images and raw (not preprocessed) images with the same optimization parameters and model. Finally, unseen data were used to test the performance of model.
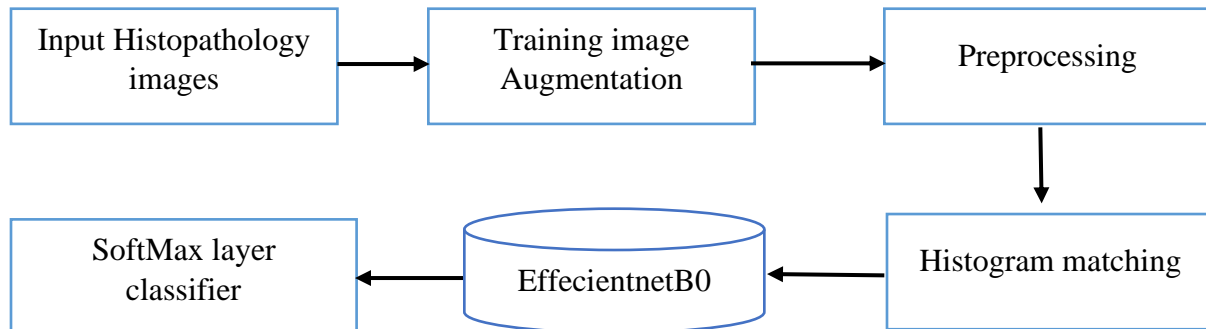
Figure 4.15: Block diagram of cervical cancer classification

4.4 Performance Evaluation Metrics

Evaluation of the model after training is important to know how well the model is working in a different aspect. Currently, the most popular performance metrics are confusion matrix, accuracy, precision, recall, F1 score, kappa score, and ROCAUC plot. The classification report was done using sklearn phyton library.

4.4.1 Confusion Matrix

The confusion matrix represents the test set summary of the classification report with respective classes in the dataset. The actual and predicted values are listed in tabular forms and give detailed information about errors being done by a classifier. More importantly, it represents the type of error being made during prediction. It works for both binary and multi-class classification problems. In this study multiclass classification problem is done so for illustration we used three-class classification evaluation metrics shown in Table 4.3 with labels A, B, and C. The blue color in the diagonal of the matrix represents a truly classified image and the rest consists of incorrectly predicted image numbers. Other performance metrics like accuracy, precision, recall, F1 score, and kappa score are estimated from confusion matrix based on TP (True Positive), TN (True Negative), FP (False Positive), and FN (False Negative) value which is represented in tabular form.

Table 4.3: Confusion matrix

It was calculated from confusion matrices as[59]:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots..(4.4)$$

$$Recall = \frac{TP}{TP+FN} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots..(4.5)$$

$$precision = \frac{TN}{TN+FP} \dotfill (4.6)$$

$$Specificity = \frac{TP}{TN+FP} \dotfill (4.7)$$

$$F - measure = \frac{2*Recall*precision}{Recall+precision} \dotfill (4.8)$$

Kappa score= (Agreement – chance agreement)/ (1-chance agreement) ……….. (4.9)

4.4.2 Receiver Operating Characteristics (ROC) and Area under ROC (AUC)

AUC ROC plot visualizes how well the model can distinguish between classes. The ROC curve can be plotted with two known methods for binary and multi-class classification which includes one versus one and one versus rest respectively. Additionally, the AUC score can be calculated for each class.

4.5 User Interface (UI)

In this study, the Tkinter platform was used to develop a user interface with an anaconda Jupyter notebook. The system was implemented using the models generated for both systems. Tkinter has many widgets used to build UI which include checkbox, button, canvas, message box, text box, etc. Overall system architecture is shown in Figure 4.16. Figure 4.17 below shows the GUI layout for cervix type classification and cervical cancer classification.
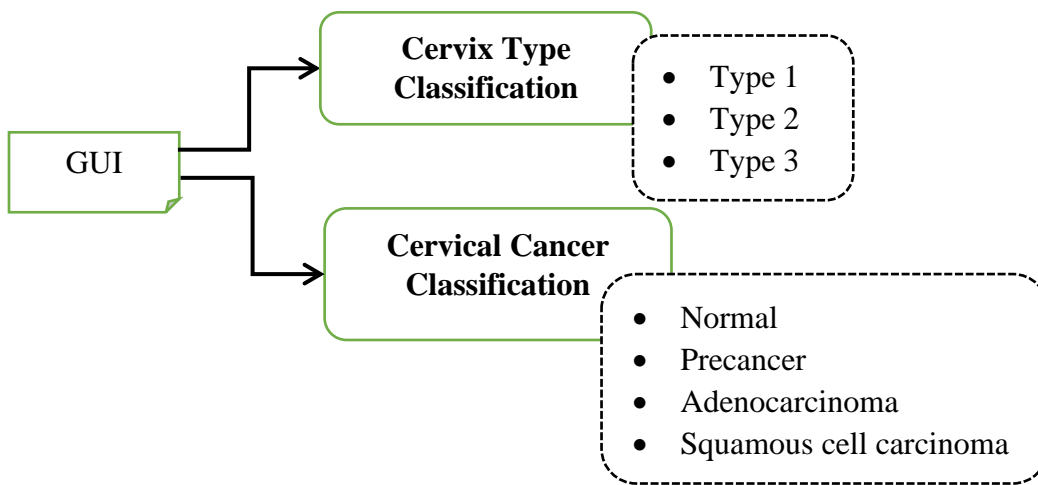

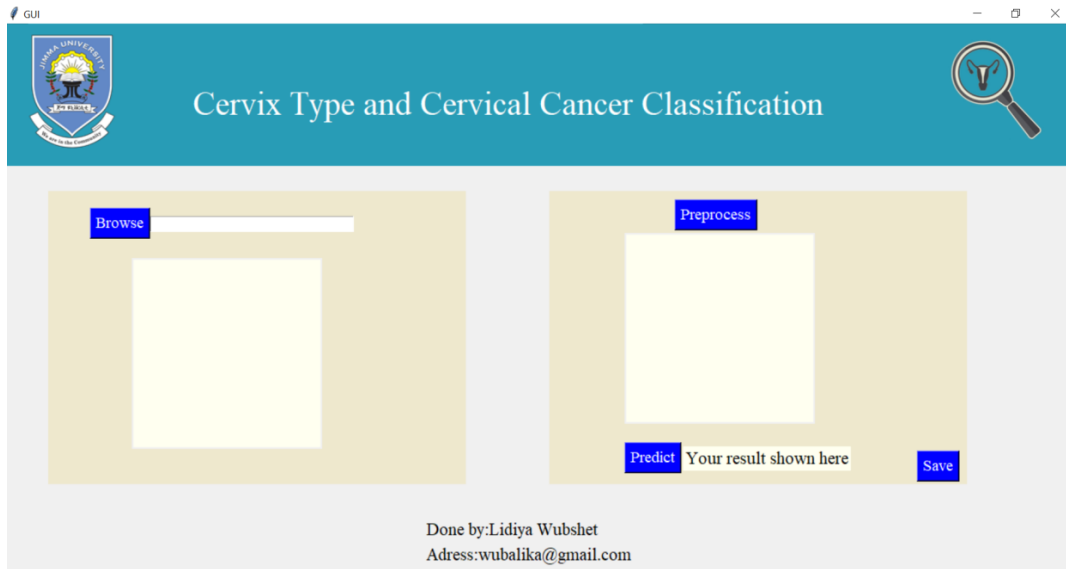Figure 4.16: The overall system architecture

Figure 4.17: GUI for classification

CHAPTER FIVE

RESULT AND DISCUSSION

5.1 Result of Cervix type classification

5.1.1 Results of pre-trained models with train-test split

Training Results for different pre-trained models are shown in Figure 5.1 - 5.4. As shown in Figure 5.1 VGG 16 model was trained for 50 epochs. During training, model accuracy increases consistently while validation accuracy is not more than 74%. In addition, training loss consistently decreases but validation loss increases this is a sign of overfitting. The lowest validation loss was achieved at the 6th epoch with a value of 0.73. The test accuracy achieved was 68.79%. The result implies that the generalization ability of the model was unsatisfactory. Similarly, for ResNet152 model in Figure 5.2 overfitting is shown and training stopped early at the 44th epoch since the validation loss was not improved for 10 epochs. The test accuracy of this model is improved compared with the former and achieved 69.25% of accuracy. Likewise, the state of art classification neural network EffecientNet_b7 was trained for 100 epochs as shown in Figure 5.3. The same overfitting problem is shown in this training result and the generalization ability of the model is low. To overcome the overfitting problem, the EfficientNet_b0 base network is trained by unfreezing 20 layers, and the result is shown in Figure 5.4.
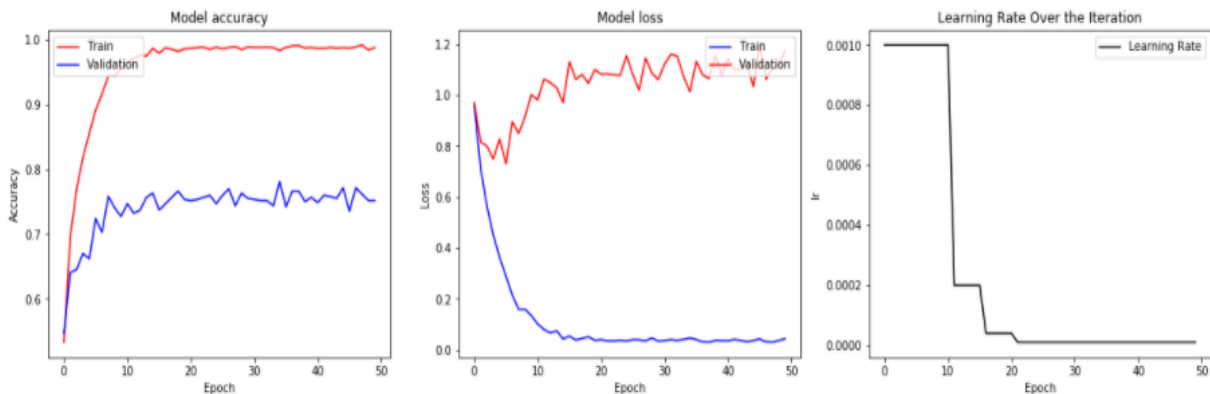

