Early Detection of Cervical Cancer Using Deep Neural Networks

by

Atoshi Akhund
18201199
Saad Ahmad
18101226
Sarwar Siddiqui Taki
18101193

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

Department of Computer Science and Engineering
BRAC University
May 2022

Declaration

It is hereby declared that

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

Student's Full Name & Signature:

Atoshi Akhund
18201199

Saad Ahmad
18101226

Sarwar Siddiqui Taki
18101193

Approval

The thesis/project titled "Early Detection of Cervical Cancer Using Deep Neural Networks" submitted by

1. Atoshi Akhund (18201199)
2. Saad Ahmad (18101226)
3. Sarwar Siddiqui Taki (18101193)

Of Spring, 2022 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on Date.

Examining Committee:

Supervisor:
(Member)

Md. Ashraful Alam, PhD
Assistant Professor
Department of Computer Science and Engineering
BRAC University

Co-Supervisor:
(Member)

Md. Tanzim Reza
Lecturer
Department of Computer Science and Engineering
BRAC University

Head of Department:
(Chair)

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

Abstract

Cervical cancer is a disease that is mostly preventable, but it is one of the major causes of cancer fatality in women worldwide. Several studies say that annually 2,60,000 women die because of cervical cancer. Chronic infections with "high-risk (HR)" human papillomavirus are the leading cause of cervical cancer (HPV). Regular cervical cancer screening, on the other hand, can help to prevent this dangerous disease. Cervical cancer screening is a procedure for detecting precancerous and cancer in women who are at risk, and it is recommended for all women aged 30 to 49. Cervical cancer can be avoided if precancerous lesions are detected and treated early. Nowadays, several tests are performed to detect cervical cancer, most of whom are time consuming and expensive. In this paper, we are approaching the development of a fast and effective system to detect cervical cancer from the cervix image in a minimum time with better accuracy using deep neural networks. First, we collected image data and classified them using VGG16, VGG19, InceptionV3, ResNet50 and ResNet101. From our result we got an accuracy rate of 88.48% from VGG16, 88.97% from VGG19, 88.09% from InceptionV3, 88.67% from ResNet50 and 89.06% from ResNet101. Then, using a mixture of classifiers with the greatest accuracy, we created ensemble models with the best overall accuracy rate of 94.20 percent for CERVIXEN V1, 95.01 percent for CERVIXEN V2, and 94.69 percent for CERVIXEN V3.

Keywords: Cervical cancer, HPV, precancerous lesions, cervix image, Deep Neural Network, VGG, ResNet, Inception.

Acknowledgement

First and foremost, we express our gratitude to Almighty Allah for allowing us to complete our Thesis on time and without any impediments.
As a result, we would like to convey our heartfelt gratitude to Dr. Md. Ashraful Alam, our excellent instructor and supervisor, for his steadfast support and persistent supervision, which helped us to complete our project.
We owe a debt of appreciation to Md. Tanzim Reza, our co-supervisor, for willingly assisting us with this thesis to the best of his capacity.
Last but not least, we would like to express our sincere appreciation to Mr. Shakib M. Dipto, CVIS lab Research Assistant, for his unwavering assistance throughout this thesis.

Table of Contents

Declaration i

Approval ii

Abstract iii

Acknowledgment iv

Table of Contents v

List of Figures vii

List of Tables viii

Nomenclature ix

1 Introduction 1
- 1.1 Motivation . . . 1
- 1.2 Aims and Objectives . . . 1
- 1.3 History of Cervical Cancer . . . 1
- 1.4 Research Methodology . . . 2
- 1.5 Research Orientation . . . 2

2 Related Work 3
- 2.1 Pap Test and Liquid Based Cytology (LBC) . . . 4
- 2.2 HPV and HPV Testing . . . 5
- 2.3 Visual inspection with acetic acid (VIA) and Lugol's iodine (VILI) . . . 6
- 2.4 Machine learning in cervical cancer . . . 7

3 Research Methodology 8
- 3.1 Our Methodology . . . 8
- 3.2 Used Architecture . . . 9
  - 3.2.1 Convolutional Neural Network (CNN) . . . 9
  - 3.2.2 VGG . . . 9
  - 3.2.3 Inception . . . 11
  - 3.2.4 ResNet . . . 13
  - 3.2.5 Convolutional Layer . . . 16

4 Implementation **21**
4.1 Dataset . . . 21
4.1.1 Source . . . 21
4.1.2 Sample Data . . . 21
4.1.3 Data Classification . . . 22
4.2 Data Pre-processing . . . 22
4.2.1 Resizing Images . . . 22
4.2.2 Normalization and Scaling Images . . . 22
4.2.3 Data Augmentation . . . 23
4.2.4 Data Scrubbing . . . 23
4.3 Architecture Training . . . 23

5 Result and Analysis **24**
5.1 Individual Model . . . 24
5.2 Cervical Cancer Analysis by Ensemble modeling Version 1 (CERVIXEN V1) . . . 25
5.3 Cervical Cancer Analysis by Ensemble modeling Version 2 (CERVIXEN V2) . . . 26
5.4 Cervical Cancer Analysis by Ensemble modeling Version 3 (CERVIXEN V3) . . . 27
5.5 CERVIXEN V1 vs CERVIXEN V2 vs CERVIXEN V3 . . . 28

6 Conclusion and Future Work **29**
6.1 Conclusion and Future Work . . . 29

Bibliography **33**

List of Figures

2.1 The life-course approach for cervical cancer prevention and control [38] 3

3.1 Workflow Diagram 8
3.2 VGG16 Architecture 10
3.3 VGG19 Architecture 10
3.4 VGG19 Layer 11
3.5 Inception Naive Model 12
3.6 Inception V3 Architecture 13
3.7 Outline of Inception V3 13
3.8 Comparison of 20-layer vs 56-layer architecture 14
3.9 ResNet Architecture 14
3.10 ReLU Shortcut 15
3.11 ResNet 50 shortcut 16
3.12 Error rates (%) of single-model results 16
3.13 Activation Function 17
3.14 ReLU-activation-function 17
3.15 Softmax formula 18
3.16 basic structure of an ensemble model 19
3.17 Output layers of the ensemble model hierarchy for CERVIXEN V1 . . 19
3.18 Output layers of the ensemble model hierarchy for CERVIXEN V2 . . 20
3.19 Output layers of the ensemble model hierarchy for CERVIXEN V3 . . 20

4.1 Sample Data 21

5.1 Comparison between used models 24
5.2 Accuracy curve of CERVIXEN V1 25
5.3 Validation loss curve of CERVIXEN V1 25
5.4 Accuracy curve of CERVIXEN V2 26
5.5 Validation loss curve of CERVIXEN V2 26
5.6 Accuracy curve of CERVIXEN V3 27
5.7 Validation loss curve of CERVIXEN V3 27
5.8 Comparison between our proposed models 28

List of Tables

5.1 Comparison between used models 24
5.2 Comparison between our proposed models 28

Nomenclature

The next list describes several symbols & abbreviation that will be later used within the body of the document

CIN Cervical Intraepithelial Neoplasia

CNN Convolutional Neural Network

$DICOM$ Digital Imaging and Communications in Medicine

HPV Human Papillomavirus

LBC Liquid Based Cytology

$LMIC$ Low and Middle Income Countries

PCA Principal Component Analysis

$ReLU$ Rectified Linear Unit

VGG Visual Geometry Group

VIA Visual inspection with Acetic Acid

$VILI$ Visual inspection with Lugol's Iodine

Chapter 1

Introduction

1.1 Motivation

Cervical cancer was one of the top two cancers affecting women worldwide in the early 2000s, particularly in developing nations [15]. Cervical cancer is the second most common type of cancer among women in Bangladesh, with 8068 cases and 5214 fatalities in 2018 [56]. Cervical cancer mortality was once high in affluent countries, but it was reduced by assuring frequent cervical cancer screening for women aged 30 to 49 years [30]. Cervical cancer can be avoided with regular screening and early detection, but the process is complicated, time-consuming, and costly. Cervical cancer screening is not very common in developing countries like Bangladesh due to the difficulty and cost of diagnosis. As a result of the lack of screening facilities to detect it early, the death rate for women's malignancies remains high. To accomplish the fastest binary classification of cervical cancer, we want to design a method that will identify the types of cervical cancer from infected persons' cervical colposcopic images by comparing them to endo-cervix and ecto-cervix images.

1.2 Aims and Objectives

The aim is to develop a deep neural network model that uses a colposcopic image of the cervix to predict cervical cancer outcomes. Our major objective is to discover the types of cervical cancer early and identify the underlying characteristics so that it can be diagnosed. Understanding the importance of features might lead to improved results in future studies.

1.3 History of Cervical Cancer

In the early 1900s, cervical cancer was the leading cause of mortality among women in the United States. This cancer develops in the tissues of the cervix, which connects the uterus to the vagina. Later in 1927-29, George Papanicolaou and Aurel Babes inspected cervical cells to detect cervical cancer and invented a screening procedure which is named as Pap-test. This was the first cervical cancer screening test that became extensively used. Until 1975, when German virologist Harald zur Hausen published his idea that HPV (human papillomavirus) is involved in cervical cancer, this test was the main method to diagnose cervical cancer. In 1980, he discovered

HPV-16 and HPV-18 in cervical cancer samples, leading to the conclusion that these two strains are responsible for nearly 70% of cervical cancer globally. The International Agency for Research on Cancer has determined that HPV-16 and HPV-18 are unquestionably carcinogenic and specifically cause cervical cancer, 15 years after Hausen's discovery [10]. The link between HPV and cervical cancer is stronger than the link between lung cancer and smoking. Sexual intercourse is how the virus spreads. Even though most women's bodies are capable of fighting HPV infection, the virus can sometimes overcome resistance, leading to cancer. Cervical cancer is a slow-growing malignancy that can go unnoticed at first.Later on, this could result in pelvic pain or vaginal bleeding. Normal cervix cells often take several years to transform into cancer cells. However, studies and doctors say that regular Pap tests, in which cells are taken from the cervix and examined under a microscope, help prevent cervical cancer. Cervical cancer rates are gradually declining in developed countries when screening is done on a regular basis. In poor countries, the situation is different; low and middle-income countries account for more than 85 percent of cervical cancer mortality. Large disparities in access to medical treatment, particularly screening, are to blame for this scenario. Most cervical cancer screening programs have depended on a large percentage of women being screened repeatedly with cytology, being referred to colposcopy, and being treated for cervical cancer precursors. However, in underdeveloped countries, such plans have proven extremely difficult to implement. Unless current approaches to the disease change, cervical cancer incidence and mortality are predicted to climb by around 20% every ten years; the majority of this increase will affect women in low- and middle-income nations [10].

1.4 Research Methodology

Initially, we observed VGG16, VGG19, InceptionV3, ResNet50, and ResNet101 among the Machine Learning models. In the first stage, we pre-processed the colposcopic image data into a well-defined form of 224x224x3. Then we used those machine learning models to train and evaluate them. Following the evaluation, we selected the classifiers which gave better results and used them to create various ensembles. Finally, we illustrate the classification after comparing the performance of the ensembles.

1.5 Research Orientation

In chapter 2, we discussed earlier research done by other researchers in the same field as ours. Then, in Chapter 3, we went over each method, convolution layer, and activation function we used in our study. The application of our thesis work is explained once again in Chapter 4. We also demonstrated how our datasets are distributed, followed by a description of the pre-processing procedures utilized on the dataset in question. We displayed the results we acquired after implementing the algorithms in Chapter 5 and then went over the results analysis. Finally, in Chapter 6, we discussed the conclusion as well as your research's future plans.

