
**Addis Ababa University**

**College of Natural and Computational Sciences**

**Department of Computer Science**

# *A Model for the Detection of Human Papilloma Virus Cancers using Deep Learning*

Solomon Gezahegn Temesgen

Thesis Submitted to the Department of Computer Science in Partial Fulfillment for the Degree of Master of Science in Computer Science

Addis Ababa, Ethiopia

August, 2021

Addis Ababa University

College of Natural and Computational Sciences

Department of Computer Science

Solomon Gezahegn Temesgen

Advisor: Solomon Atnafu (PhD)

This is to certify that the thesis prepared by Solomon Gezahegn Temesgen, titled: *A Model for the Detection of Human Papilloma Virus Cancers using Deep Learning* and submitted in partial fulfillment of the requirements for the Degree of Master of Science in Computer Science complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the Examining Committee:

**Name** **Signature**

**Date**

Advisor: _____

Examiner: _____

Examiner: _____

## Abstract

Cancer diseases caused by Human Papilloma Virus (HPV) are the most common killer infectious diseases in the world. Human papilloma virus causes cervical cancer the second most common cancer in women, oral cancer that can cause cancer of the mouth and tongue and anal cancer. Therefore, there is a need to design an automatic HPV-related disease detection model that can assist medical professionals in an early detection of the diseases with a competitive accuracy.

A convolutional neural network (CNN) is one of deep learning that has been used in computer vision is chosen, for the detection of diseases. CNNs represent an interesting approach for the processing of adaptive images. The algorithm is used for preprocessing, extraction, detection and for evaluating the model's accuracy.

The proposed model that is called human papilloma virus caused cancer detection model has a total of eight layers, five convolutions, and three dense layers. It receives 127 × 127 color images and produces two outputs. The proposed model is trained using a total of 66,336 images which includes both infected and healthy images.

The validity of our proposed model has been validated through experiments using CNN algorithms acquired from two publicly available pre-trained models namely VGG16 and Inception V3. The detection result shows the proposed model is effective in detecting HPV caused cancers. Based on the detection result the HPV caused cancer detection model has 99.3% detection accuracy and 99.4% testing accuracy. Our contribution of this work is we design a CNN model that can be used for detection of cancers caused by HPV and we have compared three CNN models and found that our CNN model performs better on those models for HPV caused cancer diseases detection.

**Keywords**: - Convolutional Neural Network, Cancer Detection, Deep Learning, Human Papilloma Virus.

## Acknowledgments

Firstly, for everything, I want to thank God. I would also like to express my sincere gratitude to my advisor Dr. Solomon Atnafu for his initial guidance and ongoing support during my thesis work.

In addition to my advisor, I would like to thank Dr. Haregewoien from Betezata Hospital for HPV-related cervical cancer and ENT specialist Mr. Samson from St. Paulo's Hospital for HPV-related oral cancer to their continuous assistance and they are always available in my study to discuss and provide informative comments on our work. Working with them was a pleasure and also a great learning experience.

Thank you to IEEE, Google Scholar, Science Direct, kaggle.com, sci-hub.se, and Stack Overflow for your assistance. Without these open-access sites, this research would not have been conducted.

Last but not least, I'd want to express my gratitude to my love Ms. Tewnet Minlargih, my family, particularly my mother Tesfa Damtie and aunt Alemtsehay Damtie, for their well wishes, which help me stay focused on my goal, and to my classmates and friends for their support.

## Table of Contents


Abstract ..... 1
List of Tables ..... iii
List of Figures ..... iv
Acronyms ..... v
Chapter 1: Introduction ..... 1
1.1 Background ..... 1
1.2 Statement of the Problem ..... 2
1.3 Motivation ..... 4
1.4 Objectives ..... 5
1.5 Methods ..... 5
1.6 Scope and Limitations ..... 7
1.7 Application of Results ..... 7
1.8 Organization of the Rest of the Thesis ..... 8
Chapter 2: Literature Review ..... 9
2.1 Cancer Detection ..... 9
2.2 Cancers Caused By HPV ..... 11
2.2.1 Cervical Cancer ..... 11
2.2.2 Oral Cancer ..... 12
2.2.3 Other HPV Related Cancers ..... 13
2.3 Types of HPV ..... 15
2.4 Deep Learning ..... 15
2.4.1 Convolutional Neural Network ..... 18
2.4.2 Application of Deep Learning in Cancer Detection ..... 26
Chapter 3: Related Work ..... 28
3.1 Introduction ..... 28
3.2 Machine Learning ..... 28
3.3 Cervical Cancer Detection ..... 29
3.4 Deep Learning ..... 31
Chapter4: Modeling Detection of HPV caused Cancer ..... 35
4.1 Model Selection ..... 35
4.2 Overview of the Model ..... 36
4.2.1 Data Collection and Dataset Preparation ..... 38
4.2.2 Data Preprocessing ..... 38

4.2.3 HPV Caused Cancer Detection CNN Model ..... 42
4.2.4 Training Components of HPV caused Cancer Detection Model ..... 46
4.3 Detection Using HPV Caused Cancer Detection Model ..... 46
4.4 Detection Using Pre-Trained Models ..... 47
4.5 Experimental Setup ..... 48
4.6 Augmentation Parameters ..... 48
4.7 Hyperparameter Settings ..... 49
Chapter 5: Experiment and Evaluation ..... 51
5.1 Development Environment and Tools ..... 51
5.2 Model Evaluation ..... 52
5.3 Pre-trained CNN Model ..... 53
5.2.1 Detection of Cancers Caused by HPV by using VGG16 Pre-trained Model .. 53
5.2.2 Result Analysis of VGG16 ..... 54
5.2.3 Detection of Cancers Caused by HPV by using InceptionV3 ..... 55
5.2.4 Result Analysis of InceptionV3 ..... 56
5.4 HPV Caused Cancer Detection CNN Model ..... 56
5.4.1 Scenario 1: Modifying the Training and Testing Dataset Ratio ..... 57
5.4.2 Scenario 2: Learning Rate Changing ..... 57
5.4.3 Scenario 3: Using Different Activation Function ..... 57
5.4.4 Scenario 4: With and Without Dataset Augmentation ..... 57
5.4.5 Result Analysis for the HPV Caused Cancer Detection Model ..... 58
5.5 Discussion ..... 59
Chapter 6: Conclusion and Future Work ..... 63
6.1 Conclusion ..... 63
6.2 Contribution ..... 64
6.3 Future Work ..... 64
References ..... 65
Annex A: The Proposed CNN Model Code ..... 75

**List of Tables**

Table 4.1: HPV Caused Cancer Detection Model Summary of Parameters for Detection of HPV caused Cancer ..... 45
Table 4.2: We Employed Augmentation Techniques ..... 48
Table 5.1: Pre-Trained Model VGG16's Mean Accuracy and Loss ..... 55
Table 5.2: Pre-Trained Model Mean Accuracy and Loss Of Inceptionv3 ..... 56
Table 5.7: The Accuracy and Loss of the HPV-Caused Cancer Detection Model ..... 59