Figure 5.1: VGG 16 model training result
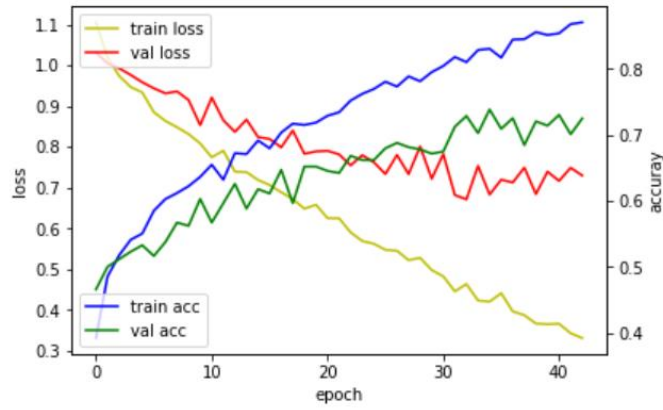
Figure 5.2: ResNet152 training result
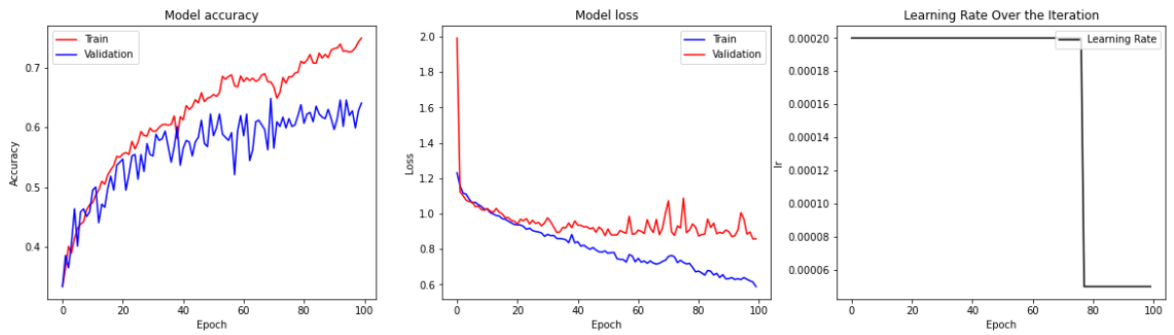
Figure 5.3: Efficientnet_b7 training result

(a)   (b)
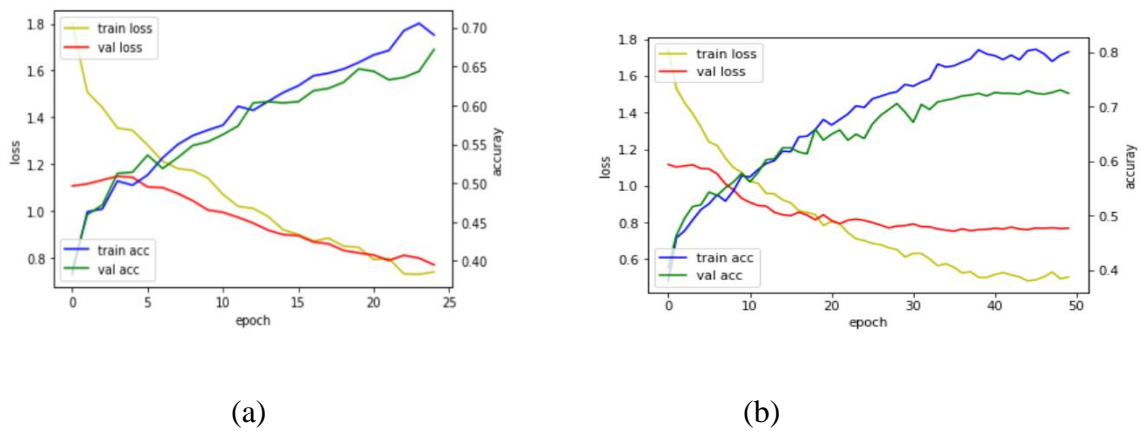
Figure 5.4: EfficientNet_b0 training result for (a) 25 epoch (b) 50 epoch

5.1.2 Result of Cervix type classification with 10-fold cross validation with ROI extraction

The three pretrained model training result shown in figure 5.5 - 5.7 with 10-fold cross validation. The vggnet16 model achieved an average accuracy of 79.67% and 1.01 loss. The Resnet50 model achieved an average accuracy of 84.17% and loss of 0.6. The effecientnetb0 model achieved an average accuracy of 89% and loss of 0.357. The effecientnetb0 yielded best result relative to other models.

Figure 5.5: Training result of EffecientnetB0
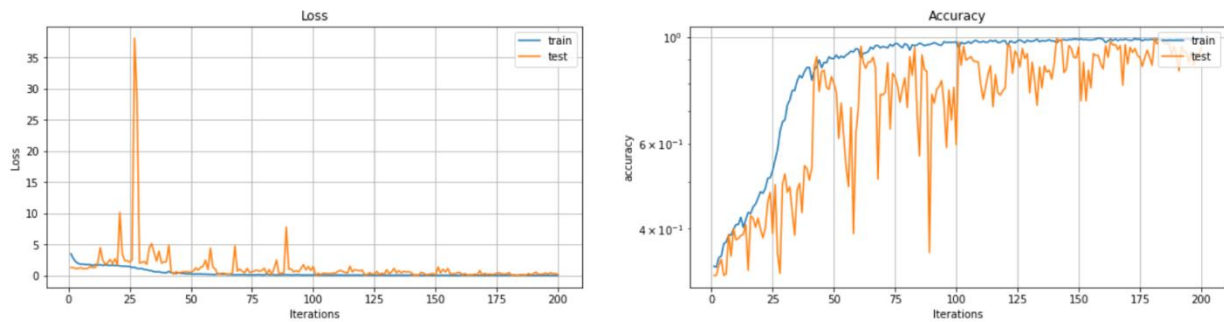
Figure 5.6: Training result of Vggnet16
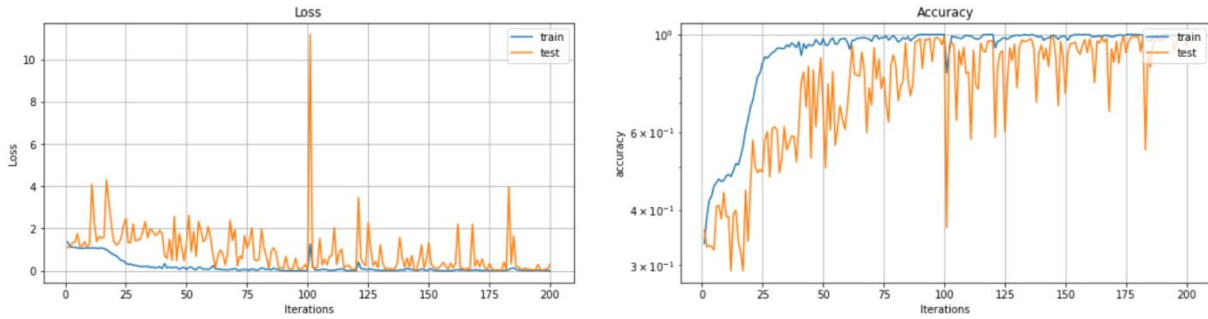
Figure 5.7: Training result of Resnet50

5.1.3 Result of Cervix type classification with 10-fold cross validation with ROI extraction

The cervix type dataset is quite difficult for differentiating since it consists of biasing features; hence the model is having difficulty in making a generalized p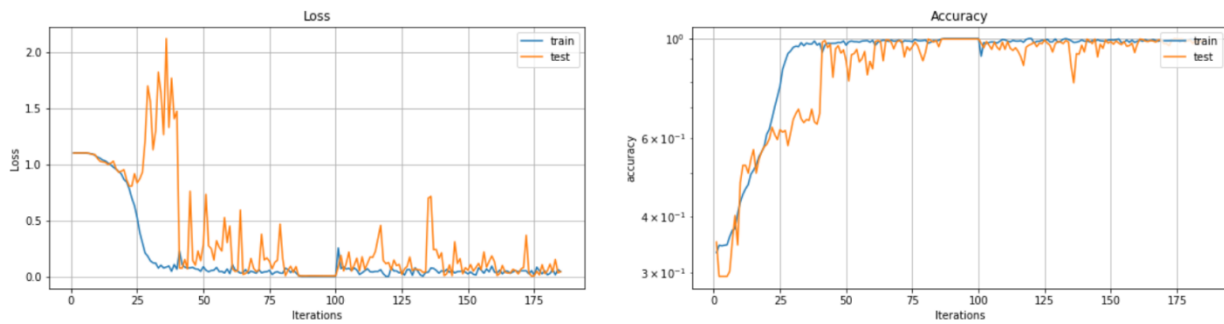rediction. Finally, with the help of ROI extraction and removing the unwanted features the model results in better performance. Additionally, using k-fold cross validation the model has a chance to train on the entire dataset which results in reduced overfitting. Furthermore, regularization techniques were used to reduce the complexity of the model, increase speed and increase the accuracy of the model.

5.1.3.1 Cervix detection

The loss versus epoch graph was shown in Figure 5.8 using a tensorboard. The result presents that as the epoch increase the train and test loss approaches zero.