Chapter 2

Related Work

Cervical cancer is a form of cancer which develops inside the cells of the cervix, which attaches the uterus to the vaginal canal. This is the second most frequent cancer in women in the world [12]. Cervical cancer normally takes a long time to develop. The tissues of the cervix undergo a process defined as dysplasia, inside which aberrant cells grow in the cervical tissue, as cancer begins. After that Later on, cancer cells start to grow and expand further into the cervix as well as surrounding cells [12]. Many forms of screening have been developed over time to avoid cervical cancer. Screening is a treatment that checks for cancer before any symptoms appear. This can help with early diagnosis of cancer. It may be simpler to cure cancer or abnormal tissue if it is detected early. Cancer may have progressed by the time symptoms occur [41]. In affluent nations, the emergence of well organized, cytology-based cervical cancer screening, as well as the quick acceptance of positive screening and prior lesion treatments, has dramatically decreased cervical cancer incidence and death. Nevertheless, because of critical problems such as insufficient equipment and quality control procedures, a low sensitivity to precancerous lesions, and poor reliability, these initiatives have been difficult to implement in Low and Middle Income Countries (LMICs). As a result, a low-cost early detection approach for this malignancy would undoubtedly be beneficial. According to [47], The World Health Assembly's global strategy for eradicating cervical cancer as a public health issue, adopted in 2020, calls for a comprehensive strategy to prevent and control cervical cancer. Interventions throughout the lifespan are among the suggested acts.

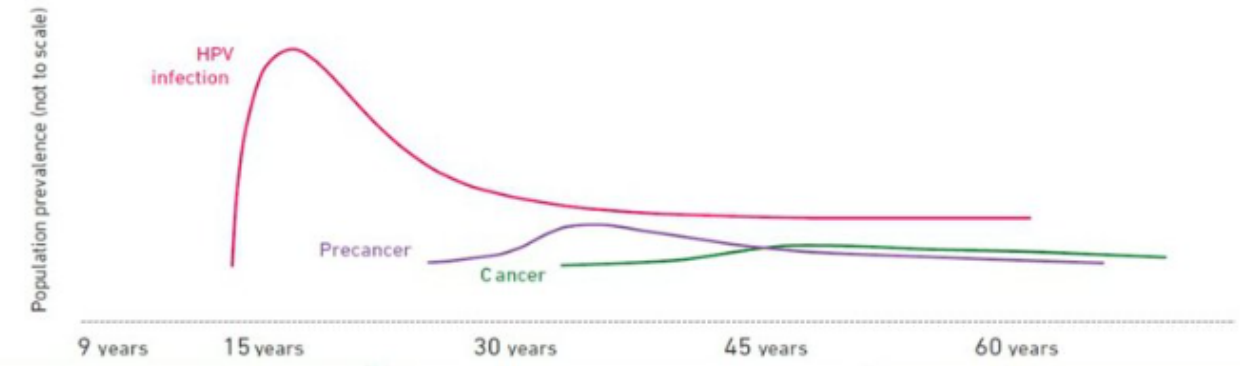

Figure 2.1: The life-course approach for cervical cancer prevention and control [38]

The primary prevention of cervical cancer is the HPV vaccination. The HPV vaccine was initially released in the United States in 2006. In 2006, the FDA approved Merck's first-generation Gardasil®, which prevented infection by four HPV strains 6, 11, 16, and 18. Gardasil®9 received FDA approval in December 2014. This vaccine defends against HPV strains 31, 33, 45, 52, and 58, as well as the 4 strains approved in the prior Gardasil vaccine. It was also the sole vaccine available in the United States until 2017 [45]. By 2016 HPV vaccine was introduced in Bangladesh. Girls between 9-13 years are the eligible category of this vaccine who yet not become sexually active. There are two doses of vaccine within 6 months of time [12]. There are still many women in developing and underdeveloped countries who do not have access to vaccination services. Even the vaccination cannot guarantee that cervical cancer will not develop. And the vaccination has only been around for a short time, and there is an age limit for it as well. So screening is currently the most prevalent method of cervical cancer prevention. This is referred to as secondary prevention of cervical cancer.

Moreover 3 types of screening are available to detect cervical cancer. They are:

1. Conventional (Pap) and liquid based cytology (LBC) [53].
2. HPV testing for high-risk (hr) HPV types [53].
3. Visual inspection with acetic acid (VIA) and Lugol's iodine (VILI) [53].

2.1 Pap Test and Liquid Based Cytology (LBC)

Cervical cancer occurrence and death have really been declining in advanced nations since the advent of Pap test in 1940s, which permitted morphological abnormalities in the cervical epithelium to be discovered swiftly [6]. This is the most basic way for detecting cervical cancer. The Pap test, introduced in the 1940s, produced a reliable and low-complexity cervical cancer screening tool [20]. The Papanicolaou test, also recognized as the Pap smear, was invented by George Nicholas Papanicolaou and revolutionized the early identification of cervical cancer [16]. Non-experimental research, notably observational and literature reviews, provide evidence for adequacy of cervical cytology in decreasing the occurrence of severe illness and deaths [1]. The finest prospective studies have compared cervical cancer fatality rates in five Northern nations in between screening programs were implemented [1] [4] [2]. The death rate was lowered from 8% to 73 percent with the advent of cytologic screening in two time periods, 1963-1967 and 1978-1982, according to studies [41]. In affluent nations, the emergence of well organized, cytology-based cervical cancer screening, as well as the quick acceptance of positive screening and prior lesion treatments, has dramatically decreased cervical cancer incidence and death. Nevertheless, because of critical problems such as insufficient equipment and quality control procedures, a low sensitivity to precancerous lesions, and poor reliability, these initiatives have been difficult to implement in Low and Middle Income Countries (LMICs). As a result, a low-cost early detection approach for this malignancy would undoubtedly be beneficial.

This procedure involves exfoliating cervical tissues, which also are subsequently preserved, examined under a microscope, then morphologically analyzed. The staining method devised for this test offered a polychromatic characterization of the nucleus and cytoplasm characteristics. The Pap test can determine whether necrotic has developed, the level of tissue degeneration found, and the development of squamous epithelium by evaluating alterations in nuclear chromatin [13] [20]. In 1990, Liquids Based Cytology was created as a cytotechnology advancement. Similar to a normal Papanicolaou smear, liquid-based test collecting entails sample and cell transfer to a liquid medium, followed by processing and analysis [35]. In the United States, the majority of tests are now performed using liquid based cytology [8]. Though LBC offers some advantages over traditional Papanicolaou smears, including enhanced cell collection and processing, blood and debris filtration, and less unacceptable findings, it has no difference in sensitivity or specificity. The specificity of the Pap smear is roughly 98 percent, which is promising [5]. In terms of sensitivity, however, it produces lower results, ranging between 55 and 80 percent [5].It is not the sole dependable solution because the sensitivity has not been promising. Although affluent nations may achieve success with good setup and frequent monitoring, it is difficult to apply in developing countries without adequate resources. Enough power for Microscopes, equipment for screening, and experienced cytopathologists to analyze the data are required for cytology. Moreover, for better results, cytology must be performed on a regular basis, which is not always achievable. The sensitivity of its results varies among geographies. As a result, the pap smear's efficacy is insufficient.

2.2 HPV and HPV Testing

Harald zur Hausen defied conventional wisdom and claimed that the oncogenic human papillomavirus (HPV) was behind cervical cancer, the second commonest malignancy in women. He found that HPV-DNA might exist in a resting form in tumors and could be detected using particular viral DNA searches. He discovered HPV to be a diverse viral family. Only a few kinds of HPV cause cancer. His finding has resulted in the classification of HPV infection's natural history, a knowledge of HPV-induced carcinogenesis mechanisms, and the creation of anti-HPV vaccinations.Harald zur Hausen proposed a role for human papillomavirus (HPV) in cervical cancer in the 1970s, contrary to popular belief at the time. If the tumor cells harbored an oncogenic virus, he anticipated that viral DNA would be integrated into their genomes. By precisely scanning cancer cells for such viral DNA, the HPV genes that promote cell proliferation should be identifiable. Harald zur Hausen spent more than a decade looking for distinct HPV strains, a search that was complicated by the fact that only sections of the viral DNA were merged into the host genome. He discovered the new, tumorigenic HPV16 type in 1983 after discovering unique HPV-DNA in cervix cancer tissues. He synthesized HPV16 and HPV18 from cervical cancer patients in 1984. In over 70% of cervical cancer biopsies around the world, HPV types 16 and 18 were regularly identified [57].
The HPV test is a cervical cancer screening test, but it does not determine if you have cancer. Rather, the test looks for HPV, the virus that causes cervical cancer, in your body. This test is not recommended before the age of 30 for women in general. Primary HPV testing will most probably become the standard treatment

for cervical cancer screening for both low and high income countries in the future decade. This shift comes as evidence emerges that HPV testing is more sensitive than other methods for detecting high-grade precancerous diseases. Furthermore, when compared to negative "Pap smears" over numerous rounds of screening, negative HPV testing has shown a decreased prevalence rate of cervical intraepithelial neoplasia (CIN) grade 3 or worse pathology (CIN3+). While many nations have started pilot programs to move to primary HPV screening, and others have entirely transitioned for particular uses, there is still much to be refined regarding this technique for cancer control, as indicated by the variety of ways it is used [26].
HPV testing is more sensitive than the pap test and is age-independent. HPV test specificity is lower than pap test and increases with age. Pap test results are examined under the microscope which surely can have any human error. On the other hand, in the HPV test the samples are examined under some highly accurate laboratory instruments that test the high risk HPV. Both HPV/Pap cotesting and HPV testing alone (also known as primary HPV testing) are more delicate than Pap testing alone in persons aged 30 and older who are receiving routine screening. As a result, someone who has a negative HPV test plus a normal Pap test—or just a negative HPV test—has a very low probability of acquiring precancerous cervical lesions in the coming years. Because of this, the recommended screening interval for Pap and HPV cotesting or primary HPV testing is 5 years: this longer interval (compared to 3 years for people receiving Pap testing alone) allows abnormalities to be detected in time to be treated while lowering the identification of HPV infections which would be successfully governed by the immune system. Pap and HPV cotesting, as well as primary HPV testing, may help detect glandular cell abnormalities, such as cervical cancer. Because Pap testing isn't as good at detecting adenocarcinoma and glandular cell abnormalities as it is at detecting squamous cell abnormalities and malignancies [55].

2.3 Visual inspection with acetic acid (VIA) and Lugol's iodine (VILI)

Researchers have established novel affordable diagnostic tools suitable for large-scale screening of cervical abnormalities, prompted by the need for efficient strategies for cervical cancer screening and based on the concepts that the large percentage of pre-invasive and invasive cervical lesions are visible by 'naked-eye' observation [7]. Visual inspection with acetic acid (VIA) and Lugol's iodine (VILI) are two low-cost screening techniques commonly employed in low-resource settings, with VIA being the more popular. These procedures are based on the fact that when acetic acid or Lugol's iodine is applied directly to the cervix, precancerous cervical lesions become visible to the naked eye by both clinicians and non-clinicians. In low-resource contexts, both VIA and VILI have been observed to have acceptable specificity and sensitivity, but not optimal [23] [14]. The only difference between Visual Inspection with Acetic Acid (VIA) and Visual Inspection with Lugol's Iodine (VILI) is the solution used to enhance the cervical lesions.