**List of Figures**

Figure 2.1: CNN Model Example ..... 19
Figure 2.2: Examples of Input Volume and Filter ..... 20
Figure 2.3: Examples of Convolution Operation ..... 21
Figure 2.4: Examples of Convolution of a 3D Input Volume ..... 22
Figure 2.5: Examples of Convolution Operation with 2 Filters ..... 23
Figure 2.6: Examples of One Convolution Layer with Activation Function ..... 24
Figure 2.7: Max Pooling Example ..... 24
Figure 2.8: Fully Connected Layer Example ..... 25
Figure 4.1: The HPV caused cancer detection model ..... 37
Figure 4.2: Image Resized ..... 40
Figure 4.3: In the HPV caused cancer detection model, Feature Extraction [69] ..... 41
Figure 5.5: HPV Caused Cancer Detection Model Training and Validation Accuracy ..... 58
Figure 5.6: HPV Caused Cancer Detection Model Training and Validation Los ..... 59
Figure 5.7: The Three Experiments Mean Accuracy ..... 60
Figure 5.8: The Three Experiments Mean Loss ..... 61

**Acronyms**

| | |
|---|---|
| **ANN** | Artificial Neural Network |
| **CNN** | Convolutional Neural Network |
| **DBN** | Deep Belief Network |
| **ENT** | Ear, Nose and Throat |
| **FCN** | Fully Convolutional Network |
| **HPV** | Human Papilloma Virus |
| **HR** | High Risk |
| **ILSVRC** | ImageNet Large Scale Visual Recognition Challenge |
| **LSTM** | Long Short-Term Memory |
| **MRI** | Magnetic Resonance Imaging |
| **NN** | Neural Network |
| **RBM** | Restricted Boltzmann Machine |
| **ReLU** | Rectified Linear Unit |
| **RNN** | Recurrent Neural Network |
| **VGG** | Visual Geometry Group |
| **VIA** | Visual Inspection with Acetic Acid |
| **WHO** | World Health Organization |

# Chapter 1: Introduction

## 1.1 Background

Cancer diseases caused by Human Papilloma Virus (HPV) are the most common killer infectious diseases in the world [1]. HPV is a significant cause of the highest risk factor for cervical cancer growth and other associated cancers [2]. In most cases, HPV-related cancers arise after chronic infections with certain forms of viruses, which are typically destroyed within a few years without any intervention by the immune system [3].

The long-term acquisition of high-risk HPV results in abnormal cell growth on the cervix surfaces, a possible precursor to cervical, vaginal and vulvar cancer growth. The most prominent are the high-risk categories, HPV-16 and HPV-18, associated with around 70% of all cervical cancers [4].

Cancers linked to HPV include the following [5]:

- Cervical cancer: HPV infection is the cause of nearly all cervical cancers.
- Oral cancer: HPV can cause cancer of the mouth and tongue.
- Other cancers: anal cancer, vulvar and vaginal cancers in women, and penile cancer in males are among the less prevalent cancers.

All of the cancers mentioned above have their own symptoms and can be described and categorized by image. In certain cases, until it develops warts, the immune system of the body fights an HPV infection. When warts emerge, depending on which form of HPV is involved, they differ in appearance [6]:

- Genital warts: These occur as flat lesions, small bumps such as cauliflower or tiny stem like protrusions.
- Common warts: are appear on the hands and fingers as rough, raised bumps and usually occur.
- Plantar warts: are difficult, grainy growths that typically occur on the feet's heels or balls.
- Flat warts: are lesions that are flat-topped, slightly elevated. They can appear anywhere, but kids usually get them on the face and in the beard zone, men tend to get them. Females prefer to have them on their legs.

Medical imaging is a technique for capturing images of bodily components for medical purposes such as observing and diagnosing diseases. Every day, millions of medical imaging operations are performed around the world. Due to advancements in image processing techniques such as image recognition, analysis, and enhancement, medical imaging is rapidly evolving. Many diseases can now be detected more easily thanks to image processing. The use of a computer to edit a digital image is known as image processing. This method has numerous advantages, including elasticity, adaptability, data storage, and communication and medical images [6, 7].

The processing of medical digital images will decrease the impact of noise, boost the image and enhance its quality for better observation. Processed images can accurately reflect the focus of disease and visually communicate medical and pathological information of the part of the body captured by image. Medical Image evaluation is important in the area of remedy [8]. The advancement of image analysis algorithms such as deep learning promises greater applications in a variety of medical sectors, particularly in the field of medical diagnostics. While it is not clear that deep learning can replace the role of doctors/clinicians in medical diagnosis at this time, it can provide valuable support for medical experts [9, 10].

There are different causes that lead to cervical, oral and anal cancers. HPV is not the only cause of those cancers, but it's a high risk type that leads to cancer that kills. Researchers focus on the detection of cervical cancer, but not on the causes and pre-cancers detection. In earlier stages of the disease, cancers are diagnosed when treatment could be more effective. In our work, we are focused on detection of cancers using deep learning approach to reduce the chance of capturing cancers only caused by HPV since it's the leading cause of cancers and pre-cancer or early stage detection of the diseases can identify the type of HPV to determine the stages of the symptom is a high risk or not.

## 1.2 Statement of the Problem

The author in [11] presents HPV caused cancer is linked all over the world with significant morbidity and mortality. It is well known that high-risk HPV strains are one of the major causative factors for cervical and oral cancer, and this form of the disease is preventable if it's detected at early stage.

The authors in [12] a research was performed by medical researchers on the global burden of HPV-related cancers, which remains a significant cause of cancer in both men and women. As a result, it has been identified that giving access to HPV identification and cervical screening to the majority of women around the world is one of the greatest goals and challenges in global health. They suggests that to increase decision making regarding to HPV related cancers detection techniques should be implemented.

R. Gorantla *et al.* [13] suggested the implementation of a fully automated technique for cervical cancer screening called CervixNet using cervigrams. In order to provide adequate care to save the patient from the clutches of excruciating death, early detection in cases of cervical and other associated cancers is crucial. So, it does not involve early detection cancers.

The authors in [14] are designing an automated HPV detection system for images captured from the Linear Array HPV Genotyping Test by Roche Molecular Diagnostics. For 37 different forms, this test provides type-specific HPV genotype results with different levels of risk for cervical cancer growth. The algorithm was checked on 17 patient's cases and found that only five of these possibilities were actually forms of HPV after testing the method. The diagnosis of patient 17 was described as types 2, 3, 6, 10, and 22. There are more than 100 different type of HPV to lead cancers i.e. cervical, oral and anal but researchers detect only 37 types. This is only for cervical cancers and other types are not included. But we include HPV caused cervical and oral cancer.

Many authors deal with the detection of cervical cancers using machine learning algorithm but not focusing on the cause of cancers such as HPV and pre-cancer or early detection. Machine learning is slow for larger dataset; it requires a large amount of time to process. One of the most problems using machine learning is the acquisition of data. Additionally, collecting data comes with a cost. Also, it so happens that when we are collecting data from medical centers, it might contain a large volume of bogus and incorrect data. Many times they do face a situation where they find an imbalance in data which leads to poor accuracy of models.

Thus, we need to use deep learning techniques to detect HPV-related cancers. The detection is performed with images of having a symptom of the diseases. Deep learning approach

provides consistent, reasonably accurate, less time consuming and cost effective solutions for clinical experts and patients to identify cancer disease.

This work will try to explore and address the following research questions:

- How can we design a model that can be used to detect HPV-related cancer and runs with a small hardware and software requirements?
- How can we design algorithms that can be used to implement the designed model?
- How can we demonstrate the validity and effectiveness of the model designed?