Figure 5.8 Train and validation loss of the detection network versus epoch

As shown in Figure 5.8 the training and validation loss versus detection network trained for 3h 2m and 31s and yielded loss values of 0.4835 and 0.3976 respectively at 60 epochs. The network mAP was 99.89% and the output of the image was generated in crop mode for classification. For

illustration, some images are shown in Figure 5.9, which is the output images generated from the network. What we observed from the output result is that the irrelevant region from the image is removed and left with only the cervix region alone.
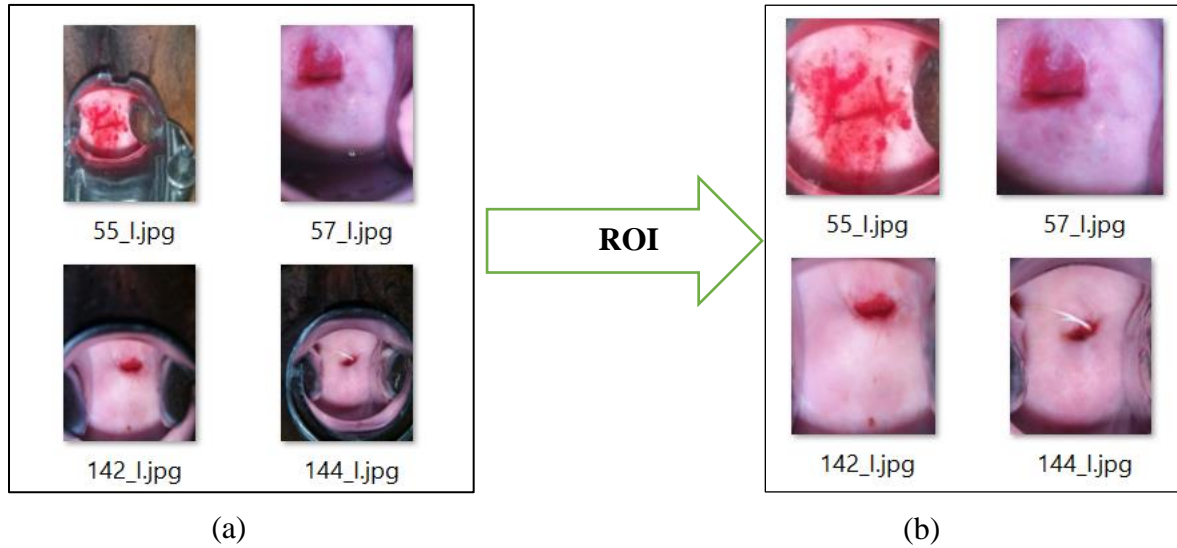
Figure 5.9 Output of the network; (a) the input image; (b) the detected region of image or output of the network
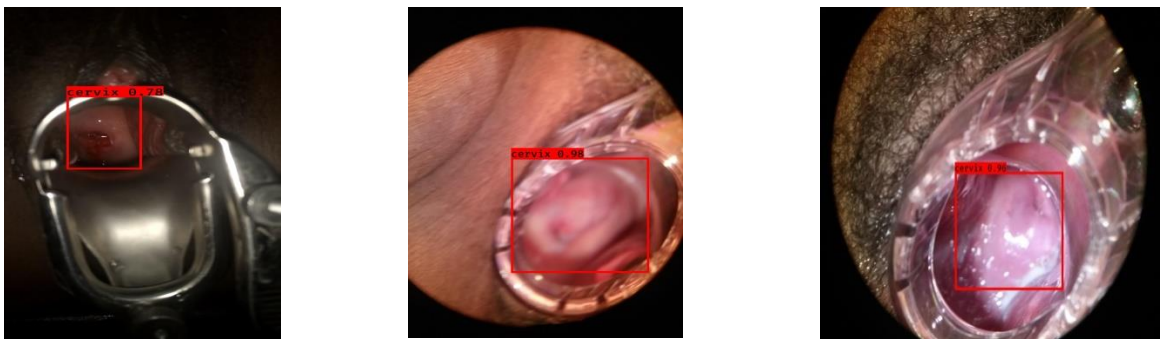
Figure 5.10: Inference images and their respective IoUscore

Some unseen images selected from the test set were shown in Figure 5.10 with a bounding box. This is an image output from detection network with rectangle on it with detection accuracy metric IoU. Additionally, the IoU score for selected images is greater than 0.5 which implies the model performed well. In comparison, the local image IoU score is less due to its low resolution.

5.1.3.2 Cervix type classification

Training results for cervix type classification are presented in Figures 5.11-5.13 using different models. For all training10-fold cross-validation was used with Adadelta optimizer with starting values of learning rate 0.00001. The ensemble of Mobilenetv2 and Resnet50 model was trained and is shown in Figure 5.11 and VGG16 is shown in Figure 5.12. The EfficientnetB0 model training was done using the same optimizer. The result in Figure 5.13(a) represents a graph of model trained with a 0.00001 learning rate and 10 epochs in each fold with batch size 32. Besides, the result in Figure 5.13(b) represents a graph of model trained with a learning rate of 0.0001 and 25 epochs in each fold with batch size 32. In the end, the result shows efficientnetb0 training results achieved the best results.
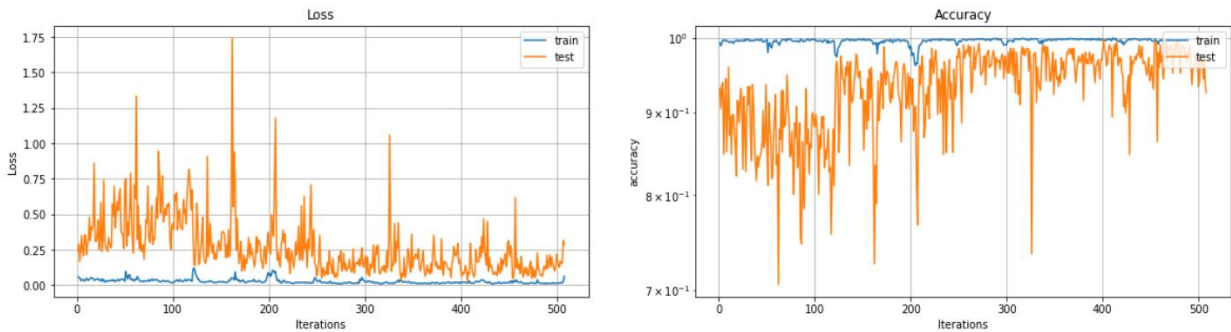
Figure 5.11: Training result of Ensemble of Mobilenetv2 and Resnet50
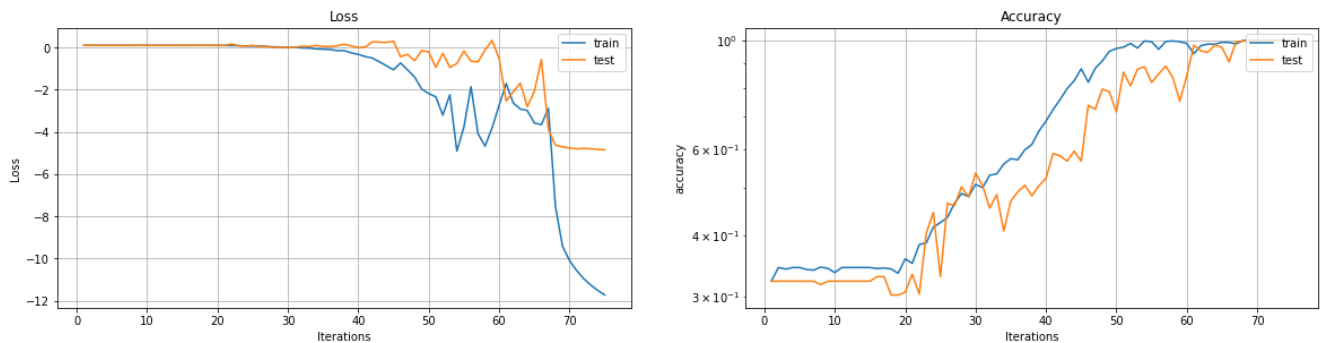
Figure 5.12: Training result of VGG16

(a)

(b)
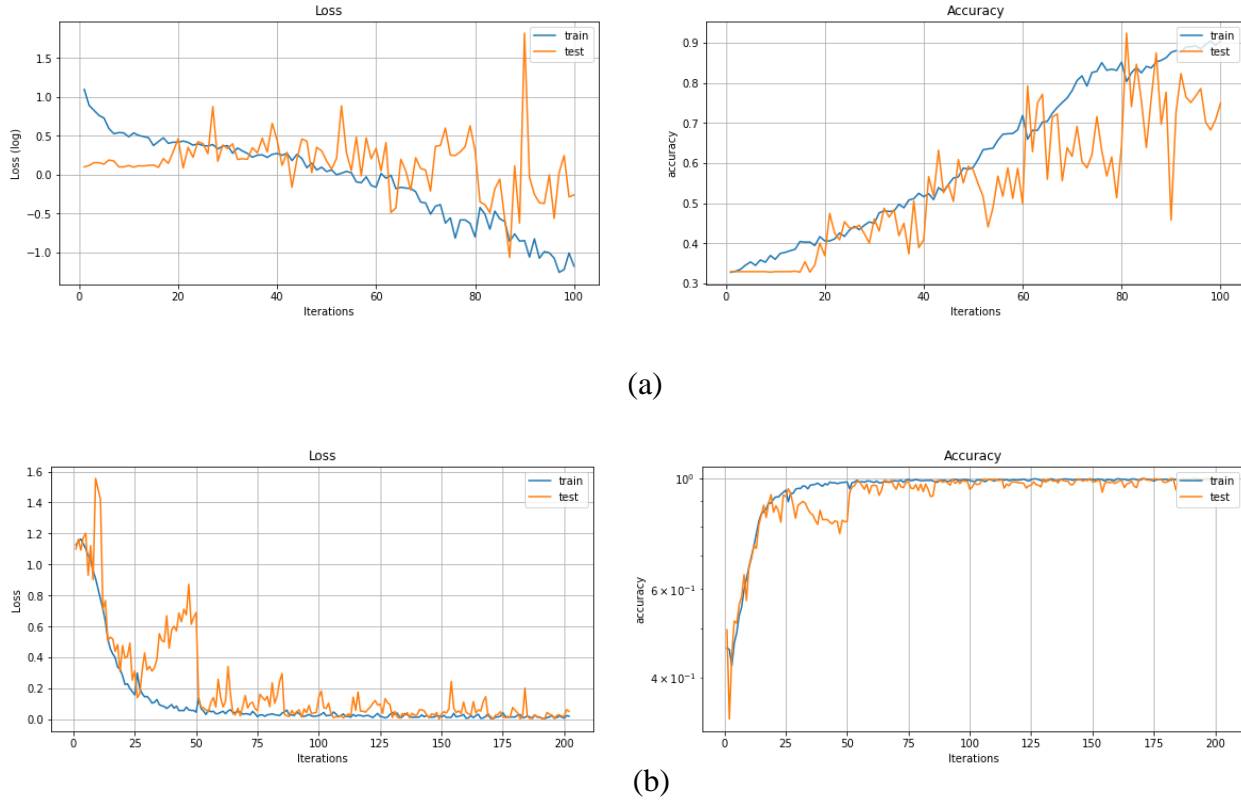
Figure 5.13:Trainning result of EfficientnetB0 (a) 5-fold cross validation (b) 10-fold cross validation