2.4 Machine learning in cervical cancer

Cervical cancer can be successfully treated if detected early via screenings. Screenings, on the other hand, are time-consuming for patients. Furthermore, due to the limited hospital resources and the vast populations who require screenings, screening tests are inefficient. A typical screening procedure cannot handle a large number of patients at once. Furthermore, the Pap test can be highly subjective and highly dependent on the clinicians' experience. After all, human decisions contain inaccuracies. Machine learning makes it more successful than traditional diagnosis procedures in resolving these challenges in the healthcare industry [51] [43]. Various methodologies following standard ML approaches, such as k-nearest neighbors (KNN), K-means clustering, and RF, have been used for cervical cancer diagnosis in previous studies [43]. WEN WU1 and HAO created a clinical decision support network for malignant development in 2017, combining a knowledge-based approach with a preliminary collection of hypotheses hereditary genetic algorithms simulations in a soft computing model [22] [43]. Reif et al. [9] reported RF effectiveness in a variety of model genomic and proteomic datasets in 2006. Based on genomic data and proteomics datasets, RF progress fails to categorize associated features [43]. Alyafeai and Ghouti [33] aim to introduce a fully integrated cervical cancer detection and screening pipeline based on cervical imaging in 2020 [43]. By automating the cervical cancer diagnosis procedure from Pap-drug photographs, William et al. [31] were able to reduce the likelihood of error. Image enhancement was done with local adaptive histogram equalization [43]. Wang et al. [49] worked in wireless medical sensor networks (WMSN) in 2021 to address physical security and over-centralized server issues . Similarly, Xiong et al. [50] proposed a blockchain-based ECDSA for blockchain-enabled IoMT with a fault-tolerant batch verification mechanism. Besides, Khamparia et al. [42] used a convolutional network and a variational encoder to classify data in 2021.

Chapter 3

Research Methodology

3.1 Our Methodology

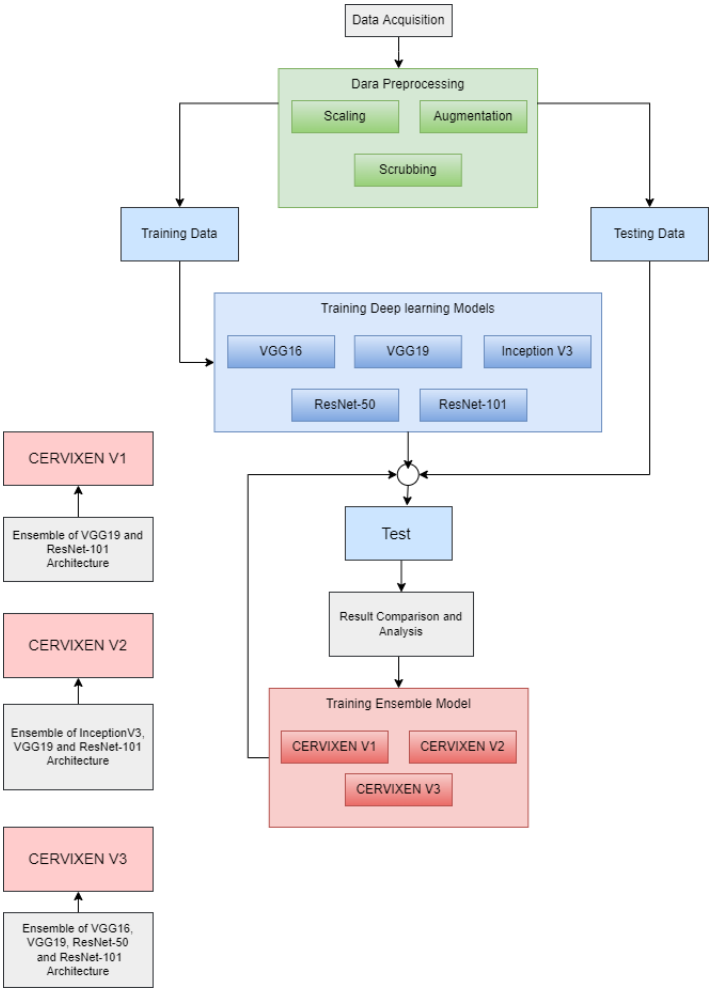

Figure 3.1: Workflow Diagram

The work flow diagram in figure 3.1 depicts an overview of each and every phase of our proposed model, from data collecting towards result extraction. To begin, we gathered our essential image data from free sources and preprocessed it with augmentation, scaling, and scrubbing. After that, we divided our dataset into training

and testing data and trained it using deep learning models (VGG16, VGG19, Inception V3, ResNet-50, and ResNet-101). We tested our dataset after finishing the training. In the next stage, We incorporated our proposed models, which were generated using a combination of existing deep learning models. To train the dataset and evaluate the data, we applied CERVIXEN V1 (Ensemble of VGG19 and ResNet-101), CERVIXEN V2 (Ensemble of Inception V3, VGG19 and ResNet-101), and CERVIXEN V3 (Ensemble of VGG16, VGG19, ResNet-50 and ResNet-101).

3.2 Used Architecture

3.2.1 Convolutional Neural Network (CNN)

Convolutional neural networks is a type of neural network which is mostly used for visual imaging [34]. Input, output, and a number of hidden layers make up a CNN. In addition to hidden layers, CNN uses convolution, pooling, and fully linked layers in its construction. The image is used as an input and convolution is used to calculate the width and height parameters. In addition, hyper parameters such as input and output channels have been finished. The convolution layer convolved the input before sending it to the next layer.

3.2.2 VGG

VGG (Visual Geometry Group) is a type of CNN architecture which is mainly used for detection and identification purposes [34]. VGG is a straightforward model consisting of many convolution layers, max-pooling layers, completely connected layers with the ReLU activation function, and the softmax activation function in the final output layer. In VGG architecture, the input to the convolutional layer is fixed to 224 X 224 X 3 image which is transferred to the multiple convolutional layers with a fixed dimension of 3 X 3.

VGG16

VGG16 is a convolutional neural network (CNN) architecture that won the 2014 ILSVRC (Imagenet) competition. It is regarded as one of the best vision model architectures ever created. VGG16 is unique in that instead of having a huge number of hyper-parameters, they focused on having 3x3 filter convolution layers with a stride 1 and use the same padding and maxpool level of 2x2 filter stride 2. All through architecture, the convolution and max pool layers are arranged in the same way. Ultimately, there are two FC (completely connected layers) and a softmax for output. The 16 in VGG16 alludes to the fact that it contains 16 layers with different weights. This network is quite huge, with approximately 138 million (estimated) parameters. [39]

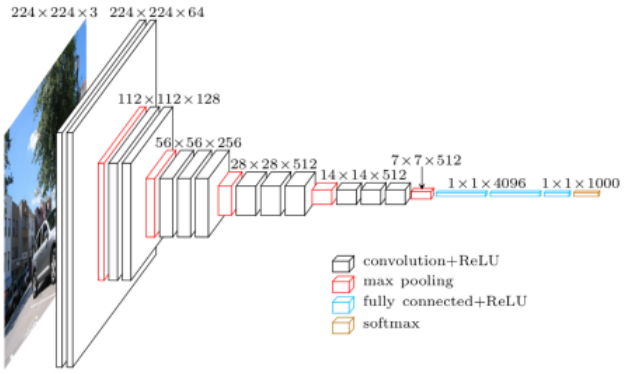

Figure 3.2: VGG16 Architecture

VGG19

In layman's terms, VGG is a deep CNN that is used to classify images. This network was given a fixed size (224 × 224) RGB picture as input, indicating that such matrix was of shape (224,224,3). The sole preprocessing was to subtract the mean RGB value from every pixel, which was calculated for the whole training set. They used kernels with a size of (3 × 3) and a stride size of 1 pixel to cover the entire image concept. To keep the image's spatial resolution, spatial padding was applied. With stride 2, maximum pooling was achieved over a 2 × 2 pixel window. This was followed by the Rectified linear unit (ReLu) to bring non-linearity into the model to improve classification and reduce processing time, while earlier models employed tanh or sigmoid functions.
Three fully connected layers were implemented, the first two of which were of size 4096, followed by a layer with 1000 channels for 1000-way ILSVRC classification, and finally a softmax function. [46]

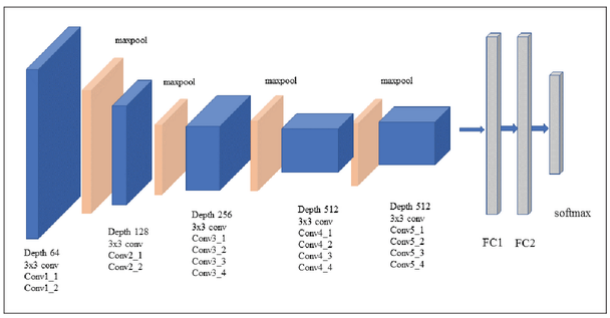

Figure 3.3: VGG19 Architecture

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 **LRN**	conv3-64 **conv3-64**	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 **conv3-128**	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 **conv1-256**	conv3-256 conv3-256 **conv3-256**	conv3-256 conv3-256 conv3-256 **conv3-256**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Figure 3.4: VGG19 Layer

3.2.3 Inception

An inception network is a deep neural network with a recurring architectural pattern known as Inception modules. Main Concepts behind Inception architecture are

1. Deep neural networks with high performance must be huge. A neural network required to have several additional layers and units inside these levels in order to be called big.

2. Extracting features at different sizes is advantageous to convolutional neural networks. The natural human visual brain works by recognizing patterns at various sizes, which combine to generate larger object experiences. As a result, multi-scale networks have the ability to learn even more.

3. The Hebbian Principle is taken into account: neurons that pulse together connect together.

The following components make up an Inception Module:

- Input layer
- 1x1 convolution layer
- 3x3 convolution layer
- 5x5 convolution layer
- Max pooling layer
- Concatenation layer

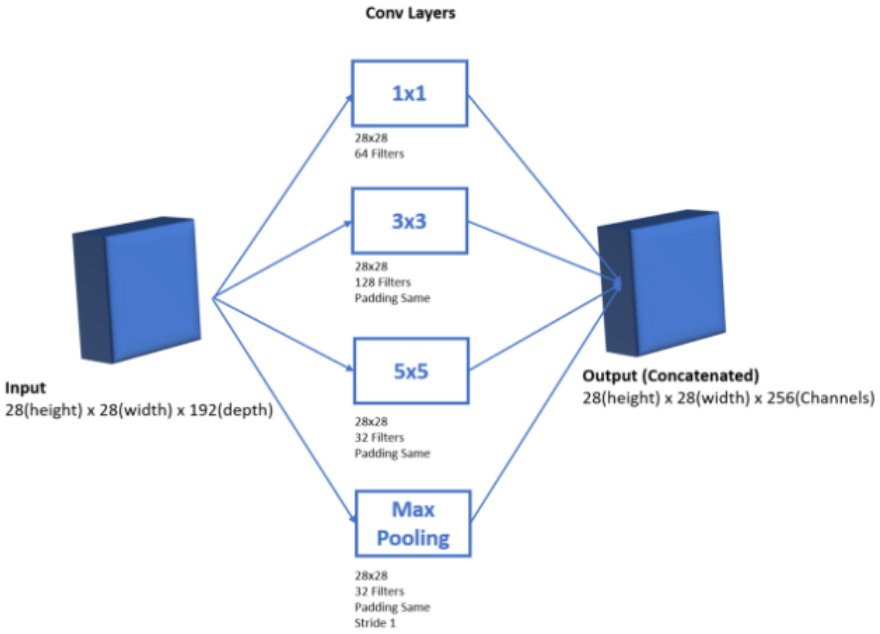

Figure 3.5: Inception Naive Model

Inception eliminates the necessity for such judgments by utilizing several filter sizes of 1x1, 3x3, and 5x5.