## 1.3 Motivation

Cervical cancer, caused by the human papilloma virus (HPV), is the second most frequent malignancy among women in developing countries, with an estimated 570,000 new cases in 2018 [5]. Cervical cancer claimed the lives of around 311,000 women in 2018, with poor and middle-income countries accounting for more than 85 percent of these deaths [5, 8]. Medical diagnostics and applications benefit greatly from today's technological breakthroughs. One of these technologies is artificial intelligence, whereas deep learning is a newer technology that can help with disease diagnosis [9].

In Ethiopia, 29.43 million women aged 15 and up are at risk of developing cervical cancer as a result of the HPV virus [10]. Each year, 6294 women are diagnosed with cervical cancer, with 4884 dying as a result of the disease, according to current report [10]. Cervical cancer is the second most common cancer among Ethiopian women, as well as the second most common cancer in women aged 15 to 44 [10].

In Ethiopia medical experts made decisions by looking at the medical images of HPV to classify than of normal or abnormal/ infected. But, it is important to support this detection by automated mechanisms to reduce mistakes, to describe the detail information about the current stage and type of HPV and to improve the performance of prediction. That is why we are motivated to help the diagnosis process of detecting HPV caused cancers diseases to decrease the patient’s death using deep learning approach.

## 1.4 Objectives

### General Objective

The general objective of this thesis is to design a model that can be used for detection of cancers diseases caused by HPV using deep learning approach.

### Specific Objective

To achieve the general objective of the thesis, the specific objectives are identified to:

- Collect, classify and analyze relevant cancer image data and prepare a dataset for model training and model testing.
- Preprocess image data and segment region of interest.
- Extract features from segmented images.
- Design detection models of cancers diseases caused by HPV that can run on a small hardware and software requirements.
- Develop a prototype to demonstrate the use of the HPV caused cancer detection model.
- Evaluate the performance of detection models.

## 1.5 Methods

We used the following approaches in order to accomplish the general and specific objectives listed above.

### Literature Review

Related literature from different sources (books, Internet, journals, etc.) will be review to understand human papillomavirus. The major activities are:

- The most recent findings in the human papillomavirus field will be reviewed.
- The most recent findings will be analyzed in the field of convolutionary neural networks.
- The most current researches in the area of detecting disease will be reviewed.
- Identifying the limitation of different types of human papillomavirus detection researches.
- Identifying the strength and weakness of solutions proposed by recent and latest researchers.

## Data Collection

When we choose to use a neural network or deep learning algorithms in science, the most important thing is to acquire the data used to train the model of the neural network.

For the model to be trained and tested, a number of image data is needed. The image data will be collected from Ethiopian Medical Centers such as St. Paulo's Hospital and Betezata Hospital in addition to images that are found on the Internet.

## Data Preprocessing

If we want to deal with image processing, preprocessing used to be the most important activity. Before feeding into the neural network or deep learning algorithm, it is a conversion of the raw information. In this thesis, images from various medical sectors in the area as well as the Internet will be used for model training and model testing. The dataset will be created using a 70/30 strategy, with 70% of the dataset being used for training and 30% for model testing.

## Data Partitioning

First and foremost, the dataset is split into two parts: training and testing. The training split is used to train the model, whereas the test split is used to test the model that is not visible during model training. The validation split will be used to assess the performance of the model created during training and to fine-tune model parameters in order to pick the optimum model performance.

## Image Detection

Detection is the key process of image recognition decision making and determining whether the image shows normal or abnormal cases. There are many detection techniques; some of them are called Linear Regression, Logistic Regression, k-Nearest Neighbor, Naive Bayes, Decision Trees, Artificial Neural Networks, Linear Classifier, Support Vector Machine, Random Forest, etc.

## Deep Learning

Deep learning (also known as deep structured learning or differential programming) is part of a wider range of machine learning techniques based on representational learning on

artificial neural networks. Learning is controlled, partially controlled, or uncontrolled. In deep learning structures algorithms to create a neural network that can learn on its own and make smart decisions in layers.

A convolutionary neural network is a subset of deep neural networks in deep learning, most widely applied to visual imagery analysis. They have image and video recognition apps, recommendation systems, detection of pictures, analysis of medical images, and processing of natural languages [9].

### Performance Evaluation

Finally, through various measurement criteria, the model will be tested to verify its performance. The HPV caused cancer detection model conducts several experiments by adjusting the ratio of training and testing datasets, using different learning rates, using the dataset before and after augmentation, using various activation functions and finally compared the performance with two pre-trained models namely InceptionV3 and VGG.

## 1.6 Scope and Limitations

Since the detection of cancers caused by HPV is a very broad field of study to be studied, it is necessary to establish some kind of task coverage boundary to achieve better results. As a result, our work here is limited to only detecting an input image that is only including cervical and oral HPV related cancers into either of two different classes i.e. Healthy and Infected. With its stage and types, we won't include all HPV-related cancers.

## 1.7 Application of Results

We all accept that technology has become stable and that there is no way we can prevent its use. Therefore, to better our living standards, we must use technology. Deep learning for the identification and detection of diseases is one of the latest technologies and is highly used by users for various purposes.

In addition to these, the significance of our thesis is described as follows:

- In the first place, the result of this research would enable medical experts to understand the importance of computer vision in the field of medicine.

- The result of this study will assist numerous hospitals to include effective steps in circumstances for the detection of cancers associated with HPV.
- It helps to check the health of human from cancers caused by HPV,
- It reduces the chance of death,
- It's applicable to both healthy as well as infected images,
- It detects cancers caused by HPV easily with a short period of time.

## 1.8 Organization of the Rest of the Thesis

The remaining part of the dissertation is formulated as follows. Chapter two covers literature reviews in the field of cancer detection, detection of the HPV types, image processing and deep learning. Chapter three covers related work that demonstrates the efforts of numerous author-related researches to identify and recognize HPV-induced cancers at different locations. The fourth chapter briefly describes and discusses the design and development of a model for the detection of cancers caused by HPV along with the preparation. The experiment and evaluation of the models and algorithms using the implemented model is discussed in chapter five. Finally, our findings are summarized and future studies are discussed in chapter six.

# Chapter 2: Literature Review

This chapter offers a description of the state of the art concepts linked to our work. The key components of our work, such as Cancer Detection, Cancers Caused by HPV, Types of HPV are reviewed and try to list out some applications of deep learning in cancer detection.

## 2.1 Cancer Detection

Cancer detection is a method that is multiphase. Sometimes, because of some symptom or another, the patient will go to physicians. Cancer is often detected by chance or by screening [15]. In cancer diagnosis, early detection plays a crucial role and may increase long-term survival rates. A very important technique for early detection and diagnosis of cancer is medical imaging [15]. The rising burden of cancer is attributed to a variety of factors, including residential development and aging, as well as the shifting commonality of some cancer causes linked to social and economic progress [16].

Cancer is a major global public health concern and is the second leading cause of death [17]. Growth is slowing for cancers that are vulnerable to early detection by screening (i.e. breast, prostate, and other cancers) and major ethnic and regional inequalities exist for cancers that are highly preventable, such as cervical and oral cancers [17]. Early diagnosis, prognosis and monitoring of the health of patients are considered to be crucial concerns for effective care and prevention of side effects, helping to minimize morbidity, mortality and also improve the quality of life of patients. Because of the dynamic existence of cancer, at least two biomarkers must be detected or quantified simultaneously for the correct decision-making, which focuses on the evolution of multiplex lateral flow assays [18].