The validation accuracy of the dataset was calculated in each fold and finally, the average accuracy was calculated. Table 5.1 presents the accuracy and loss of training in each fold. The average accuracy of the result is 92.1956 and an average loss of 0.2724. The best accuracy score was achieved at the 10th fold with the smallest loss of 0.1313. As shown in Figure 5.13 better validation accuracy was achieved compared with another fold accuracy as well as loss value.

Table 5.1:Scores per fold

Folds	Loss	Accuracy
1	0.2733	90.7514%
2	0.3839	89.0173%
3	0.4624	88.4393%
4	0.2254	92.4855%
5	0.1659	95.3757%

6	0.3366	90.7514%
7	0.2892	92.4855%
8	0.2590	91.6184%
9	0.1974	93.9306%
10	0.1313	97.1014%

The test result is presented in terms of a confusion matrix. From the result, we observed that the model performs well. The correctly classified images are in the diagonal with dark blue color. Figure 5.14 shows the confusion matrix for actual and predicted images of test set.
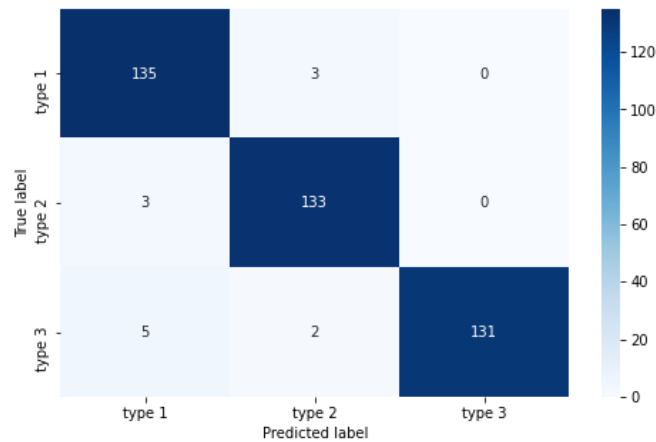
Figure 5.14: Confusion matrix for test set

The precision, recall, and f1-score are calculated using test images as shown in table 5.2. The macro average and weighted average are equal since the datasets are balanced. The value illustrates that the model performs well. Figure 5.15 shows the ROCAUC curves with AUC score for each class 0, class 1, and class 2 on top of curve respectively. Table 5.3 illustrates the summary of result of selected models and the effecientnet_b0 outperformed the other models which are trained with a low learning rate. The result shows that training model with extracted ROI achieved better compared with the model trained without extracting ROI trained with the same model and hyperparameters. The missed class represents the similarity among images for example 3 images in type 1 are misclassified as type 2 since type 1 and type 2 classes are similar. Likewise, 3 type 2

class images are misclassified as type 1. Both classes are not misclassified as type 3 since the image in type 1 and type 2 is not close in similarity to type 3.

Table 5.2: Result of cervix type classification with different metrics for the final model

Classes	Precision	Recall	F1-score	Number of images
Type 1	0.94	0.98	0.96	138
Type 2	0.96	0.98	0.97	136
Type 3	1.00	0.95	0.97	138
Macro avg	0.97	0.97	0.97	412
Weighted avg	0.97	0.97	0.97	412
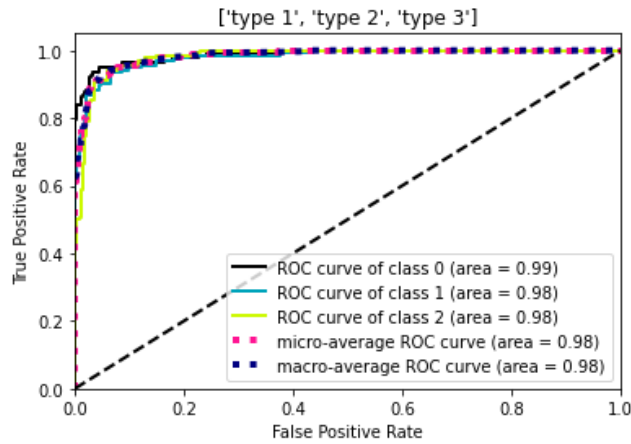
Figure 5.15: ROCAUC score for cervix type classification

Table 5.3: Comparison of results using different models in this study

Model	Optimizer	The starting point of the Learning rate	Batchsize	Performance (accuracy)
Training Pretrained models without ROI and train test validation				
ResNet50	Adadelta	0.00001	32	84.17%
VGG16	Adadelta	0.00001	32	79.67%

Efficiennetb0	Adadelta	0.00001	32	89%
Training pretrained model with detected ROI and 10 Fold cross-validation				
ResNet50	Adadelta	0.00001	32	90.73%
VGG16	Adadelta	0.00001	32	89%
Ensembing ResNet50 with MobilenetV2	Adadelta	0.00001	32	92%
Efficiennetb0	**Adadelta**	**0.00001**	**32**	**92.19%**

5.2 Result of Cervical cancer classification

5.2.1 Preprocessing histopathologic image

The first step of cervical cancer classification was preprocessing histopathologic images. The image is resized to 224x224 height and width using Keras. The resized image is shown in Figure 5.16.
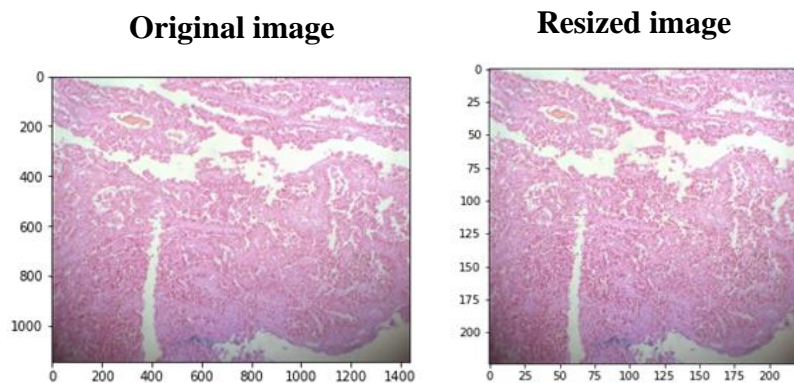
Figure 5.16: Resized Image

The class balance was done using the augmentation technique of 0^0 45^0, 90^0, 180^0, and 270^0 on the original image. Not all rotation is applied for the image except 45^0 rotation which is applied for all class images. Figure 5.17 presents different degrees of augmentation techniques used in this study. Figure 5.18 represents the proportion of augmented images used for training.

| Original image | 45⁰ Rotation | 90⁰ Rotation | 180⁰ Rotation | 270⁰ Rotation |

Original image   **45^0 Rotation**   **90^0 Rotation**   **180^0 Rotation**   **270^0 Rotation**
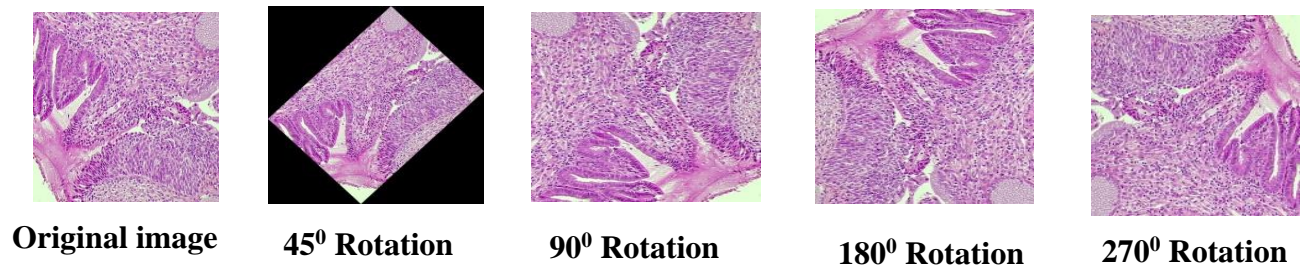
Figure 5.17 Image Augmentation using different degree of rotation

Augmented Images for Training

■ Adinocarcinoma ■ Normal ■ Precancer ■ Squamous cell carcinoma
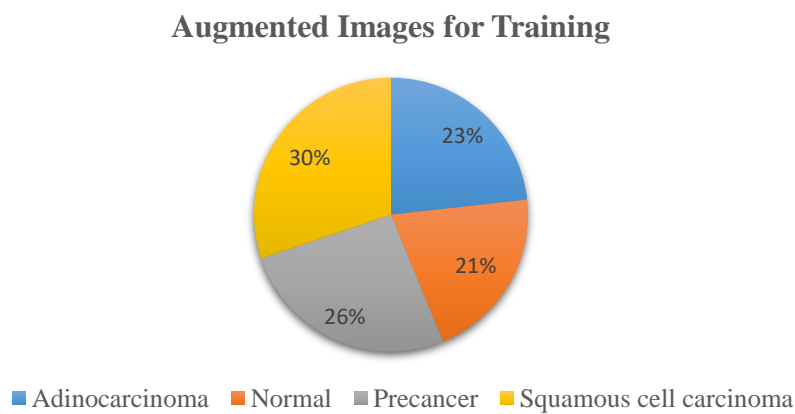
Figure 5.18: Augmented images for training