- 1x1 learns patterns through out input's depth.
- 3x3 and 5x5 discover spatial patterns throughout the input's three dimensions (height, breadth, and depth)

When all of the patterns learned from the various filter sizes are combined, the representational power increases. [32]

Inception V3

According to [19], Inception-v3 is an extended version of the popular GoogLeNet [18], which has demonstrated strong classification performance in a variety of biomedical applications by employing transfer learning. Inception-v3 suggested an inception model that concatenates multiple different sized convolutional filters into a new filter, similar to GoogLeNet. This architecture reduces the amount of parameters that must be taught, lowering computational complexity.
The Inception v3 model, which was launched in 2015, features 42 layers and a reduced margin of error than its predecessors. The Inception V3 model has undergone significant changes. Such as:

1. Factorization into Smaller Convolutions
2. Spatial Factorization into Asymmetric Convolutions
3. tility of Auxiliary Classifiers
4. Efficient Grid Size Reduction

The finished Inception V3 version looks like this after all of the optimizations

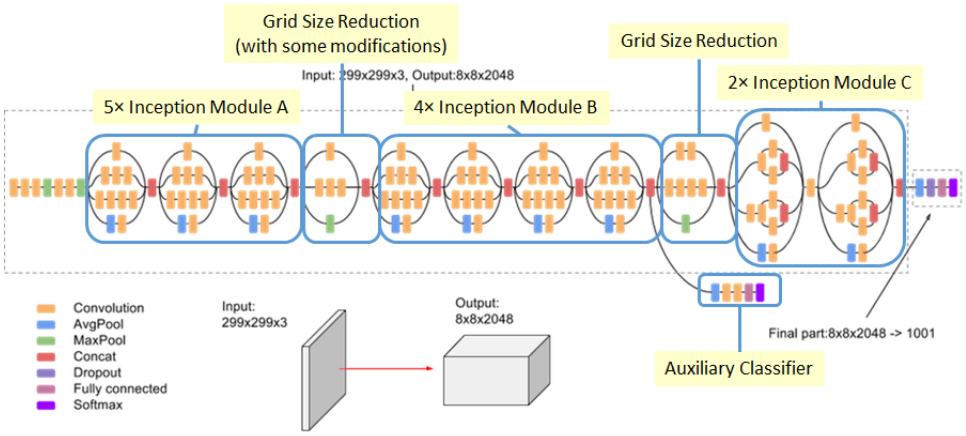

Figure 3.6: Inception V3 Architecture

The inception V3 model has 42 layers in total, which is somewhat more than the preceding inception V1 and V2 models. However, this model's efficiency is truly remarkable. We'll get to it in a minute, but first, let's take a look at the components that make up the Inception V3 model. [48]

TYPE	PATCH / STRIDE SIZE	INPUT SIZE
Conv	3×3/2	299×299×3
Conv	3×3/1	149×149×32
Conv padded	3×3/1	147×147×32
Pool	3×3/2	147×147×64
Conv	3×3/1	73×73×64
Conv	3×3/2	71×71×80
Conv	3×3/1	35×35×192
3 × Inception	Module 1	35×35×288
5 × Inception	Module 2	17×17×768
2 × Inception	Module 3	8×8×1280
Pool	8 × 8	8 × 8 × 2048
Linear	Logits	1 × 1 × 2048
Softmax	Classifier	1 × 1 × 1000

Figure 3.7: Outline of Inception V3

3.2.4 ResNet

Every consecutive winning design uses more layers in a deep neural network to minimize the error rate during the first CNN-based architecture (AlexNet) won the ImageNet 2012 competition. This works for a small number of layers, but as the number of layers grows, a typical problem in deep learning called Vanishing/Exploding gradient emerges. As a result, the gradient becomes 0 or too huge. As a result, as the number of layers increases, so does the testing data error rate.[54]

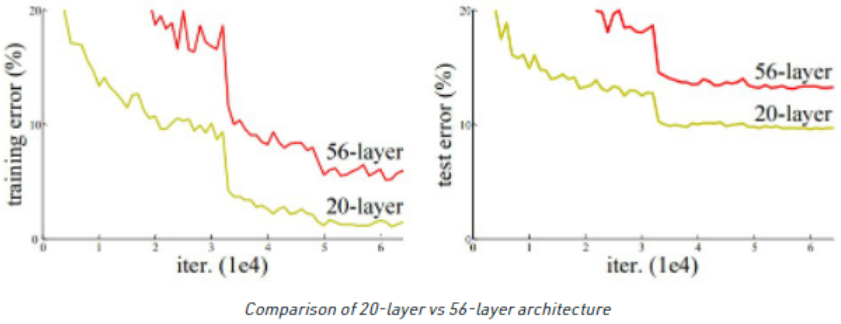

Figure 3.8: Comparison of 20-layer vs 56-layer architecture

In the graph above, we can see that a 56-layer CNN has a greater error rate on both the training and testing datasets than a 20-layer CNN architecture. If this was due to overfitting, the 56-layer CNN should have a lower training error, but it also has a higher training error. The authors came to the conclusion that the vanishing/exploding gradient is to blame for the high error rate.
ResNet, a new architecture presented by Microsoft Research in 2015, established a new architecture called Residual Network.
Residual Block: This architecture introduces the concept of the Residual Network to solve the challenges of the vanishing/exploding gradient. We employ a method called skip connections in this network. The skip connection bypasses a few stages of training and links the output port.

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
		3×3 max pool, stride 2				
conv2_x	56×56	[3×3, 64; 3×3, 64] ×2	[3×3, 64; 3×3, 64] ×3	[1×1, 64; 3×3, 64; 1×1, 256] ×3	[1×1, 64; 3×3, 64; 1×1, 256] ×3	[1×1, 64; 3×3, 64; 1×1, 256] ×3
conv3_x	28×28	[3×3, 128; 3×3, 128] ×2	[3×3, 128; 3×3, 128] ×4	[1×1, 128; 3×3, 128; 1×1, 512] ×4	[1×1, 128; 3×3, 128; 1×1, 512] ×4	[1×1, 128; 3×3, 128; 1×1, 512] ×8
conv4_x	14×14	[3×3, 256; 3×3, 256] ×2	[3×3, 256; 3×3, 256] ×6	[1×1, 256; 3×3, 256; 1×1, 1024] ×6	[1×1, 256; 3×3, 256; 1×1, 1024] ×23	[1×1, 256; 3×3, 256; 1×1, 1024] ×36
conv5_x	7×7	[3×3, 512; 3×3, 512] ×2	[3×3, 512; 3×3, 512] ×3	[1×1, 512; 3×3, 512; 1×1, 2048] ×3	[1×1, 512; 3×3, 512; 1×1, 2048] ×3	[1×1, 512; 3×3, 512; 1×1, 2048] ×3
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

Figure 3.9: ResNet Architecture

Instead of layers learning the underlying mapping, this network allows the network to fit the residual mapping. Instead of using H(x) as the initial mapping, use F(x):= H(x) – x, which provides H(x):= F(x) + x.

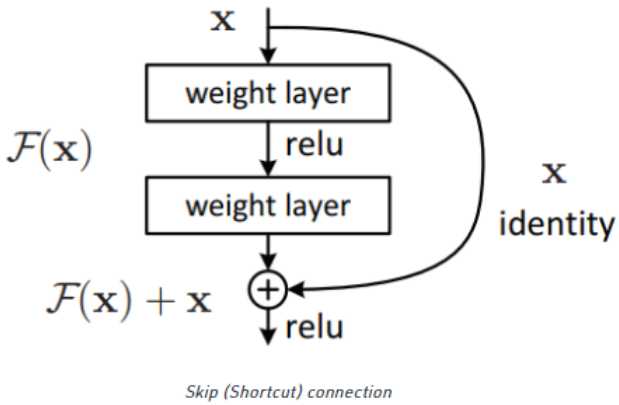

Figure 3.10: ReLU Shortcut

The benefit of including this kind of skip connection is that any layer that degrades architecture performance will be bypassed by regularization. As a result, very deep neural networks can be trained without the issues caused by vanishing/exploding gradients. On the CIFAR-10 dataset, the authors explored 100-1000 layers.
A related concept is known as "highway networks," and these networks use skip connections as well. These skip connections, like LSTM, make use of parametric gates. The amount of data that travels through to the skip connection is determined by these gates. However, this architecture does not deliver superior accuracy than ResNet architecture.

ResNet 50

We get one layer from a convolution with a kernel size of 7×7 and 64 distinct kernels, all with a stride of size 2. Following that, we have max pooling with a stride size of 2.
Following there comes a 1×1,64 kernel, followed by a 3×3,64 kernel, and finally a 1×1,256 kernel. These three layers are repeated three times in total, giving us nine levels in this stage. Next, we see a kernel of 1×1,128, followed by a kernel of 3×3,128 and finally a kernel of 1×1,512. This phase was done four times, totaling 12 layers. Then there's a 1×1,256 kernel, followed by 3×3,256 and 1×1,1024 kernels, which are repeated six times for a total of 18 layers.Then a 1×1,512 kernel was added, followed by two more 3×3,512 and 1×1,2048 kernels, for a total of nine layers. After that, we do an average pool and finish with a fully linked layer with 1000 nodes and a softmax function, giving us one layer. The activation functions and the max/average pooling layers are not counted.
In conclusion, we have a Deep Convolutional network with 1 + 9 + 12 + 18 + 9 + 1 = 50 layers. For the ResNet 50 and higher, a little adjustment was made in that previously, shortcuts connections skipped two layers, but now they skip three levels, and 1×1 convolution layer was added, which we will go over in depth with the ResNet 50 Architecture.[37]

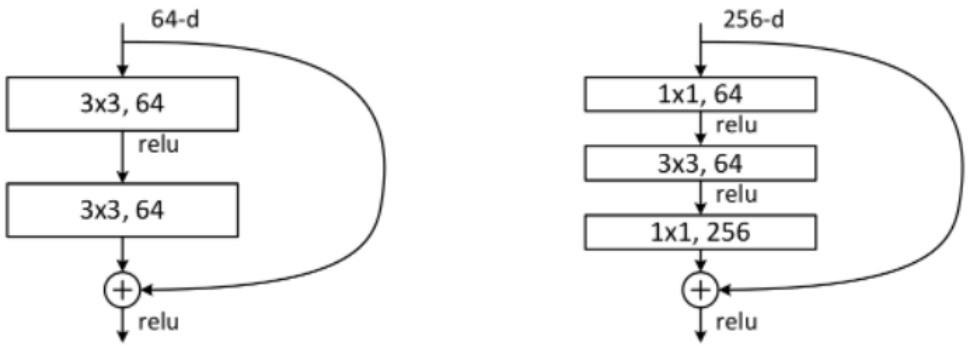

Figure 3.11: ResNet 50 shortcut

ResNet 101

ResNet-101 is a 101-layer deep convolutional neural network. The ImageNet database contains a pre - trained deep variant of the network that has been trained on over a million photos. The network can classify photos into 1000 different object categories, including keyboards, mice, pencils, and a variety of animals. As a consequence, the network has learnt about various of rich feature descriptions for a variety of images. The network's picture input size is 224×224 pixels. They use more 3-layer blocks to build ResNets with 101 and 152 layers (above table). The 152-layer ResNet (11.3 billion FLOPs) has less complexity than VGG-16/19 nets (15.3/19.6 billion FLOPs) even after the depth is raised. [27]

method	top-1 err.	top-5 err.
VGG [41] (ILSVRC'14)	-	8.43†
GoogLeNet [44] (ILSVRC'14)	-	7.89
VGG [41] (v5)	24.4	7.1
PReLU-net [13]	21.59	5.71
BN-inception [16]	21.99	5.81
ResNet-34 B	21.84	5.71
ResNet-34 C	21.53	5.60
ResNet-50	20.74	5.25
ResNet-101	19.87	4.60
ResNet-152	**19.38**	**4.49**