It takes a lot of time to detect vast quantities of medical images with human labor, and even for professionals, precision is not well assured. Therefore, a quick and precise automated cancer disease diagnosis system is highly needed. Diseases such as cancer should be detected as soon as possible where the time factor is significant. Medical images include variables such as noise, unclear limits of the tumor and large variations in the presence of the tumor, making it difficult to find exact regions of the tumor [19, 20].

With regard to the fact that the amount of contextual information is of great significance for detecting abnormalities from images and provided that the fusion of multiple sources of

image information can improve the detection efficiency. One of the easiest ways to avoid patient mortality is to identify and analyze cancers correctly and accurately [21]. Cancer is among the leading cause of mortality among women in developed as well as under-developing countries. Detecting cancer in the early stages of its development can allow proper care for patients. Algorithms for histopathological images are increasingly evolving in medical image processing, but there is still a strong demand for an automated method to achieve reliable and highly accurate performance [22].

Based on microscopic analysis of tissue/cells (e.g. differentiation capacity, cell pleomorphic, nuclear to cytoplasm ratio) or clinical biological markers, tumors can be categorized into benign and malignant tumors [23]. For females with the highest morbidity, it is one of the most destructive diseases. Moreover, the trajectory of cancer evolves rapidly. Thus, delayed diagnosis may have a big effect on patients. A successful tool to identify indeterminate lesions early is cancer screening. Imaging diagnosis, which includes magnetic resonance imaging (MRI), mammography, and ultrasound images, is the common method of cancer screening. Similar signs are correlated with various approaches to imaging. The MRI is particularly susceptible to soft tissue lesions for screening. It is expensive, however, with a relatively long scan time and a greater rate of false positives. As a result, MRI is prescribed specifically for women at high risk of breast cancer. Mammography is particularly susceptible to calcification detection, but with disadvantages for individuals with dense breast tissues.

Ultrasound transforms electrical signals into ultrasound signals using a transducer. The reflected sound waves will generate an image through computer processing based on the different magnitude of reflected ultrasound waves and echo time. As a result, no ionizing radiation and real-time analysis have the benefit of ultrasound. Ultrasound is used clinically for echo driven biopsy exams. Mammography and ultrasound are probably the most common approaches to screening [23].

In any object recognition process, detection is the key and concluding step. The aim of recognition is to recognize or identify an object of interest and to make a decision on the basis of its features, characteristics and properties. By categorizing objects into one or more

clusters or classes, grouping goes one step further (like whether a cell is benign or malignant) [24].

## 2.2 Cancers Caused By HPV

Almost all cervical, oral and other cancers, including anal cancer, vulvar and vaginal cancers in women, and penile cancer in men, are caused by HPV [5]. The only HPV-associated cancer for which screening is routinely recommended is cervical cancer. Recommendations recommend that women aged 21-65 years are routinely screened for cervical pre-cancers and cancers [25]. Infection in both the genital and oral regions, regardless of HPV type, has been characterized as concurrent genital and oral HPV infection. A concurrent high-risk genital and oral HPV infection was found in both the genital and oral regions, as well as any high-risk type. Infection with a certain type of HPV in the vaginal and oral sites was characterized as concurrent type-specific infection [26]. It is a DNA virus that infects the mucous membranes and skin. To date, more than 100 different HPV types have been categorized [27].

### 2.2.1 Cervical Cancer

Nearly all occurrences of cervical cancer (99%) are caused by high-risk human HPV infection, a virus that is spread through sexual contact [5]. Cervical cancer is the fourth most frequent malignancy in women. According to WHO projections, 570,000 persons were diagnosed with cervical cancer worldwide in 2018, with 311,000 women dying as a result of the disease [5, 27], Cervical cancer is the most common genital cancer in women globally, with over half a million new cases each year [27].

After breast cancer, cervical cancer is the silent killer that affects the sociability of women's disease. The key limitation of these commercial detections has been time-consuming and needs a specialist to predict and detect the cancer-causing HPV infection [28]. This is a known virus causing cancer that has been mainly associated with cervical cancer [29].

For the development of cervical cancer, the persistence of high-risk HPV infection is important and represents the most significant established risk factor. Other risk factors include: sexual activity, oral contraceptive cocktail use, multi-party use, smoking and immunosuppression [30]. Including menarche age, first delivery age, menopause status, and

long-term estrogen exposure status, there are several independent risk factors for causing cervical cancer. The short distance between the menarche and the beginning of sexual activity raises the risk of cervical cancer development [31].

HPV infection is very common because of its type of transmission, but very few infections are advanced and malignant and lead to cervical cancer. HPV is the primary cause of cervical carcinoma, although it is suggested that some risk factors are associated with cervical cancer progression [31].

Cervical cancer is the world's second most common cancer and the most prevalent in sub-Saharan Africa among women. East Africa has a particularly high risk of HPV-induced cervical cancer. Via early detection and treatment, cervical cancer is highly preventable. Screening based on HPV is now the primary WHO suggested screening strategy for low-resource settings, but widespread implementation is hindered by cost and availability [32].

### 2.2.2 Oral Cancer

The scientific health community has drawn attention to a dramatic rise in the number of cases of head and neck cancers in both the U.S. and internationally. With regard to the spread of HPV, the public's impression may be that this problem falls to doctors, and it is possible to overlook the link to dental professionals. Whenever a patient comes into their office, dental practitioners have access to oral cancer tests [33].

Infections with HPV are responsible for a large part of cancer's global burden. Epidemiological studies have shown rising worldwide trends in HPV-related oral cancers. To be able to offer accurate advice to their patients, dental professionals need detailed, up-to-date HPV-related information. Globally, HPV-attributable oropharyngeal cancer accounts for around 4.5 percent of all malignancies, with an estimated 630,000 new cases identified each year. The causes of oropharyngeal cancer have historically been linked to modifiable risk factors such as smoking or drinking alcohol. Recent research has linked oropharyngeal cancer with other risk factors, such as infection with HPV [34].

The high oral cancer-related mortality rate is related to the late presentation of a significant proportion of advanced disease patients. Thus, in order to obtain a better outcome in patients, early diagnosis tends to be of primary importance. While oral cancer is an easily

accessible location for clinical review, lack of knowledge precludes early detection of precancerous and early cancer lesions in both patients and health care professionals [35]. The vast majority of patients in the clinical community have correctly defined the tongue and floor of the mouth with respect to the potential anatomical locations for oral cancer [35]. HPV may be known as a particular oral cancer causative agent and may be a possible risk factor for oral cancer [36]. A subset of oral carcinomas is associated etiologically with HR-HPV infection, with the most common form of HR-HPV found in these carcinomas being HPV-16 [36].

### 2.2.3 Other HPV Related Cancers

Other cancers associated with HPV include anal cancer, female vulvar and vaginal cancers, and male penile cancer [5].