5.2.2 Histogram Matching

The histogram matching effect was shown in Figure 5.19 with a template image. It is a lightweight color normalization technique in image processing and matches the histogram of the template image to the original image.
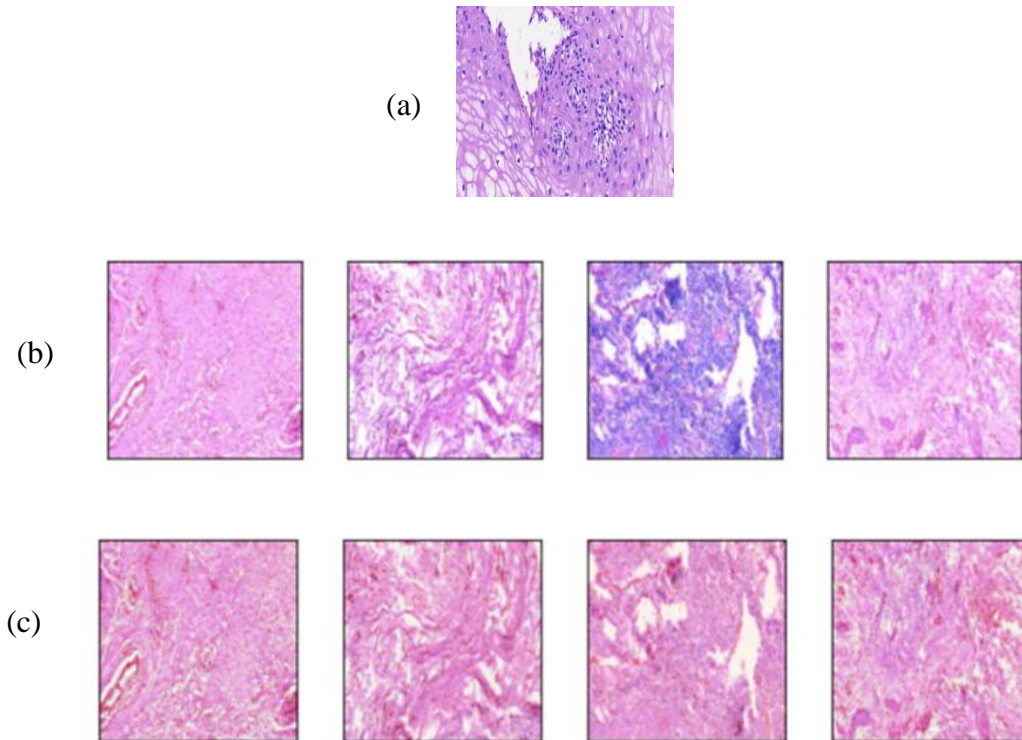
Figure 5.19: Result of histogram matching (a) Template image (b) Input image (c) Matched image with a reference image

5.2.3 Training result

The pre-trained efficientnet_b0 model was used to classify a cervical histopathologic image after organizing the data well, splitting the data, and pre-processing the image. Figure 5.20 shows the model accuracy and loss curves during training. The accuracy vs. epoch curve in the figure consistently approaches 100 for both train and validation data. Unlike the model accuracy curve loss vs. epoch, the curve approaches zero which implies that the model performs well. The highest accuracy achieved during training is 0.9780 at the 50th epoch with the smallest loss value of 0.0784.

(a)   (b)
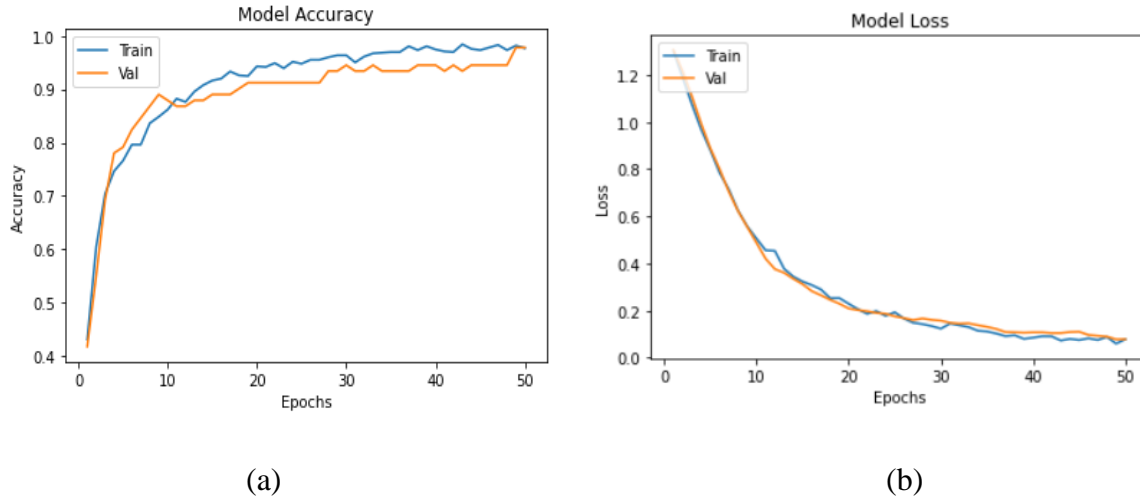
Figure 5.20: Result of Training the model

5.2.4 Test Result

The confusion matrix presented in Figure 5.21 implies that the actual value is listed vertically and the predicted value horizontally. The diagonal numbers in the matrix represent correctly classified images and the other blocks represent incorrectly classified images. The misclassified image represents that there are some similar features in some level of degree since the confusion matrix is done on a threshold value. Table 5.4 presents the accuracy, precision, recall, and F1 score with macro average and a weighted average value for each class. The macro average gives each prediction a similar weight while calculating loss but the weighted average gives importance to some predictions more for imbalanced data. The last column represents the number of images with respective classes used to test the model. Figure 5.22 shows the AUC-ROC using the one versus rest approach. The AUC score for all classes is near to 1 which confirms the model has a good measure of separability. The ROC curve is above the diagonal line where a model has good diagnostic ability at varied thresholds. The list of classes above the plot under square bracket implies class 0, class 1, class 2, and class 3 respectively for AUC score.

Generally, the test result shows the model performs well as shown in different performance measuring metrics. Two images were incorrectly classified as shown in the confusion matrix at a specific threshold value. Additionally, the ROCAUC represents the classification ability at varying threshold values.
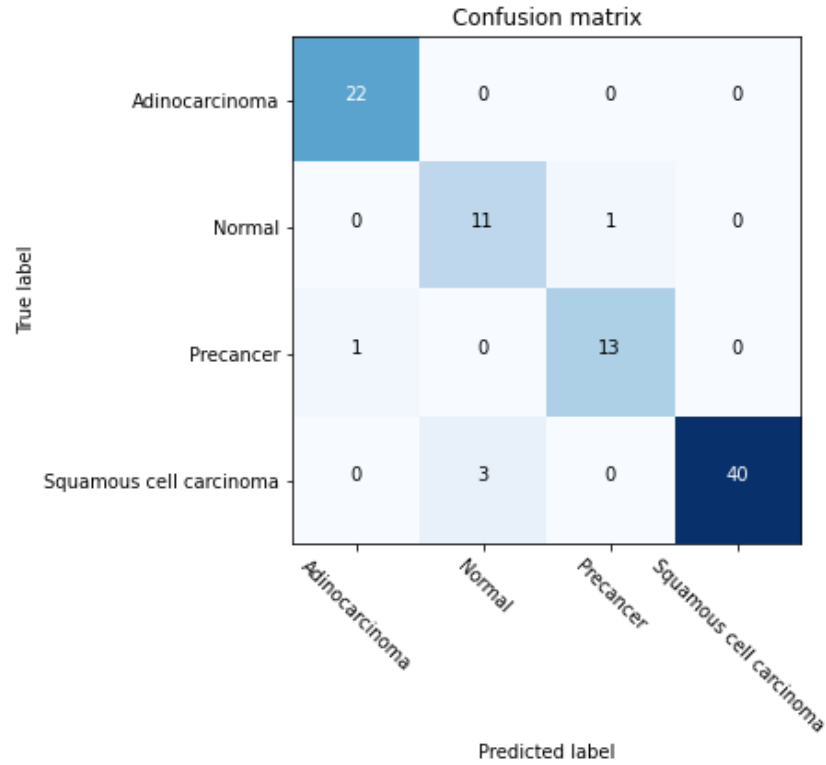
Figure 5.21: Confusion matrix for cervical cancer classification

Table 5.4: Result of cervical cancer classification with different metrics

Classes	Precision	Recall	F1-score	Number of images
Adinocarcinoma	0.96	1.00	0.98	22
Normal	0.79	0.92	0.5	12
Precancer	0.93	0.93	0.93	14
Squamous cell carcinoma	1.00	0.93	0.96	43
Macro avg	0.92	0.94	0.93	91
Weighted avg	0.95	0.95	0.95	91
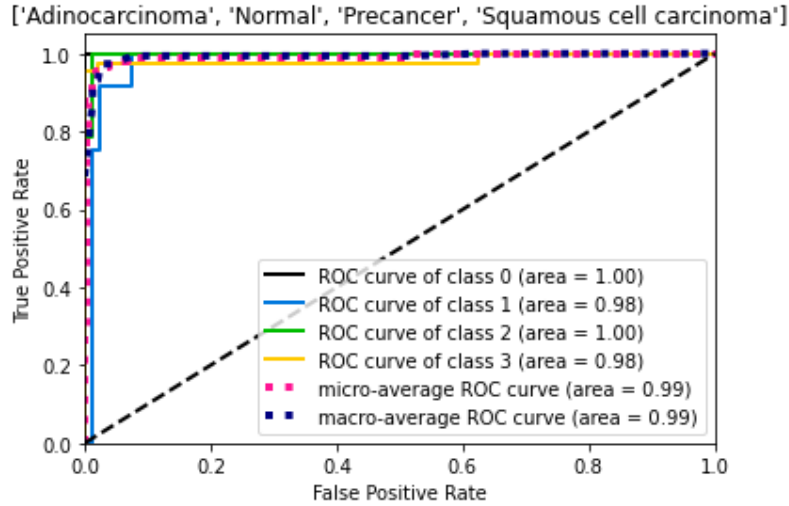
Figure 5.22 ROCAUC plot

The comparison of the test result for images preprocessed with and without histogram matching was done during the training and testing phase as shown in table 5.5. The test result obtained for images preprocessed with histogram matching yielded better than the image without applying histogram matching.