Figure 3.12: Error rates (%) of single-model results

3.2.5 Convolutional Layer

A CNN's fundamental building block is a convolutional layer. It contains a set of filters (or kernels) whose parameters must be learned during the course of the training. The filters are typically smaller than the original image. Each filter develops an activation map by combining with the image. The convolution layer's output

volume is created by stacking the activation maps of each filter along the depth dimension. Every component of the activation map can be viewed as a neuron's output. As a result, each neuron is connected to a small local region in the input image, the size of which is equal to the filter's size. An activation map's neurons share parameters as well. Because of the convolutional layer's local connection, the network is driven to train filters that have the best response to a specific region of the input.[44]

Activation Function

The activation function is a component that is placed at the end or in the middle of a Neural Network. They aid in determining whether or not the neuron will fire.[29]

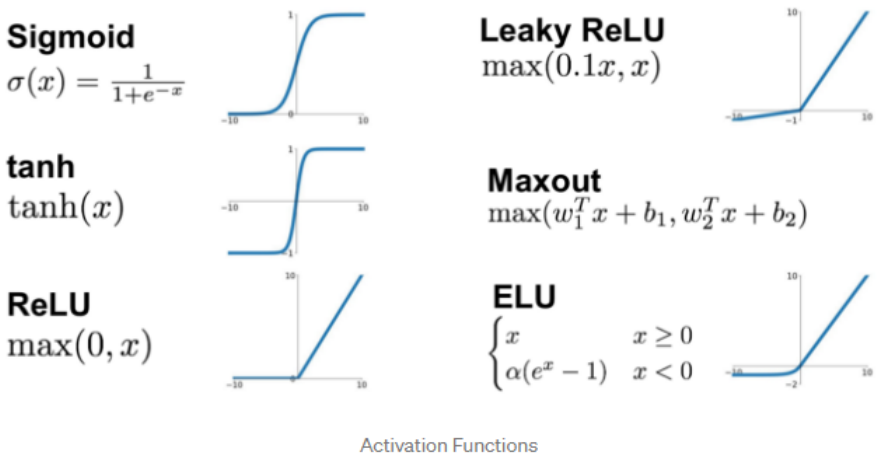

Figure 3.13: Activation Function

Rectified Linear Unit (ReLU)

In today's neural networks, the ReLU function has been the most extensively utilized activation function. One of the most significant advantages of ReLU above other activation functions is it doesn't simultaneously stimulate all neurons. We can see from the above figure that the ReLU function transforms all negatives inputs to zero and does not activate the neuron. Because just a few neurons are stimulated at a time, it is incredibly computationally efficient. The positive zone does not saturate. ReLU converges six times quicker than the tanh and sigmoid activation functions in practice.

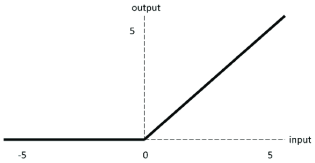

Figure 3.14: ReLU-activation-function

Softmax

The softmax function reduces a vector of K true values to a vector of K true values that add up to 1. The softmax turns the input values, which might be positive, negative, zero, or higher than one, into values between 0 and 1, allowing them to be understood as probabilities. If one of the inputs is tiny or negative, the softmax converts it to a small probability; if one of the inputs is high, it converts it to a large probability; nonetheless, it will always be between 0 and 1. The softmax formula is:

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$$

Figure 3.15: Softmax formula

Where all of the zi values are input vector elements and thus can accept any real value. The normalization factor at the bottom of the calculation guarantees that all of the function's output values add up to 1, resulting in a legitimate probability distribution.[28]

Optimizer: Adam

Adam is an alternate optimization approach that uses repeated cycles of "adaptive moment estimation" to produce better efficient neural network weights. Adam uses stochastic gradient descent to solve non-convex problems more easily and with less resources than many other optimization algorithms. It works best in very large data sets since the gradients are kept "tighter" across many learning iterations.
Adam combines the benefits of Adaptive Gradients and Root Mean Square Propagation, two additional stochastic gradient approaches, to offer a novel learning methodology for optimizing a variety of neural networks.[28]

Ensemble Modeling

Ensemble learning is a broad meta technique to machine learning that combines predictions from different models to improve predictive performance.
Although you may create an apparently infinite number of ensembles for any predictive modeling issue, the subject of ensemble learning is dominated by three approaches. So much so that, rather than algorithms themselves, each has produced a plethora of specialized ways.
Bagging, stacking, and boosting are the three primary types of ensemble learning methods, and it's critical to have a thorough grasp of each and to use them in your computational modeling project.
[40]

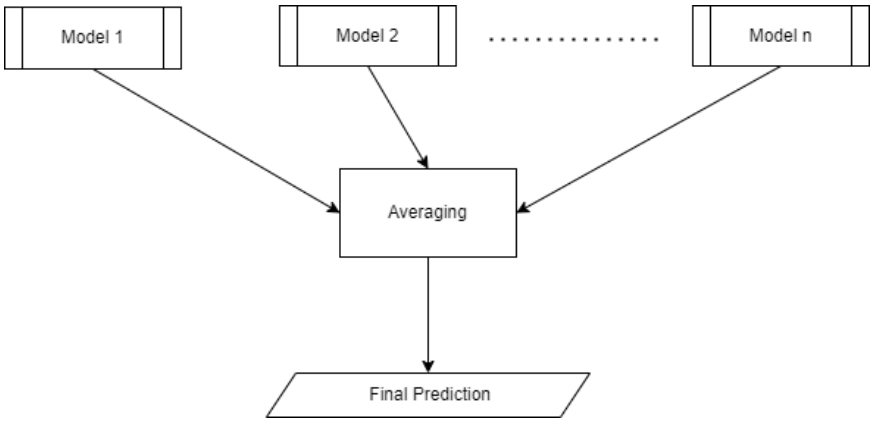

Figure 3.16: basic structure of an ensemble model

Averaging Layer

We used the fusion of VGG19 and ResNet-101 architectures in our proposed "CERVIXEN V1" model, the fusion of Inception V3, VGG16, and ResNet-101 architectures in "CERVIXEN V2", and the fusion of VGG19, VGG16, ResNet-50, and ResNet-101 architectures in "CERVIXEN V3," and the fusion of VGG19, VGG16, ResNet-50, and ResNet The output probability from various architectures with the same output dimension will be fed into the averaging layer. It will also calculate an average probability for the two labels: Ecto and Endo cervical cancer.

Output Layer

Our proposed "CERVIXEN V1", "CERVIXEN V2", and "CERVIXEN V3" forecasts either the sample picture indicated to Endo cervical cancer or Ecto cervical cancer based on colposcopic image to the model of the individual, depending on the output probability of the averaging layer. Our proposed model will output the average of all models employed in our proposed architecture in the averaging layer.

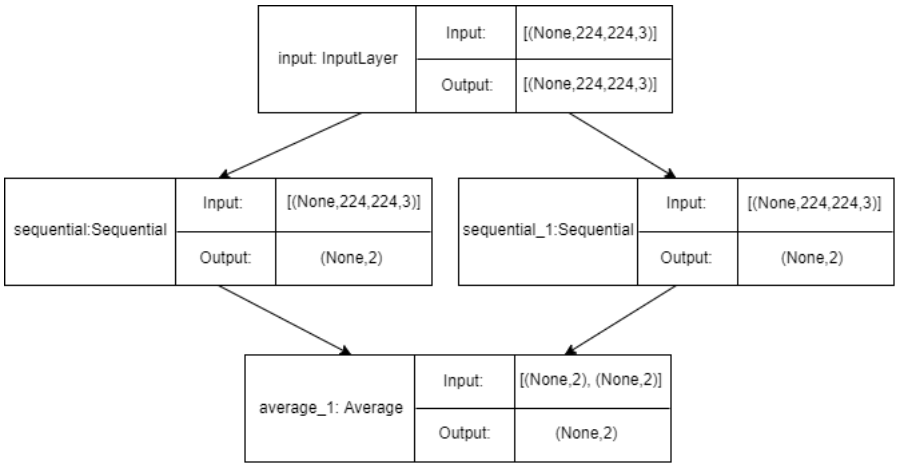

Figure 3.17: Output layers of the ensemble model hierarchy for CERVIXEN V1

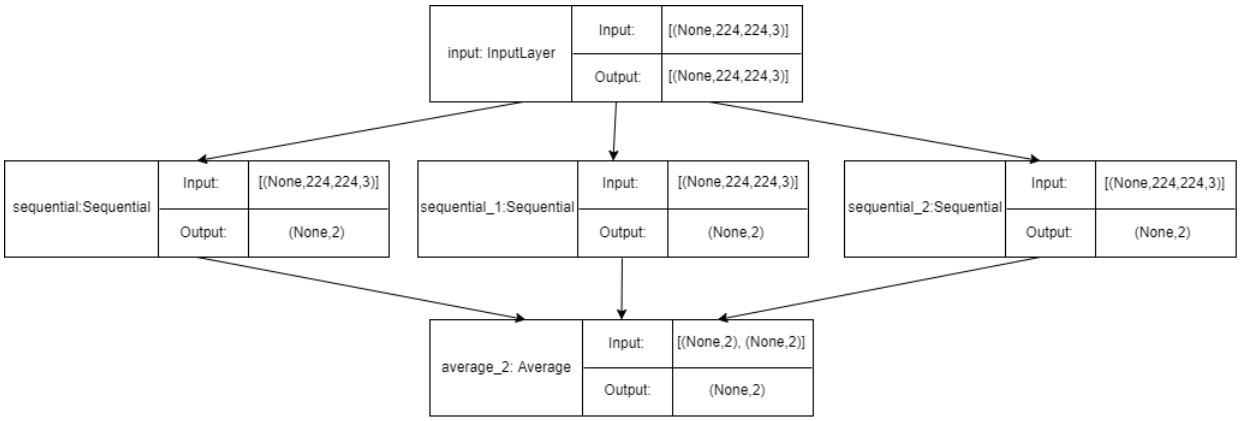

Figure 3.18: Output layers of the ensemble model hierarchy for CERVIXEN V2

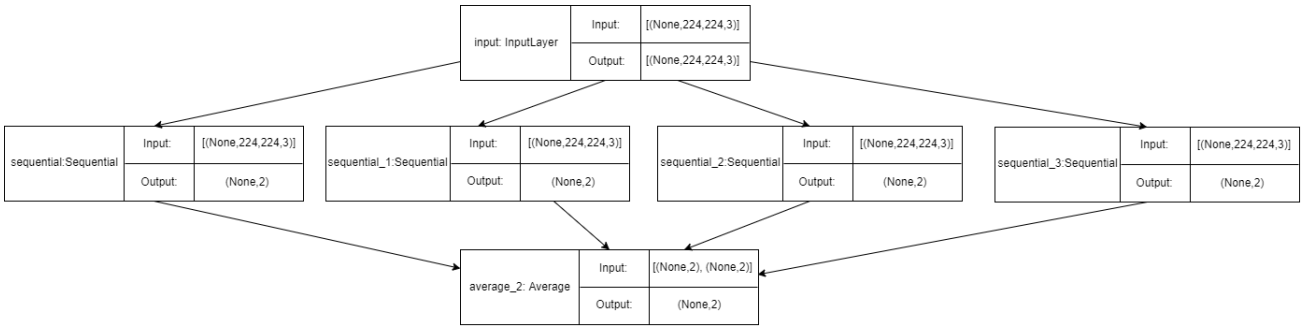

Figure 3.19: Output layers of the ensemble model hierarchy for CERVIXEN V3

Chapter 4

Implementation

4.1 Dataset

4.1.1 Source

Colposcopic image collections: Constructed in [52].
Cervical Colposcopy images: Constructed in [36].