**Anal Cancer**

There are an estimated 27,000 new cases of anal cancer worldwide annually, with a female to male ratio as high as 5:1. Depending on the anatomical region involved, the appearance of the disease varies; often HPV infected people do not have genital warts or other symptoms of infection. During an active HPV infection, there are different pathologies that may be present; these may include: common warts, plantar warts, and flat warts. If there are genital warts, then there are several different potential appearances and several numbers or single numbers may be present [37]. Anal cancers have many parallels to cervical cancers, including, but not limited to, both squamous cell cancers and related forms of HR HPV. Due to the comparable rates of cervical cancer prior to the implementation of the screening protocol in combination with the existing rates of anal cancer in sexual intercourse, there is an increased understanding that screening HR groups for anal cancer precursors will be advantageous [38].

In both sexes, HPV infection is very prevalent in the perianal area and the anal canal. The highest prevalence of anal HPV (almost 100 percent) is observed in males who have sex with HIV infected patients. Bleeding, discomfort or mass sensation occur in the majority of patients with anal cancer [39]. The anal canal has a transition zone that is vulnerable to dysplasia from HPV infection, close to cervical cancer. Awareness and knowledge of anal cancer and its connection with HPV in HIV-positive women is largely unknown,

considering the increasing incidence of anal cancer in HIV-positive individuals [40]. With increasing severity of lesions, anal cancer risk increased and is particularly high among HIV-positive people [41].

**Vulvar, Vaginal and Penile Cancers**

HPV is an oncogenic virus that is related to the development of many cancers in humans. Primary adenocarcinomas of vaginal, vulvar and penile cancer are uncommon and have rarely been shown to be associated with HPV infection to date, and certain cancers occur less often [42]. The reasons for the rise in penile and vulvar cancers are not well known, but there is an increase in the number of HPV infections [43]. Vulvar cancer is a rare female genital tract tumor which accounts for approximately 5% of all gynecological malignancies. In younger women, HPV-related cancer is more frequent and is often followed by a pre-malignant lesion, i.e. vulvar high-grade intraepithelial squamous lesion [44]. A squamous intraepithelial lesion closely associated with the development of invasive carcinoma is characterized by vulvar cancer. [45].

HPV is associated with the majority of vaginal cancers and a smaller proportion of vulvar cancers. The incidence rates of vulvar cancer in younger women have been identified in different studies, likely due to an increased prevalence of types of high-risk HPV. In comparison, vaginal cancer incidence rates, despite the higher HPV-attributable fraction compared to vulvar cancer, were lower and more stable [46].

In the development of penile cancer, HPV infection tends to play a significant role. Penile carcinoma, which accounts for <1% of adult male cancers, is a rare malignant tumor [47]. Penis carcinoma is a rare neoplasm in developed countries, but in the developing countries of Africa, Asia and South America, the prevalence of this malignancy is much greater. Penile carcinoma is found more commonly in men between the ages of 50 and 70, and any male may be confirmed to be affected. Early stage treatment of penile cancer is important for successful long-term outcomes and maintains the quality of life mainly [48]. The exact cause of penile cancer is unclear, but there are many contributory factors, including age-increasing HPV infection, smoking and immunodeficiency [49].

## 2.3 Types of HPV

There is a number given to each HPV virus, which is called an HPV type. Since certain HPV types cause papillomas (warts), which are non-cancerous tumors. Over 100 different forms of HPV are present, of which at least 14 are cancer-causing [5, 50]. Each HPV has a number or type of its own. The word "papilloma" refers to a type of wart that occurs from certain forms of HPV. Two high-risk (cancer-causing) and low-risk (wart-causing) forms of HPV [50] may be divided into HPVs.

**High-risk**

High-risk strains of HPV include HPV 16 and 18, which are responsible for about 70% of cervical cancers. HPV types of 6, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68 and a few others are other high-risk HPV [50, 51].

**Low-risk**

Around 90 percent of genital warts that seldom grow into cancer are caused by low-risk HPV strains, such as HPV 11. Such advances can look like bumps. They're shaped like a coli-flower, sometimes. The warts may appear weeks or months after intercourse with an infected partner has occurred [51].

## 2.4 Deep Learning

For larger datasets, machine learning fast for model training and model testing it takes a long time to process. The acquisition of data is one of the most difficult problems when using machine learning. Furthermore, when we collect data from medical centers, it is possible that it will contain a large amount of bogus and incorrect data. They frequently encounter situations in which they discover an imbalance in data, resulting in poor model accuracy [13, 14].

Deep learning is a subfield of machine learning that employs a neural network model and bases its learning on a data representation algorithm rather than a simple job algorithm [52]. The use of neural networks has increased at a quicker rate than ever in the recent decade, owing to the availability of many powerful processors and a large amount of data (inexpensive processing units such as GPU). The Artificial Neural Network has one or more processing layers (ANN). The amount of layers we utilize in the network defers depending

on the problem, which we want to address. If there are only two or three layers in the network, it is referred to as a shallow model. Because the ANN model has a large number of layers, the network is dubbed a deep model, and deep learning refers to this deep NN model [53]. Multilayer networks have been around since the 1980s, but for a variety of reasons, they haven't been utilized to train a neural network with several hidden layers [54]. The curse of dimensionality was the main problem preventing the use of multilayer networks at that time, i.e. if the number of dimension features increases, the amount of configuration increases. The number of data samples for training increases exponentially as the number of configurations increases. The selection of adequate training datasets was thus time consuming and the use of storage space was not cost-effective [54, 55]. Most neural networks today are sometimes referred to as deep neural networks and are widely used. As a large amount of data, as well as storage space, and computer resources are available, we can train a neural network with many hidden layers. Before the machine learning level, the conventional machine learning algorithm requires separate hand-tuned feature extraction. Only one neural network process has deep learning. The layers learn to identify the fundamental characteristics of the data at the beginning of the neural network and to feed data to the other layers of the network for additional network computation [54]. ANN is influenced by the human brain. Computer vision is one of the key applications of deep learning, inspired by the human visual system. In the last two decades, deep learning has yielded great success in computer vision and speech recognition [52]. Deep learning models have also been used in a wide range of problem areas, including text detection, speech recognition (natural language processing), visual object recognition (computer vision), object detection, and a variety of other fields such as drug discovery and genomics [55, 56]. The quantity and type of problems that a neural network can tackle are based on numerous deep learning methods created during the last two decades. Deep learning models that are widely utilized include recurrent neural networks (RNN), long short-term memory (LSTM), convolutional neural networks (CNN), deep belief networks (DBN), and auto encoders.

The RNN was one of the earliest deep learning models to lay the groundwork for future deep learning algorithms. It is frequently used in speech recognition and natural language generation [57]. The purpose of RNN is to identify the data's sequential features (remembering previous entries). While we evaluate time-critical data, the network contains

memory (hidden state) to store previously examined data. The main disadvantage is that RNN must use current information (short-term reliance) to complete the current objective. RNN varies from a neural network in that it takes a set of data and processes it over time [57].

An LSTM is a form of RNN designed to tackle the long-term dependency problem by reminding the model of values across time intervals of any length. Gradient disappearance and gradient bursting are two major RNN issues. The gradient is the relationship between the change in weight and the change in error. It is well suited to processing and predicting time series with indeterminate lengths and temporal delays. If we want to forecast a thousand intervals instead of ten, RNN forgets the model, whereas LSTM remembers such actions. The fundamental reason LSTM can remember its input for so long is that it has a memory that acts similarly to a machine memory, allowing knowledge to be read, written, and deleted by LSTM [58, 59]. Text recognition, handwriting recognition, voice recognition, gesture detection, and image captioning are the most common applications.