Table 5.5: Comparison of model test result with and without applying histogram matching on the image

Preprocess image	Precision	Recall	F1-score	kappa score	Accuracy	Number of images
With Histogram matching	0.92	0.94	0.93	0.920	0.95	91
Without Histogram matching	0.90	0.95	0.92	0.874	0.91	91

5.3 Graphical user Interface

The result button triggers the classification result to be displayed for each work using the respective models. The developed GUI includes canvas, buttons to browse, and load images then the loaded image seen in the first canvas. Additionally, the preprocessed and cervix image buttons are responsible to show user histogram matched histopathology image and ROI of cervix respectively as shown in figure 5.23 and 5.24. In the end, the save button is used to save the diagnosis result and the exit button is quite the GUI as shown in figure 5.25. Table 6.6 represents the sample questions for experts concerning developed GUI.
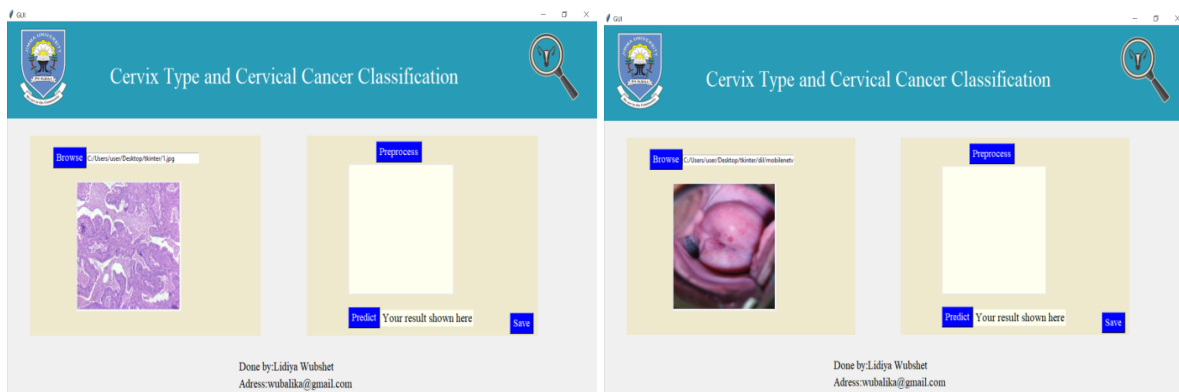
Figure 5.23: Browse Image for both cervix type and cervical cancer
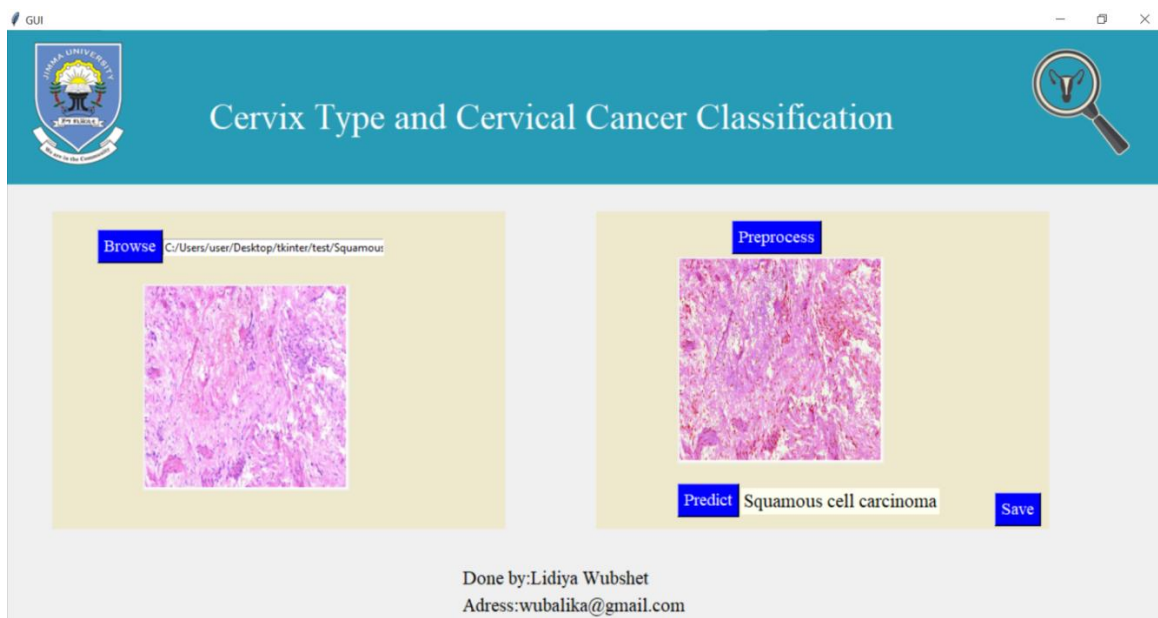
Figure 5.24: End to end histopathology image classification
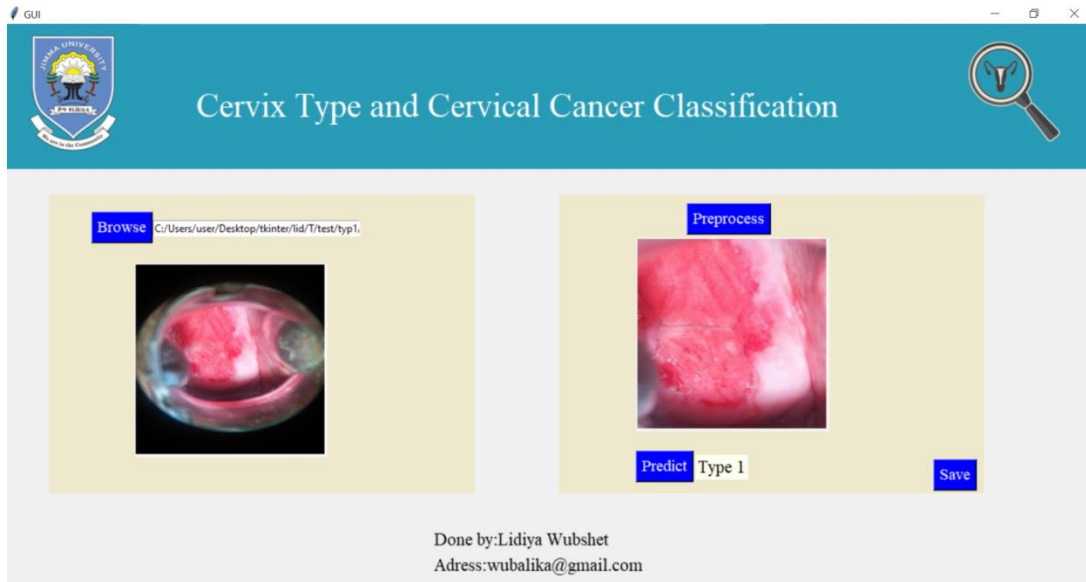
Figure 5.25: End to end cervix type classification

Hereunder in table 5.6 summary of user acceptance testing is presented. Ten gynecologists and pathologists participated in the evaluation. Accordingly, the result shows that the proposed prototype is accepted by all evaluators.

Table 5.6: Sample questions for experts about GUI

Questions	The number of people who participated	
	Accept	Reject
1. Easy to use	10	0
2. Speed of system	10	0
3. Storage requirement	10	0
4. Layout	10	0
5. Result of Acceptance	Accepted	

The feedback from gynecologists and pathologists about GUI includes extending the user interface to a mobile app or web system, sub-class classification on cervical histopathology image classification, and adding the result to be printed with patient information. We take the feedback as a comment and it will be included in the feature or extended by another researcher.

5.4 Discussion

This study aims to classify cervix type and cervical cancer using deep learning techniques. Cervix type is classified based on transformation zone location as type 1, type 2, and type 3. While cervical cancer is classified as Normal, Precancer, Squamous cell carcinoma, and Adenocarcinoma. For cervix type classification different methods were proposed to have the best model of classification. Of this, the first was done using the pre-trained model without extracting the region of interest. In this, the low accuracy with overfitting was observed as shown in Figure 5.1 -5.4. In order to solve the overfitting problem, 10-fold cross-validation and L2 regularization were used. Besides, to increase the generalization ability of the model, a two-step solution was proposed. Then, cervix type classification was done by extracting the region of interest followed by the classification model. This method performs better than the model trained without extracting ROI with the same optimization parameters and models. MobileNetv2-YOLOv3 end-to-end object detection was used for extracting ROI based on the concept of regression, which gives us the ROI of an image used as an input for classification. The model input size is 416 x 416 which is used for large objects and the detection was done at three different scales by downsampling the image by 32, 16, and 8. Then the model crop ROI using bounding box vertices as shown in Figure 5.9 (b). Whereas, the classification was done by comparing the popular deep learning models such as VGG16, ResNet 50, an ensemble of ResNet50 and Mobilenetv2, and EfficientNet families. The classification networks were trained with a very low learning rate of 0.00001 as a starting point and L2 weight decay regularizes applied in each layer of selected models with randomly initialized weights. The model was trained for 25 epochs in each fold with Adadelta optimizer, 32 batchsize, and an automatically adjustable learning rate. The training result for selected models is presented in Figure 5.11-5.13. For cervical cancer classification, the current state of art classification neural network effecientnet_b0 is used to solve the classification problem. The softmax layer of effecientnet_b0 was trained with hyperparameter 0.0001 learning rate, adam optimizer for 50 epochs, and batchsize of 32. The model with and without applying histogram matching was compared using the same hyperparameter for both. The model performance was evaluated using confusion matrix, accuracy, recall, precision, F1 score, Kappa score, and ROC AUC plot for both classification tasks.