4.1.2 Sample Data

Figure 4.1: Sample Data

4.1.3 Data Classification

For the study, the dataset has been splitted into an 8:2 ratio, with 80% of the dataset utilized to train the model and 20% used for the test set.

Train Dataset

The machine learning algorithm method is given a labeled sampling dataset with response or output labels in this stage.

Test Dataset

In some cases, when the algorithm converges to enhance efficiency and a sequence of samples is utilized for real world research, this might acquire particular properties of the training set. Decent results for an unknown test group will increase trust that now the algorithm can also provide accurate response in the actual world.

4.2 Data Pre-processing

4.2.1 Resizing Images

This model was created particularly to get the best accurate diagnosis of Cervical cancer colposcopy pictures. Therefore, as shown, some models were categorized and trained individually. Endocervix and ectocervix colposcopy pictures are included in the collection, with endocervix referring to the inner section of the cervix and ectocervix referring to the segment of the cervix that extends into the vagina. Every picture is scaled to a specified size of 224x224 for network training. TensorFlow [17] will thus be utilized for this model. The ImageDataGenerator [21] [24] class in Keras has been used to adjust pixel values in image datasets for modeling. This method collects the images data for creation, verification, or evaluation, and restores images to the algorithm in batches as well as performs scaling processes as required. When modeling utilizing neural networks, this provides a robust and rational technique to scale visual input. The ImageData Generator handles a variety of feature selection techniques and also pixel scaling options based on percentages. The ImageData Generator methods provides a connection to leveling since it mostly utilizes the mean determined from the training data as feature-wise centering. Stats must also be verified before regression on the training dataset.

4.2.2 Normalization and Scaling Images

A multivariate approach for examining a data table is Principal component analysis (pca) [3] [11] which occurrences are defined by a set of quantified response variables that are mostly interconnected. To correct every colposcopy picture projections, the far more significant eigen flat fields are linearly merged. Tests show that as compared to the usual flat field corrections, the recommended dynamic flat field correction lowers systematic errors in projection intensity normalization by a significant amount. This would be accomplished using Keras' ImageDataGenerator method. The technique of reducing data to a 0-1 scale is known as normalization.

This may be done by changing the rescale inputs to a ratio that each pixel will multiply. The required range could be obtained by translating the rescale parameter to a ratio that could be multiplied from each pixel.

4.2.3 Data Augmentation

The dataset was supplemented with a series of different adjustments to the photographs to boost the effectiveness of the model. To improve the worth of the dataset, some of the photos were resized. Some photographs were additionally cleaned because they were not clear enough. Moreover, several images have been rotated (90, 180, and 270 degrees), translated [25], and flipped vertically and horizontally. The dataset was created from two different data sources. To enhance and improve the dataset, it was marged. For all this, Keras' ImageDataGenerator class is used. The Keras deep learning neural network framework's ImageDataGenerator class enables fitting models with image data augmentation. To enhance the amount of training data and enhance the model's generalization ability, image data augmentation is necessary. Analytically, the photos inside the data are not identified. The model only gives improved photos as an alternative. For training motives, randomized augmentation of pictures enables the generation of modified and near copy pictures. A second ImageDataGenerator instance is mostly used for data augmentation, that might or might not have the very same pixel scaling option as the ImageDataGenerator instance used during training data. The technique of shifting all the pixels of an image for one direction, either horizontally and vertically, while keeping the image proportions the same is known as "image shifting." The ImageDataGenerator constructor's width and height shifting range arguments control the amount of horizontal and vertical shifting in a sequential way.

4.2.4 Data Scrubbing

We gathered the dataset from several sources, and there were many images in the dataset that were not clear at all. Some of the images were blurry and others were cropped. Moreover, several images exhibited an excessive amount of noise. As a result, we had to remove certain images from our training dataset in order to enhance our total dataset.

4.3 Architecture Training

The initial stage in training a Neural Network is determining the appropriate weights for the Neural Connections. The data was separated into two sections required to train neural networks. "Train" will be the first section. The 80 percent data set which was used to fit the model. It was used by VGG16, VGG19, ResNet50, ResNet101, and InceptionV3. These neural network topologies were run using a batch size of 16 and an epoch of 30. The remaining of the data is for "Test," and it accounts for 20% of the total. The following configuration parameters were used to run the dataset: CPU Intel Core i5 8500, 16GB ram, and GPU NVIDIA GTX1070 Ti 8GB.

Chapter 5

Result and Analysis

5.1 Individual Model

Architecture	val_categorical_accuracy	val_loss
VGG16	88.48%	0.2580
VGG19	88.97%	0.8186
Inception V3	88.09%	0.8558
ResNet-50	88.67%	1.0417
ResNet-101	89.06%	1.4071

Table 5.1: Comparison between used models

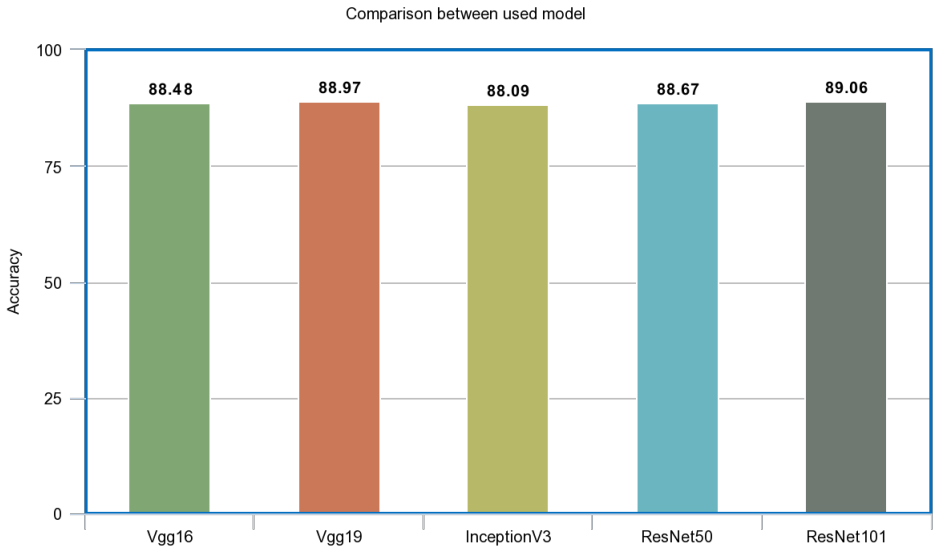

Figure 5.1: Comparison between used models

According to 5.1, VGG19, ResNet50, and ResNet101 have val_catagorical_accuracy of 88.97%, 88.67%, and 89.06%, respectively.These three architectures demonstrated the highest accuracy rate. For the following implementation, two, three, and four types of architecture from distinct models will be integrated to form an ensemble.

5.2 Cervical Cancer Analysis by Ensemble modeling Version 1 (CERVIXEN V1)

The result for the initial ensemble modeling, named "CERVIXEN V1," were obtained by combining these two architectures, VGG19 and ResNet101. In the VGG and ResNet models, both VGG19 and ResNet101 architectures provided the highest accuracy rate. The results are presented in the graph below.

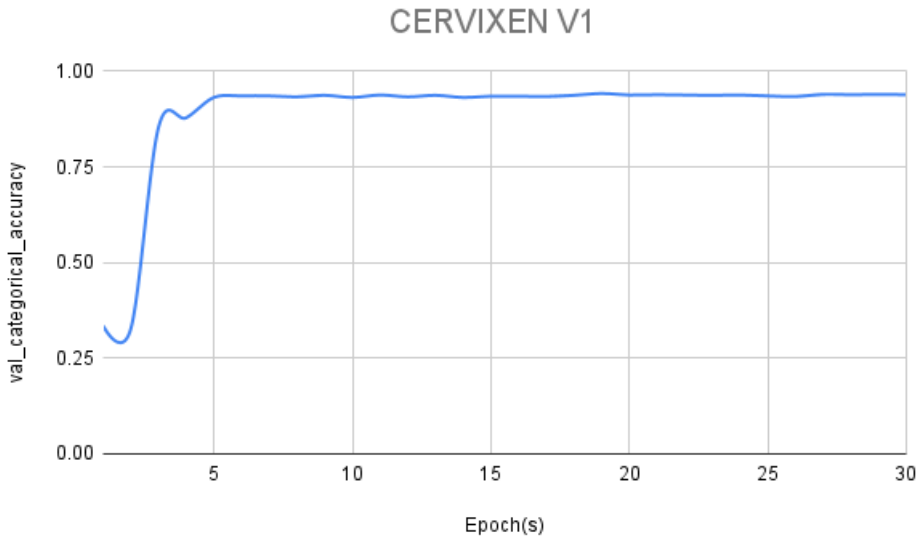

Figure 5.2: Accuracy curve of CERVIXEN V1

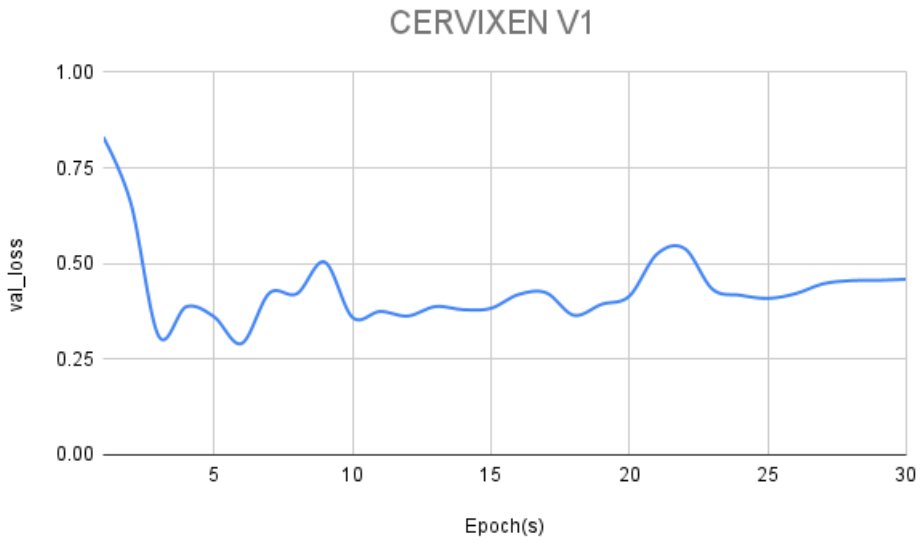

Figure 5.3: Validation loss curve of CERVIXEN V1

CERVIXEN V1 has reached 94.20% validation accuracy based on its learning curves.

5.3 Cervical Cancer Analysis by Ensemble modeling Version 2 (CERVIXEN V2)

"CERVIXEN V2" was ensembled by combining the best findings from each model, such as vgg19 from VGG, ResNet101 from ResNet, and inceptionv3 from Inception. The results are presented in the graph below.