CNN is the common deep learning model for various computer vision tasks, especially for image recognition. It is a network of multilayers and is inspired by the visual cortex. CNN is used for the detection of cancers caused by HPV in our thesis, and the details are given in Section 2.4.1.

DBN is a class of deep neural networks with several hidden layers in which each layer of the network is linked to each other, but not bound to each other by the neurons in the layers. DBN preparation happens in two steps. It consists of layers of Restricted Boltzmann Machines (RBMs) for the supervised fine-tuning process for the unsupervised pre-training and feed forward network. It learns a layer of operation in the input layer during the training of the first step (pre-training). The fine-tuning process starts after the pre-training is done. It accepts the features of the input layer as input in the fine-tuning process and learns features in the second hidden layer. To train the entire network, including the final layer [60], back propagation or gradient descent is then used. DBN is used to recognize images, retrieve information, understand natural languages, and recognize video sequences.

Auto Encoders are a type of neural feed-forward network that's designed for unsupervised learning, or when the data isn't labeled. The inputs and outputs of auto encoders are

identical. The input is taken and compressed into a lower-dimensional code, which is then rebuilt from the compressed code's output. The encoder, the code, and the decoder are the three components of an auto encoder. The encoder accepts the input and produces the output, whereas the decoder creates the output using the code. One of the most prevalent auto encoder applications is anomaly detection [56].

Deep learning techniques are cutting-edge and evolving at a rapid pace. Deep learning today outperforms other machine learning algorithms thanks to the availability of a large amount of data and high-performance computing system components such as GPU [61]. Deep learning techniques, in contrast to traditional machine learning approaches, use multi-layer (too many hidden layers) processing for improved performance accuracy, and there is no explicit extraction of features, i.e. features are automatically extracted from raw data in deep learning models, and We are capable of feature extraction and detection (it may be recognition depending on our problem) Numerous writers have demonstrated that deep learning can attain state-of-the-art performance for a variety of challenges that artificial intelligence and machine learning have faced for a long time in the fields of computer vision, natural language processing (NLP), and robotics [61, 62]. To address the model's difficulties in learning complicated qualities, deep learning techniques employ back propagation algorithms, loss functions, and a high number of parameters.

### 2.4.1 Convolutional Neural Network

A convolutional neural network (CNN), often known as a convnet, is a multilayer deep learning model that is similar to the feed-forward NN that is commonly used to assess visual imagination. It is similar to ordinary neural networks; however it is specifically built for computer vision challenges. CNNs are derived from the conventional neural network, which is commonly used in jobs requiring recurring patterns, such as image recognition [61]. Because images include a lot of information most of the time, there is a dimensionality problem in a traditional neural network for image processing or computer vision applications. For example, a grayscale image with a dimension of 1280 by 720 pixels has 921,600 pixels. If the pixel intensity of this image is detected as an input by a totally linked network, the weight required by the neuron is 921,600. For example, a 1920 by 1080 picture will necessitate 2,073,600 weights. If the picture is colored, the color amount is multiplied

by three (polychrome). As a result, as the picture size grows, so does the network's number of free parameters. As a result, as the model grows increasingly complex, the network's performance deteriorates, resulting in overfitting [63]. Overfitting is an issue with machine learning techniques when the size of the network grows and there is no data to match the model. The issue restricts the machine learning model's ability to generalize. CNN overcomes this issue by providing layers of neurons that are organized in three dimensions (3D): width, height, and depth. Each CNN layer receives a 3D volume of input data (in this example, an image), and a separate function outputs another 3D volume of data as an output [64].

CNN's core concept was inspired by the receptive field, a scientific term [61, 62]. Edges, for example, are receptive fields that are responsive to a stimulus and are found in sensory neurons. The term receptive field is extensively used in the context of ANN, most notably in relation to CNN, where biological calculations in computers using convolution processes are modeled using convolution processes. In computer vision, the use of convolution procedures to achieve various observable effects will filter images. CNN use convolution filters to distinguish specific elements in a photograph, such as edges that resemble the biological receptive field. CNN has achieved notable success in handwritten digit detection and facial recognition since the late 1980s and early 1990s [65].

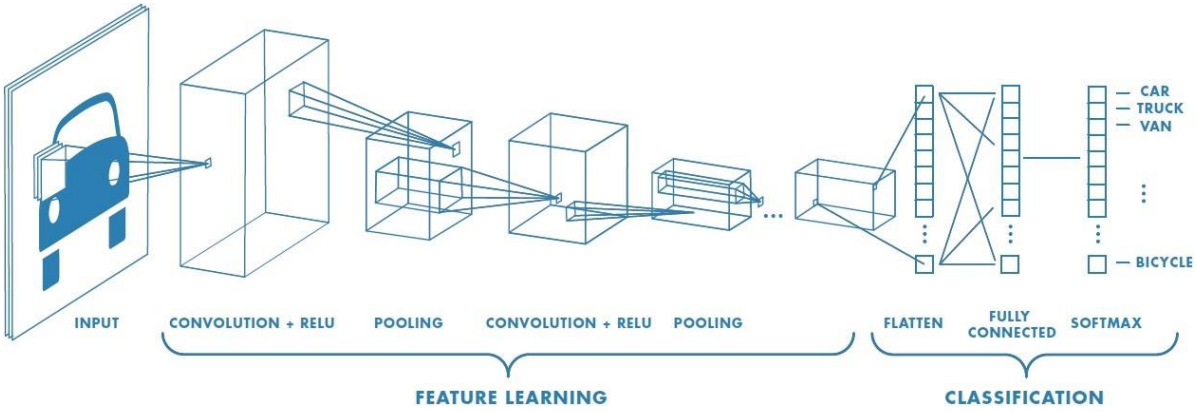


**Figure 2.1:** *CNN Model Example*

As shown in Figure 2.1 [66], CNN is a machine learning algorithm that uses contaminated and healthy images as input to detect and recognize disease. For the detection procedure, CNN is employed, which is made up of several successive layers, each of which changes

one activation volume into another using a different function [67]. The basic and extensively utilized layers of CNN are the convolution layer, the pooling layer, and the totally linked layer [54].

**Convolution Layer**

The basic goal of the convolution layer is to extract relevant properties from the input image. Every image on a computer is represented as a pixel value matrix. A basic digital camera produces three channels: red, green, and blue (RGB). This image is made up of three 2D matrices (one for each color) piled on top of one another, each with a pixel value ranging from 0 to 255. The convolution layer consists of a number of convolutionary filters (also known as kernels or feature detectors) with modest matrix values such as 3 x 3, 9 x 9, and so on [67]. The filters are thought to be neuron parameters that can be learned. Because each filter is lower in dimension (width and height) than the input volume, it widens the depth to the same extent (input image). A normal filter, for example, would measure 5 x 5 x 3 inches (5 widths, 5 heights, and 3 depths for the three-color channels). Because linking all of the pixels in the image would be too expensive to compute, portion 1 of the image is linked to the next convolution layer. Convolution is performed by shifting the filter from left to right across the width and height of the input picture and measuring the dot product at each point between the filter and the input image. The procedure's result is known as a feature map (aka convolved feature or activation map). After that, the filters are applied to the input image in order to extract meaningful information. The retrieved features or the function map will vary if the filter parameters are modified. In the example below, we created a 2D input picture with kernel sizes of 5 x 5 and 3 x 3 (Figure 2.2) [67].