Automating cervix type classification was done using a CAD system in previous work with the Kaggle dataset [50]. It was done either by extracting the cervix region before classification [12, 37–42] or fine-tuning the pre-trained model [34–36]. The best result was achieved by extracting the region of interest using different methods. This was done to remove unwanted features which made the image too different since it includes speculum, vagina wall, hands of professional, and hair around the vagina. The best result achieved in past work was done using Mask RCNN for ROI extraction and HCME algorithm for classification of cervix type [12]. Even though the result is 96.77% of accuracy and 95.1 of kappa score the study used computationally complex cervix segmentation techniques. Compared with this study our method achieved a relatively better result with a test accuracy of 97% and a kappa score of 95.2. Additionally, the real-time object detection of the lightweight model achieved a mAp of 99.89% and which ease the model deployment. Cervical cancer classification using CAD system is compared with previously done works using the dataset collected from their respective hospitals [43–49], it performs better. This study was done to further classify cervical histopathology images unlike previous works done for specific class classification. Moreover, this study overcomes magnification factor dependency but in previous studies, specific magnification factors are shown.

In this study, different methods were proposed to achieve the best result with some selected models for comparison of cervix type classification. Similarly, a comparison was done for cervical cancer classification with and without applying preprocessing methods. For both classification task test result of the model was presented in terms of confusion matrix, accuracy, recall, F1 score, kappa score, and AUCROC curve. The performance metrics show that the model performs well on the test set. Finally, the cervix type and cervical cancer model user interface were done using Tkinter.

5.5 Contribution of the study

The main contribution of this study is making one system for diagnosis of cervical cancer and the cervix type classification for a decision of treatment option method. Clinically, the cervix type classification is important to decide treatment options and to take biopsy as well from the transformation zone since it's the region where cervical cancer starts. On the other hand, a biopsy test is a gold standard or final confirmation to have cervical cancer. Therefore, taking the biopsy from the region of the cervix with help of cervix type classification and confirmation to have cancer and then deciding the treatment option.

CHAPTER SIX

CONCLUSION AND RECOMMENDATION

6.1 Conclusion

Cervical cancer is one cause of cancer-related death for women worldwide. Currently, the available screening and diagnosis infrastructure of cervical cancer are expensive with a limited number of pathologists and gynecologists. Due to this, the incidence and mortality rate is still a burden for low-income countries. In addition to that existing diagnosis, a method is subjective since it's done manually. The CAD system developed recently used a limited number of images and specific class classification has been done for cervical cancer classification.

The proposed study includes two unrelated image classifications but is clinically dependent and significant. This includes the cervix type and cervical cancer classification which is done using DNN. The cervix type class includes type1, type2, and type3 while the cervical cancer class includes Normal, Precancer, Adenocarcinoma, and squamous cell carcinoma. The cervix from the image is detected using MobileNetv2-Yolov3 and10-fold cross-validation with efficientnetb0 model outperforms the existing as well as experiments done throughout the study period. The test result with unseen data for similar model (Effecientnetb0) with and without extracting the region of interest the model with extracted ROI improved 3.19% of accuracy. This study improved accuracy by 0.33% with a lightweight ROI extraction method compared with past work. For cervical cancer classification, EfficientNetb0model achieved the best result with a preprocessed image using histogram matching improved 4% accuracy compared to a model trained without applying histogram matching. The past work proposed to classify cervical cancer was mainly focused on the specific class classification that is by taking one class and classifying it to sub-class and most of them focused on machine learning. This study outperforms the pat work which is similar by class classification with an accuracy of 7%. The developed GUI system assists gynecologists, gynecologist oncologists, and pathologists.

In this study, we tried to automate cervix type into three classes and cervical cancer classification into four major classes. However, for cervical cancer classification, all the dataset was locally collected and not enough to do more experiments. In addition, due to time limitations, the cytology

test or pap smear was not included in this study. Besides, we developed locally used GUI or desktop app.

6.2 Recommendation

Cervix type classification was done using public data with some more images locally collected. Our recommendation for this is to use more data as much as possible since having more image result in generating a better model and allow to do more experiments. The cervical cancer classification was done by locally collected small images from JUMC and St.Paul hospital. This allows us to only classify major classes as normal precancer and two cancer types. This classification task will be extended to sub-class classification is important for professionals to know the degree of cancer and prognosis of the disease. In addition, in order to address the current problems of difficulties in early screening, diagnosis and treatment of cervical cancer and to make the system full by including papsmear image or cytology image is also important. Moreover, developing a web system for user interface by upgrading our GUI is more important to address all hospitals.

Reference

[1]   W. S. Jr, M. A. Bacon, A. Bajaj, and L. T. Chuang, "Cervical Cancer : A Global Health Crisis," *Cancer*, pp. 1–9, 2017, doi: 10.1002/cncr.30667.

[2]   G. Jassim, A. Obeid, and H. A. Al Nasheet, "Knowledge , attitudes , and practices regarding cervical cancer and screening among women visiting primary health care Centres in Bahrain," *BMC Public Health*, pp. 1–6, 2018, doi: 10.1186/s12889-018-5023-7.

[3]   "Human Papillomavirus (HPV) and Cervical Cancer," 2019. https://www.who.int/en/news-room/factsheets/detail/human-papillomavirus-(hpv)-and-cervical-cancer.

[4]   F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA. Cancer J. Clin.*, vol. 68, no. 6, pp. 394–424, 2018, doi: 10.3322/caac.21492.

[5]   M. Simon M, T. Alan T, H. Warner K, N. Calvin, and L. Margaret I, "Type 3 Transformation Zone of the Cervix and Risk of Missed Lesions during Cervical Cancer Screening with Visual Methods: A Case Report from Cameroon," *Obstet. Gynaecol. Cases - Rev.*, vol. 8, no. 1, pp. 8–11, 2021, doi: 10.23937/2377-9004/1410195.

[6]   T. B. Tad, M. Ezgi, L. W. Donald, G. E. Joann, and MPH4, "Accuracy is in the Eyes of the Pathologist: The Visual Interpretive Process and Diagnostic Accuracy with Digital Whole Slide Images," *Physiol. Behav.*, vol. 176, no. 12, pp. 139–148, 2017, doi: 10.1016/j.jbi.2017.01.004.Accuracy.

[7]   G. Okubazgi, H. T. L. Ascp, B. Berhane, M. Nigussie, A. Tsegaye, and F. Hassen, "Status of Histopathology Services in Ethiopia," *Am J Clin Pathol*, vol. 153, pp. 3–4, 2020, doi: 10.1093/AJCP/AQZ144.

[8]   F. Ruddies *et al.*, "Cervical cancer screening in rural Ethiopia: A cross- sectional knowledge, attitude and practice study," *BMC Cancer*, vol. 20, no. 1, pp. 1–10, 2020, doi: 10.1186/s12885-020-07060-4.

[9]   B. Nathalie *et al.*, "WHO guidelines for screening and treatment of precancerous lesions for cervical cancer prevention," *World Heal. Organ.*, pp. 1–36, 2013.

[10]   A. Grover and D. Pandey, "Anatomy and physiology of cervix," in *Colposcopy of Female Genital Tract*, Springer Science and Business Media Singapore, 2017, pp. 1–16.

[11]   "Understanding Cervical Changes: A Health Guide for Women - National Cancer Institute." https://www.cancer.gov/types/cervical/understanding-cervical-changes (accessed Nov. 06, 2020).

[12]   R. Gorantla, "Cervical Cancer Diagnosis using CervixNet - A Deep Learning Approach," *IEEE 19th Int. Conf. Bioinforma. Bioeng. (BIBE)At Athens, Greece*, doi: 10.1109/BIBE.2019.00078.

[13]   J. Payette, J. Rachleff, and C. Van De Graaf, "Intel and MobileODT Cervical Cancer Screening Kaggle Competition : Cervix Type Classification Using Deep Learning and Image Classification," 2017.

[14]   S. R. Prendiville W, "Colposcopic terminology : the 2011 IFCPC nomenclature," in *Colposcopy and Treatment of Cervical Precancer*, 2017, pp. 51–57.

[15]   H. A. Robbins *et al.*, "Cervical cancer screening intervals and management for women living with HIV : a risk benchmarking approach," *AIDS*, no. February, pp. 1035–1044, 2017, doi: 10.1097/QAD.0000000000001450.

[16]   S. R. Prendiville W, "Squamous intraepithelial lesions: cytology–histology correlation," in *Colposcopy and Treatment of Cervical Precancer*, no. 45, International Agency for Research on Cancer, Lyon (FR), 2017, pp. 23–27.

[17]   N. Colombo, S. Carinelli, A. Colombo, C. Marini, D. Rollo, and C. Sessa, "Cervical cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up," *Ann. Oncol.*, vol. 23, no. 7, pp. vii27–vii32, 2012, doi: 10.1093/annonc/mds268.

[18]   M. Ardahan and A. B. Temel, "Visual inspection with acetic acid in cervical cancer screening," *Cancer Nurs.*, vol. 34, no. 2, pp. 158–163, 2011, doi:

10.1097/NCC.0b013e3181efe69f.