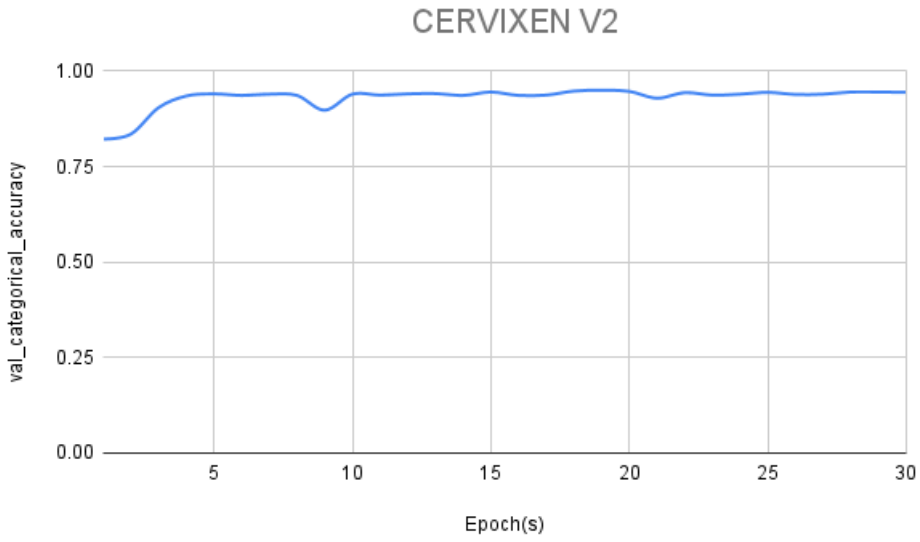

Figure 5.4: Accuracy curve of CERVIXEN V2

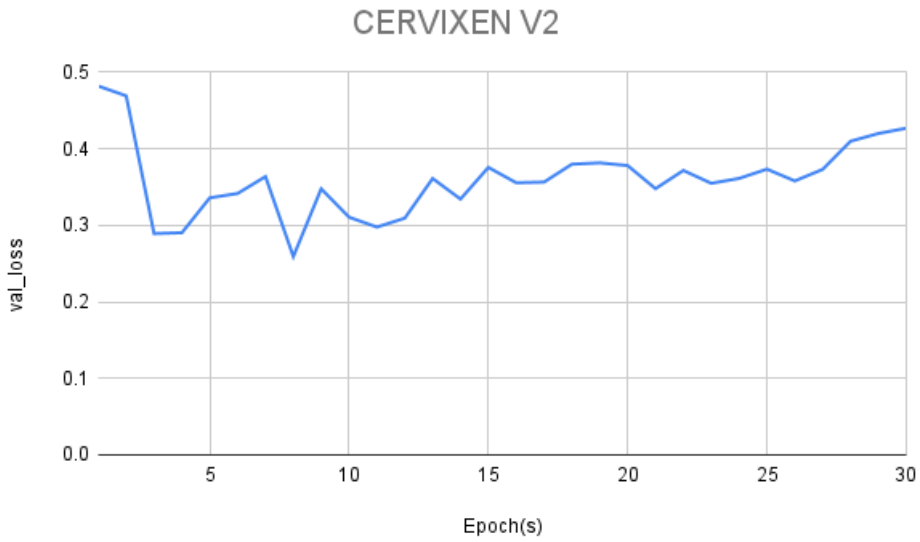

Figure 5.5: Validation loss curve of CERVIXEN V2

"CERVIXEN V2" has reached 95.01 percent validation accuracy based on its learning curves.

5.4 Cervical Cancer Analysis by Ensemble modeling Version 3 (CERVIXEN V3)

"CERVIXEN V3" was ensembled by only taking VGG and ResNet models. Vgg16, vgg19, ResNet50 and ResNet101. The results are presented in the graph below.

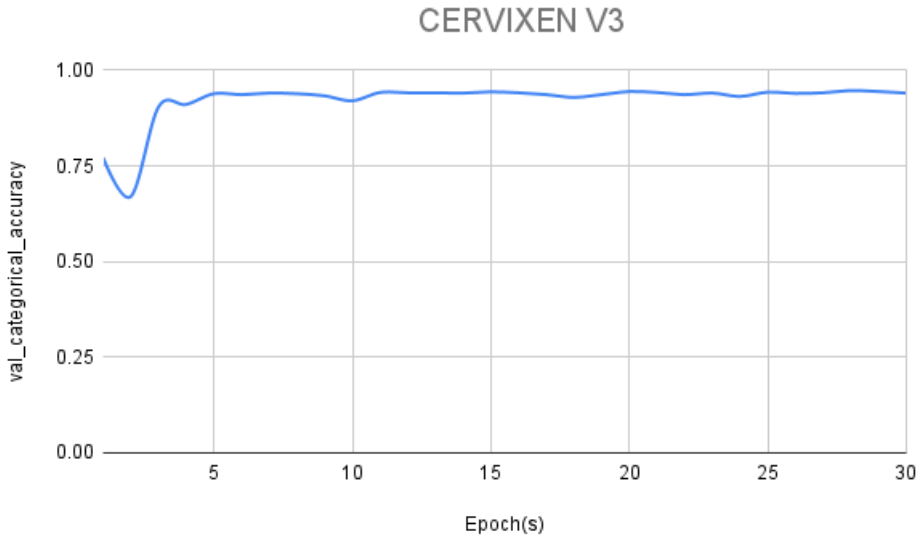

Figure 5.6: Accuracy curve of CERVIXEN V3

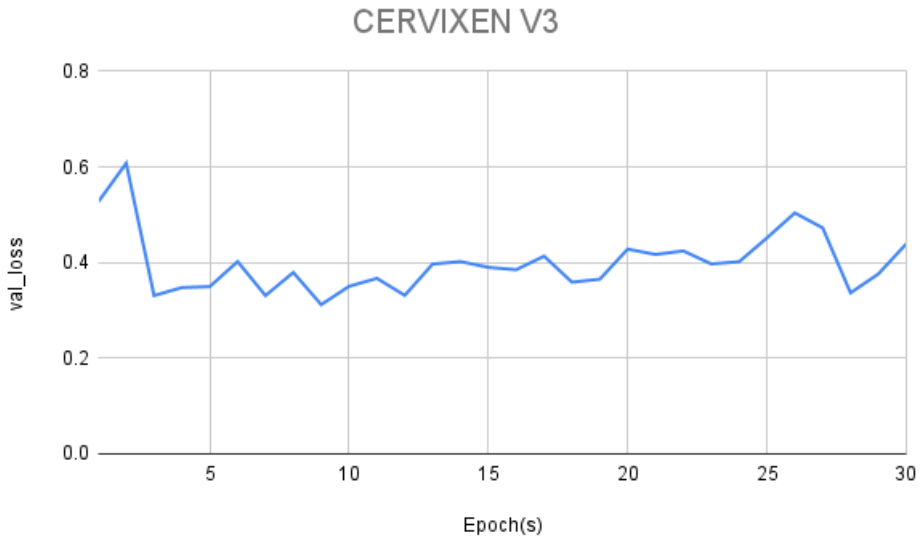

Figure 5.7: Validation loss curve of CERVIXEN V3

"CERVIXEN V3" has reached 94.69 percent validation accuracy based on its learning curves.

5.5 CERVIXEN V1 vs CERVIXEN V2 vs CERVIXEN V3

Model	val_categorical_accuracy	val_loss
CERVIXEN V1	94.20%	0.3936
CERVIXEN V2	95.01%	0.3816
CERVIXEN V3	94.69%	0.3367

Table 5.2: Comparison between our proposed models

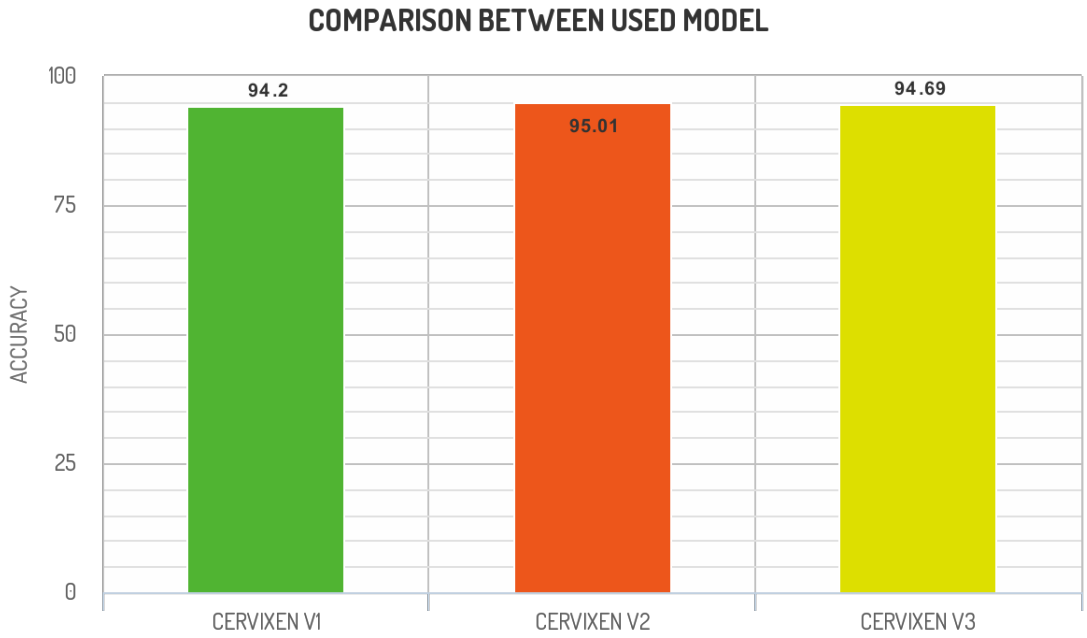

Figure 5.8: Comparison between our proposed models

It is clear from 5.2 that "CERVIXEN V2" represents 95.01 percent validation accuracy which is highest among all the models. Therefore, out of the three models, "CERVIXEN V2" performs the best. As a result, "CERVIXEN V2" was chosen for the project.

Chapter 6

Conclusion and Future Work

6.1 Conclusion and Future Work

Cervical cancer is the second most common type of cancer in women worldwide, and many women suffer greatly as a result of it. The number of cervical cancer patients, as well as the number of deaths, is not reduced significantly especially in developing countries. The majority of treatments are time-consuming, expensive, and require regular testing. It takes longer to receive the results, and it also takes much longer to determine the cancer stage. To help alleviate the situation, our model will be able to identify cancerous cells and determine whether they are in the endocervix or ectocervix by evaluating the patients' colposcopy images. Additionally, our system will be able to quickly perform binary categorization of the cervical stage of cancer. We used the dataset for training VGG16, VGG19, InceptionV3, ResNet50, and ResNet101 architectures, and compared with the results to our own architectures CERVIXEN V1 (Ensemble of VGG19 and ResNet101 architecture), CERVIXEN V2 (Ensemble of VGG19, InceptionV3, and ResNet101 architecture), and CERVIXEN V3(Ensemble of VGG16, VGG19, ResNet50 and ResNet101 architecture). We will endeavor to collect additional data in the future to enhance our model. We will also provide Dicom images to get more accurate results for different kinds of cervical data. We will also include Explainable AI, which will help us improve our outcomes. As a result, we will be able to develop a method for contributing to the cervical cancer problem in an efficient manner.

Bibliography

[1] E. Lăără, N. Day, and M. Hakama, "Trends in mortality from cervical cancer in the nordic countries: Association with organised screening programmes," *The Lancet*, vol. 329, no. 8544, pp. 1247–1249, 1987.