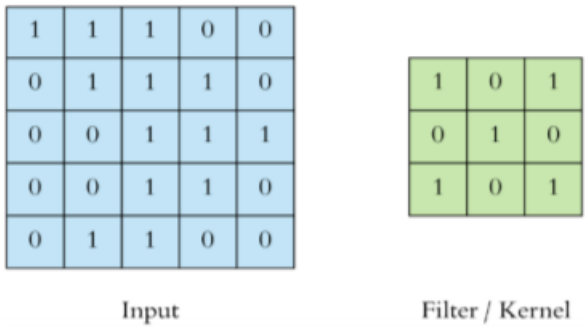


**Figure 2.2:** *Examples of Input Volume and Filter*

The input and filter are provided; the convolution process is then performed by sliding or convolving the filter over the input. At each place, the dot product is calculated (by multiplying the element wise matrix and summing the result) and the result is saved in a new matrix called the function map (Figure 2.3). These results are added to the feature map since the output of the first convolution operation is 4 and the output of the second is 3, as seen in the accompanying image. The entire process is completed by shifting the filter to the right and adding the result to the feature map.

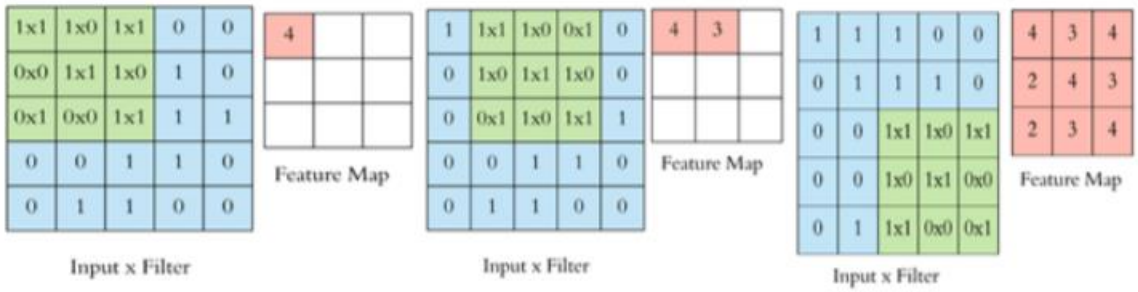


**Figure 2.3:** *Examples of Convolution Operation*

The receptive field, seen in Figure 2.3 [67], is the area where the convolution operation is carried out, and it is 3 3 in size since the filter size is always the same. We use multiple filters to execute as many convolution operations on the input as we can, resulting in a range of feature maps. Finally, the convolution sheet is complete when all of the function mappings are piled together (nadre.ethernet.edu.et). Depth, stride, and padding are three hyper-parameters that control the size of the output neuron (nadre.ethernet.edu.et) (feature map) (aka zero paddings). These parameters must be specified before the convolution process can begin [67].

- The depth refers to how many filters are used in the convolution process. The more filters we use, the better the model we construct, but the larger the parameter count, the greater the risk of overfitting. We can build three different feature maps if we employ three different filters throughout the convolution process. These feature maps were eventually stacked as 2D matrices, resulting in three detailed feature maps.
- Stride is the total number of pixels that the filter glides over the input volume at one time. When the stride is 1, the filter matrix slides 1 pixel at a time to the input volume.

When the stride is 2, the filter hops 2 pixels at a time to the input volume, and so on. The volume of production is lower when the number of steps is greater.

- Padding in the input volume adds zeros around the edges. It's simple to pad the input volume across boundaries with zeros. It allows additional data to be held beyond input borders and the size of the function map to be controlled.

CNN typically uses hyper parameters with a size of 3, stride with 2, and padding with 1, but depending on the input volume we have, we can change this hyper parameter [67]. Because the matrix only has one depth for grayscale images (Figure 2.3), these convolutions are done in 3D in this work because color images acquired with a digital camera are represented as a 3D matrix with width, height, and depth dimensions (the depth represents the three color channels). For example, with a 6 x 6 x 3 input and a 3 x 3 x 3 filter size (input depth and filters stay constant), we may do convolution, with the only change being that the total of matrix multiplication is 3D rather than 2D, as illustrated in Figure 2.4 below.

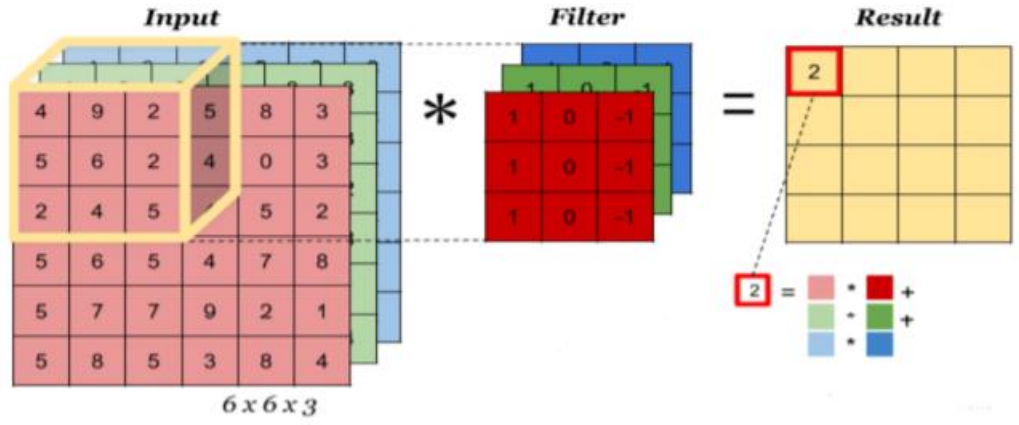


**Figure 2.4:** Examples of *Convolution of a 3D Input Volume*

Figure 2.4 [68] depicts a 6 x 6 x 3 input volume and a 3 x 3 x 3 filter [68]. Both the phase and the filter number are one. It just travels one pixel at a time. In the convolution layer, they may employ a variety of filters to detect a range of characteristics, and the output of the convolution layer will have the same number of channels as the number of filters. The figure below (Figure 2.5) is identical to Figure 2.4, but with two filters [68].