[19]   "Cervical Cancer Screening (PDQ®)–Patient Version - National Cancer Institute." [Online]. Available: https://www.cancer.gov/types/cervical/patient/cervical-screening-pdq.

[20]   M. Slaoui and L. Fiette, "Histopathology Procedures : From Tissue Sampling to Histopathological Evaluation," *Natl. Cent. Biotechnol.*, pp. 69–82, 2011, doi: 10.1007/978-1-60761-849-2.

[21]   S. I. Lee and M. Atri, "2018 FIGO staging system for uterine cervical cancer: Enter Cross-sectional Imaging," *Radiology*, vol. 292, no. 1, pp. 1–10, 2019, doi: 10.1148/radiol.2019190088.

[22]   Y. R. Park, Y. J. Kim, W. Ju, K. Nam, S. Kim, and K. G. Kim, "Classification of cervical cancer using deep learning and machine learning approach," *Res. Sq.*, pp. 1–26, 2021.

[23]   L. Nanni, S. Ghidoni, and S. Brahnam, "Handcrafted vs. non-handcrafted features for computer vision classification," *Pattern Recognit.*, vol. 71, pp. 158–172, 2017, doi: 10.1016/j.patcog.2017.05.025.

[24]   L. Cai, J. Gao, and D. Zhao, "A review of the application of deep learning in medical image classification and segmentation," vol. 8, no. 4, 2020, doi: 10.21037/atm.2020.02.44.

[25]   T. G. Debelee, S. R. Kebede, F. Schwenker, and Z. M. Shewarega, "Deep Learning in Selected Cancers' Image Analysis—A Survey," *J. Imaging*, vol. 6, no. 11, pp. 1–40, 2020, doi: 10.3390/jimaging6110121.

[26]   R. Girshick, "Fast R-CNN," *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 1440–1448, 2015, doi: 10.1109/ICCV.2015.169.

[27]   J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 779–788, 2016, doi: 10.1109/CVPR.2016.91.

[28]   X. W. Zhong-Qiu Zhao, Peng Zheng, Shou-tao Xu, "Object Detection with Deep Learning :

A Review," *Comput. Vis. Pattern Recognit.*, pp. 1–21, 2019, [Online]. Available: https://arxiv.org/abs/1807.05511.

[29]   A. Z. Karen Simonyan, "Very Deep Convolutonal Networks for Large-Scale Image Recognition," *Comput. Vis. Pattern Recognit.*, pp. 1–14, 2015, [Online]. Available: https://arxiv.org/abs/1409.1556.

[30]   H. A. Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," *Comput. Vis. Pattern Recognit.*, 2017.

[31]   J. S. Kaiming He, Xiangyu Zhang, Shaoqing Ren, "Deep Residual Learning for Image Recognition," *Comput. Vis. Pattern Recognit.*, 2015, [Online]. Available: https://arxiv.org/abs/1512.03385.

[32]   M. Tan and Q. V Le, "EfficientNet : Rethinking Model Scaling for Convolutional Neural Networks," *Int. Conf. Mach. Learn.*, 2020, [Online]. Available: https://arxiv.org/abs/1905.11946.

[33]   J. Wang, H. Zhu, S. H. Wang, and Y. D. Zhang, "A Review of Deep Learning on Medical Image Analysis," *Mob. Networks Appl.*, vol. 26, no. 1, pp. 351–380, 2021, doi: 10.1007/s11036-020-01672-7.

[34]   A. Mamta, D. Sanjeev, and S. Kulvinder, "Deep Learning in Health Care: Automatic Cervix Image Classification Using Convolutional Neural Network," in *Mobile Radio Communications and 5G Networks*, N. Marriwala, C. . Tripathi, D. Kumar, and S. Jain, Eds. Springer, Singapore, 2020, pp. 145–151.

[35]   C. Sy, Stuart and Y. H. Asawa, "Deep Learning Approaches for Determining Optimal Cervical Cancer Treatment," *Stanford Univ. Rep.*, pp. 1–8, 2017.

[36]   S. D. and K. S. Mamta Arora, "Exploring Deep Convolution Neural Networks with Transfer Learning for Transformation Zone Type Prediction in Cervical Cancer," in *Soft Computing: Theories and Applications*, Springer, Singapore, 2020, pp. 1127–1138.

[37]   D. A. Cruz, E. G. , Carmen Villar-Patiño, and   and M. Martinez-Alanis, "Cervix Type Classification Using Convolutional Neural Networks," in *VIII Latin American Conference on Biomedical Engineering and XLII National Conference on Biomedical Engineering*, vol. 75, 2020, pp. 377–384.

[38]   L. Lei, R. Xiong, and H. Zhong, "Identifying Cervix Types using Deep Convolutional Networks," *Stanford Univ. Rep.*, 2017.

[39]   N. Kaur and N. Panigrahi, "Automated Cervical Cancer Screening Using Transfer Learning," vol. 6, no. 8, pp. 2110–2119, 2017.

[40]   A. M. P. Jayaraj, M. Bijoy, "Segmentation Based Preprocessing Techniques for Predicting the Cervix Type Using Neural Networks," in *Computational Vision and Bio-Inspired Computing*, S. Smys, J. M. R. S. Tavares, V. E. Balas, and A. M. Iliyasu, Eds. 2020, pp. 717–726.

[41]   X. Zhang and S.-G. Zhao, "Cervical image classification based on image segmentation preprocessing and a CapsNet network model," *Int J Imaging syst Technol*, pp. 1–10, 2018, doi: 10.1002/ima.22291.

[42]   L. Ghouti, "A Fully-Automated Deep Learning Pipeline for Cervical Cancer Classification," *Expert Syst. Appl.*, vol. 141, 2019, doi: 10.1016/j.eswa.2019.112951.

[43]   P. Guo *et al.*, "Nuclei-Based Features for Uterine Cervical Cancer Histology Image Analysis with Fusion- based Classification," *IEEE J Biomed Heal. Inf.*, vol. 20, no. 6, pp. 1595–1607, 2015, doi: 10.1109/JBHI.2015.2483318.

[44]   H. A. Almubarak, R. J. Stanley, R. Long, S. Antani, R. Zuna, and S. R. Frazier, "Convolutional Neural Network Based Localized Classification of Uterine Cervical Cancer Digital Histology Images .," *Procedia Comput. Sci.*, vol. 114, pp. 281–287, 2017, doi: 10.1016/j.procs.2017.09.044.

[45]   H. A. Almubarak *et al.*, "A Hybrid Deep Learning and Handcrafted Feature Approach for Cervical Cancer Digital Histology Image Classification," vol. 14, no. 2, pp. 66–87, 2019, doi: 10.4018/IJHISI.2019040105.

[46] S. Sornapudi *et al.*, "Cervical Whole Slide Histology Image Analysis Toolbox," 2020, doi: 10.1101/2020.07.22.20160366.

[47] L. Wei, Q. Gan, and T. Ji, "Cervical cancer histology image identification method based on texture and lesion area features," *Innov. Biomed. Sci. Eng.*, vol. 22, no. 1, pp. 186–199, doi: 10.1080/24699322.2017.1389397.

[48] M. Wu, C. Yan, H. Liu, Q. Liu, and Y. Yin, "Automatic classification of cervical cancer from cytological images by using convolutional neural network," *Natl. Libr. Med.*, vol. 38, no. 6, pp. 1–9.

[49] Y. Tian *et al.*, "Computer-aided Detection of Squamous Carcinoma of the Cervix in Whole Slide Images," *Comput. Vis. Pattern Recognit.*, [Online]. Available: https://arxiv.org/abs/1905.10959.

[50] "Intel & MobileODT Cervical Cancer Screening | Kaggle." https://www.kaggle.com/c/intel-mobileodt-cervical-cancer-screening (accessed Jan. 08, 2021).

[51] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," 2018, [Online]. Available: http://arxiv.org/abs/1804.02767.

[52] H. Zhao *et al.*, "Mixed YOLOv3-LITE: A lightweight real-time object detection method," *Sensors (Switzerland)*, vol. 20, no. 7, 2020, doi: 10.3390/s20071861.

[53] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 4510–4520, 2019, doi: 10.1109/CVPR.2018.00474.

[54] K. Dong, C. Zhou, Y. Ruan, and Y. Li, "MobileNetV2 Model for Image Classification," *Proc. - 2020 2nd Int. Conf. Inf. Technol. Comput. Appl. ITCA 2020*, pp. 476–480, 2020, doi: 10.1109/ITCA52113.2020.00106.

[55] H. B. Moss, D. S. Leslie, and R. Paul, "Using J - K -fold Cross Validation to Reduce Variance When Tuning NLP Models," *Proc. 27th Int. Conf. Comput. Linguist.*, pp. 2978–

2989, 2018.

[56] P. Weill and J. W. Ross, "Acheiving Strong Regularisation For Deep Neural Networks," pp. 1–13, 2018.

[57] T. J. B.Lakshmanan, S.Anan, "Stain Removal Through Color Normalization of Haematoxylin and Eosin Images : A Review," *J. Phys. confrence Ser.*, vol. 1362, no. 012108, pp. 1–12, 2019, doi: 10.1088/1742-6596/1362/1/012108.

[58] O. Marques, *Practical Image and Video Processing Using MATLAB*. Singapore: John Wiley & Sons, Inc., Hoboken, New Jersey, 2011.

[59] T. Saito and M. Rehmsmeier, "The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets," *PLoS One*, vol. 10, no. 3, pp. 1–21, 2015, doi: 10.1371/journal.pone.0118432.

Appendix

Appendix A

Clinical setup for cervix type dataset collection

Appendix B

Materials used during histopathological image

a) Amsope camera   **b) Camera mounted Microscope**

Appendix C

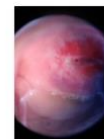Sample of some annotation of the cervix