[2] K. Magnus, F. Langmark, and A. Andersen, "Mass screening for cervical cancer in øtfold county of norway 1959-77," *International journal of cancer*, vol. 39, no. 3, pp. 311–316, 1987.

[3] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and intelligent laboratory systems*, vol. 2, no. 1-3, pp. 37–52, 1987.

[4] D. M. Eddy, "Screening for cervical cancer," *Annals of Internal Medicine*, vol. 113, no. 3, pp. 214–226, 1990.

[5] H. Soost, H. Lange, W. Lehmacher, and B. Ruffing-Kullmann, "The validation of cervical cytology. sensitivity, specificity and predictive values.," *Acta cytologica*, vol. 35, no. 1, pp. 8–14, 1991.

[6] H. M. Shingleton, R. L. Patrick, W. W. Johnston, and R. A. Smith, "The current status of the papanicolaou smear.," *CA: a cancer journal for clinicians*, vol. 45, no. 5, pp. 305–320, 1995.

[7] R. Sankaranarayanan, R. Wesley, T. Somanathan, *et al.*, "Visual inspection of the uterine cervix after the application of acetic acid in the detection of cervical carcinoma and its precursors," *Cancer: Interdisciplinary International Journal of the American Cancer Society*, vol. 83, no. 10, pp. 2150–2156, 1998.

[8] E.-K. Yim and J.-S. Park, "The role of hpv e6 and e7 oncoproteins in hpv-associated cervical carcinogenesis," *Cancer research and treatment: official journal of Korean Cancer Association*, vol. 37, no. 6, p. 319, 2005.

[9] D. M. Reif, A. A. Motsinger, B. A. McKinney, J. E. Crowe, and J. H. Moore, "Feature selection using a random forests classifier for the integrated analysis of multiple data types," in *2006 IEEE Symposium on Computational Intelligence and Bioinformatics and Computational Biology*, IEEE, 2006, pp. 1–8.

[10] P. E. Castle, "Invited commentary: Is monitoring of human papillomavirus infection for viral persistence ready for use in cervical cancer screening?" *American journal of epidemiology*, vol. 168, no. 2, pp. 138–144, 2008.

[11] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley interdisciplinary reviews: computational statistics*, vol. 2, no. 4, pp. 433–459, 2010.

[12] W. H. Organization *et al.*, "Who guidance note: Comprehensive cervical cancer prevention and control: A healthier future for girls and women," 2013.

[13] A. Diamantis and E. Magiorkinis, "Pioneers of exfoliative cytology in the 19th century: The predecessors of george papanicolaou," *Cytopathology*, vol. 25, no. 4, pp. 215–224, 2014.

[14] M. J. Huchko, J. Sneden, J. M. Zakaras, *et al.*, "A randomized trial comparing the diagnostic accuracy of visual inspection with acetic acid to visual inspection with lugol's iodine for cervical cancer screening in hiv-infected women," *PloS one*, vol. 10, no. 4, e0118568, 2015.

[15] P. C. Mpata, "Student nurses' risk perception of contracting cervical cancer in zimbabwe," Ph.D. dissertation, 2015.

[16] S. Y. Tan and Y. Tatsumura, "George papanicolaou (1883–1962): Discoverer of the pap smear," *Singapore medical journal*, vol. 56, no. 10, p. 586, 2015.

[17] M. Abadi, P. Barham, J. Chen, *et al.*, "{Tensorflow}: A system for {large-scale} machine learning," in *12th USENIX symposium on operating systems design and implementation (OSDI 16)*, 2016, pp. 265–283.

[18] A. Lavin and S. Gray, "Fast algorithms for convolutional neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 4013–4021.

[19] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.

[20] N. Chantziantoniou, A. D. Donnelly, M. Mukherjee, M. E. Boon, and R. M. Austin, "Inception and development of the papanicolaou stain method," *Acta cytologica*, vol. 61, no. 4-5, pp. 266–280, 2017.

[21] A. Gulli and S. Pal, *Deep learning with Keras*. Packt Publishing Ltd, 2017.

[22] W. Wu and H. Zhou, "Data-driven diagnosis of cervical cancer with support vector machine-based approaches," *IEEE Access*, vol. 5, pp. 25 189–25 195, 2017.

[23] P. Basu, S. Mittal, D. B. Vale, and Y. C. Kharaji, "Secondary prevention of cervical cancer," *Best Practice & Research Clinical Obstetrics & Gynaecology*, vol. 47, pp. 73–85, 2018.

[24] N. K. Manaswi, N. K. Manaswi, and S. John, *Deep learning with applications using python*. Springer, 2018.

[25] J. Brownlee, *Generative adversarial networks with python: deep learning generative models for image synthesis and image translation*. Machine Learning Mastery, 2019.

[26] C. Gilham, A. Sargent, H. C. Kitchener, and J. Peto, "Hpv testing compared with routine cytology in cervical screening: Long-term follow-up of artistic rct.," *Health Technol Assess*, vol. 23, no. 28, pp. 1–44, 2019.

[27] *Resnet (34, 50, 101): Residual cnns for image classification tasks*, Jan. 2019. [Online]. Available: https://neurohive.io/en/popular-networks/resnet/.

[28] *Softmax function*, May 2019. [Online]. Available: https://deepai.org/machine-learning-glossary-and-terms/softmax-layer.

[29] U. Udofia, *Basic overview of convolutional neural network (cnn)*, Sep. 2019. [Online]. Available: https://medium.com/dataseries/basic-overview-of-convolutional-neural-network-cnn-4fcc7dbb4f17#:~:text=The%5C%20activation%5C%20function%5C%20is%5C%20a,of%5C%20neurons%5C%20as%5C%20input.%5C%E2%5C%80%5C%9D%5C%20%5C%E2%5C%80%5C%94.

[30] E. A. Van Dyne, "Establishing baseline cervical cancer screening coverage—india, 2015–2016," *MMWR. Morbidity and mortality weekly report*, vol. 68, 2019.

[31] W. William, A. Ware, A. H. Basaza-Ejiri, and J. Obungoloch, "Cervical cancer classification from pap-smears using an enhanced fuzzy c-means algorithm," *Informatics in Medicine Unlocked*, vol. 14, pp. 23–33, 2019.

[32] R. Alake, *Deep learning: Understand the inception module*, Dec. 2020. [Online]. Available: https://towardsdatascience.com/deep-learning-understand-the-inception-module-56146866e652#:~:text=An%5C%20inception%5C%20network%5C%20is%5C%20a,details%5C%20of%5C%20the%5C%20inception%5C%20module.

[33] Z. Alyafeai and L. Ghouti, "A fully-automated deep learning pipeline for cervical cancer classification," *Expert Systems with Applications*, vol. 141, p. 112 951, 2020.

[34] T. A. Amma, A. R. Sunny, K. Biji, and M. Mohanan, "Lung cancer identification and prediction based on vgg architecture," *International Journal of Research in Engineering, Science and Management*, vol. 3, no. 7, pp. 88–92, 2020.

[35] S. L. Bedell, L. S. Goldstein, A. R. Goldstein, and A. T. Goldstein, "Cervical cancer screening: Past, present, and future," *Sexual medicine reviews*, vol. 8, no. 1, pp. 28–37, 2020.

[36] O. Harel, *224 224 cervical cancer screening*, May 2020. [Online]. Available: https://www.kaggle.com/datasets/ofriharel/224-224-cervical-cancer-screening.

[37] A. Kaushik, *Understanding resnet50 architecture*, Jul. 2020. [Online]. Available: https://iq.opengenus.org/resnet50-architecture/.

[38] W. H. Organization *et al.*, "Global strategy to accelerate the elimination of cervical cancer as a public health problem," 2020.

[39] R. Thakur, *Step by step vgg16 implementation in keras for beginners*, Nov. 2020. [Online]. Available: https://towardsdatascience.com/step-by-step-vgg16-implementation-in-keras-for-beginners-a833c686ae6c.

[40] J. Brownlee, *A gentle introduction to ensemble learning algorithms*, Apr. 2021. [Online]. Available: https://machinelearningmastery.com/tour-of-ensemble-learning-algorithms/#:~:text=Ensemble%5C%20learning%5C%20is%5C%20a%5C%20general,the%5C%20predictions%5C%20from%5C%20multiple%5C%20models..

[41] *Cervical cancer screening (pdq®)–patient version*, 2021. [Online]. Available: https://www.cancer.gov/types/cervical/patient/cervical-screening-pdq.

[42] A. Khamparia, D. Gupta, J. J. Rodrigues, and V. H. C. de Albuquerque, "Dcavn: Cervical cancer prediction and classification using deep convolutional and variational autoencoder network," *Multimedia Tools and Applications*, vol. 80, no. 20, pp. 30 399–30 415, 2021.

[43] M. Mehmood, M. Rizwan, S. Abbas, *et al.*, "Machine learning assisted cervical cancer detection," *Frontiers in public health*, p. 2024, 2021.

[44] S. Mostafa and F.-X. Wu, *Diagnosis of autism spectrum disorder with convolutional autoencoder and structural mri images*, Jul. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B978012822822700003X.

[45] 2. Published: Jul 12, *The hpv vaccine: Access and use in the u.s.* Jul. 2021. [Online]. Available: https://www.kff.org/womens-health-policy/fact-sheet/the-hpv-vaccine-access-and-use-in-the-u-s/#:~:text=Since%5C%20HPV%5C%20vaccines%5C%20were%5C%20first,include%5C%20boys%5C%20and%5C%20young%5C%20men.

[46] D. I. Sec., *Vgg-19 convolutional neural network*, Mar. 2021. [Online]. Available: https://blog.techcraft.org/vgg-19-convolutional-neural-network/.

[47] H. Sung, J. Ferlay, R. L. Siegel, *et al.*, "Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a cancer journal for clinicians*, vol. 71, no. 3, pp. 209–249, 2021.

[48] A. N. T, *Inception v3 model architecture*, Oct. 2021. [Online]. Available: https://iq.opengenus.org/inception-v3-model-architecture/.

[49] W. Wang, C. Qiu, Z. Yin, *et al.*, "Blockchain and puf-based lightweight authentication protocol for wireless medical sensor networks," *IEEE Internet of Things Journal*, 2021.

[50] H. Xiong, C. Jin, M. Alazab, *et al.*, "On the design of blockchain-based ecdsa with fault-tolerant batch verication protocol for blockchain-enabled iomt," *IEEE Journal of Biomedical and Health Informatics*, 2021.

[51] Q. Yin, "The application of machine learning in cervical cancer prediction," in *2021 6th International Conference on Machine Learning Technologies*, 2021, pp. 12–19.

[52] Y. Yu, J. Ma, W. Zhao, Z. Li, and S. Ding, "Msci: A multistate dataset for colposcopy image classification of cervical cancer screening," *International journal of medical informatics*, vol. 146, p. 104 352, 2021.

[53] *Cervical cancer - screening and prevention*, May 2022. [Online]. Available: https://www.cancer.net/cancer-types/cervical-cancer/screening-and-prevention.

[54] *Residual networks (resnet) - deep learning*, Jan. 2022. [Online]. Available: https://www.geeksforgeeks.org/residual-networks-resnet-deep-learning/.

[55] *Hpv and pap testing*. [Online]. Available: https://www.cancer.gov/types/cervical/pap-hpv-testing-fact-sheet.

[56] *Lung mesothelioma*. [Online]. Available: http://lung-mesothelioma101.blogspot.com/2008/?cv=1.

[57] *The nobel prize in physiology or medicine 2008*. [Online]. Available: https://www.nobelprize.org/prizes/medicine/2008/press-release/.