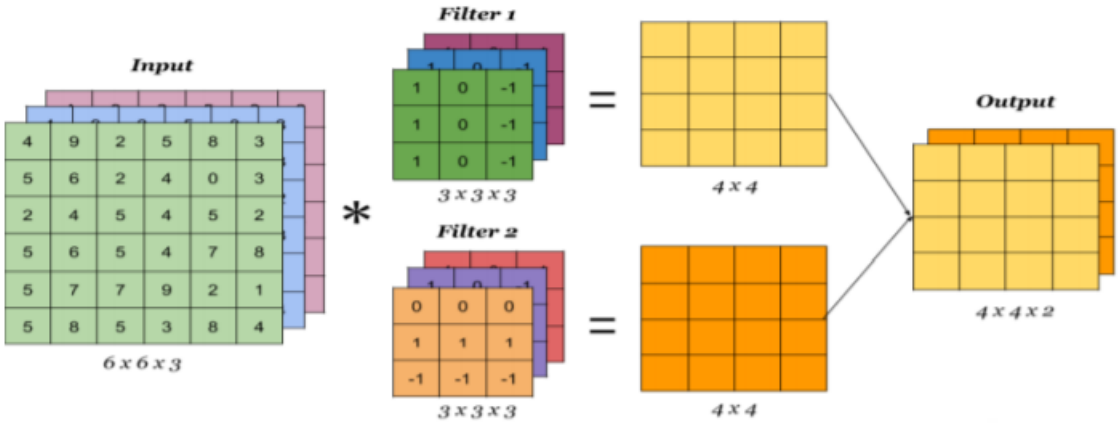


**Figure 2.5:** *Examples of Convolution Operation with 2 Filters*

As shown in Figure 2.5 [68], the number of filters equals the depth of the function map. To regulate the amount of free parameters in the convolution layer, a systematic approach called parameter sharing is used. If a function is effective for computing a specific spatial location, it should be beneficial in other situations as well. If we employ the same filter (often referred to as weights) in all portions of the input volume, the number of free parameters reduces. In the convolutional layer, the neurons share their parameters and are only coupled to select regions of the input volume (local connectivity). Convolutional parameter sharing adds to CNN's translation invariance; for example, if the input volume has a specific orientated structure and they want CNN merely exchange the parameters and identify the layer connected locally to learn various characteristics in different spatial positions [67]. To build a single convolution layer, they must apply the activation function (ReLU) and bias (b) to the output volume. Figure 2.6 depicts a single CNN convolution layer with the ReLU activation feature [68].

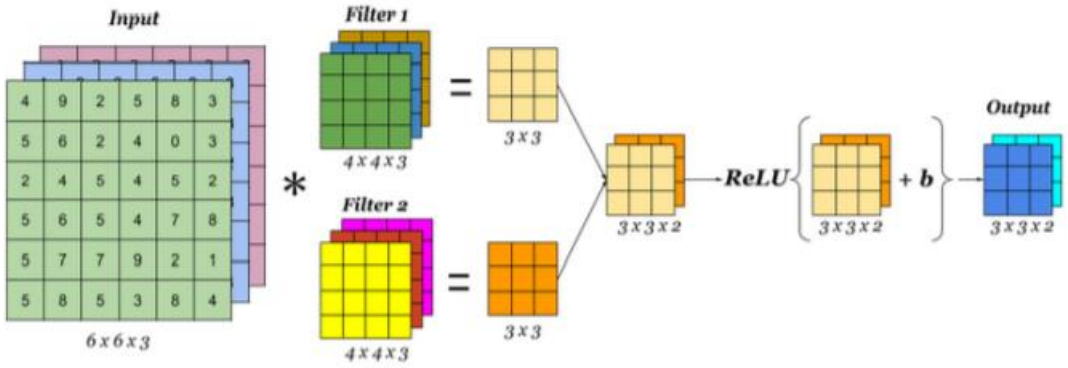


**Figure 2.6:** *Examples of One Convolution Layer with Activation Function*

**Pooling Layer**

Pooling layers (also known as subsampling or down sampling) are used in CNN to minimize the number of parameters, extract dominating features in certain spatial regions, progressively lower the spatial size of the convoluted feature, and monitor the overfitting problem in the network [67]. This layer contributes to the reduction of computer resources required for network training. The pooling action is carried out by sliding the filter across the converted feature.

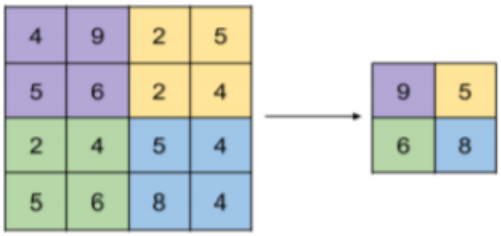


**Figure 2.7:** *Max Pooling Example*

As shown Figure 2.7 [67], there are three types of polling: maximum pooling, average polling, and the sum pooling method, which is less widely used. The most widely used polling operation is max polling (Figure 2.7), and its output is the maximum value of the portion of the image protected by the filter. Average pooling returns the average of all the image values covered by the filter and, essentially, the sum pooling returns the sum of all the image values covered by the filter. The max-pooling conducts de-noising along with reduction of dimensionality, but average polling is only used to reduce dimensionality [67].

Maximum polling, therefore, is higher than average pooling. After the convolution operation, the pooling operation is implemented in all the depth slices of the image and the widely used filter is 2 x 2 and strides 2, but we can adjust accordingly. For instance, if we take the widely used 2 x 2 filter (as shown in Figure 2.7); it returns the maximum value of the four values for the max pool [67].

**Fully Connected Layer**

A fully connected neural network is composed of a set of layers that are fully connected. Every neuron in the previous layer is connected to every neuron in the next layer in a completely connected layer. This layer embraces the output of the convolution or pooling layer, which is the input volume of high-level features. These high-level features are in the form of a 3D matrix, but a 1D vector of numbers is approved by the completely connected layer. Therefore, the 3D data volume needs to be transformed into a 1D vector called flattening, which becomes the input to the fully linked layer. The flattening vector is given to the completely linked layer and, like any ANN, it performs mathematical computation. To apply non-linearity in these layers, activation functions such as ReLU in the hidden layers are used. The last layers (output layer) of the completely connected layer perform detection (probability of inputs being in a specific class) based on the training data class by using the sigmoid activation function for binary class and by using softmax activation function for more than two classes.

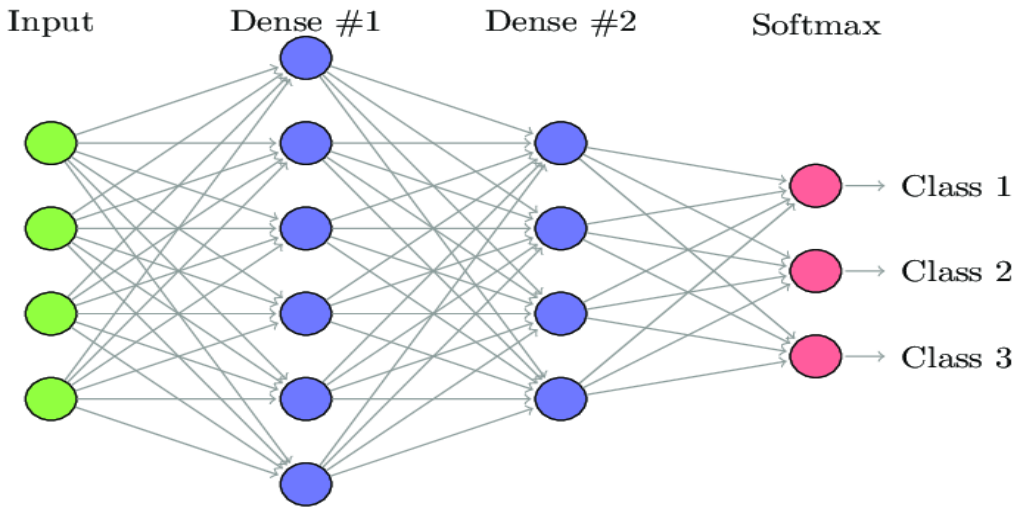


**Figure 2.8:** *Fully Connected Layer Example*

As shown in Figure 2.8 [69], a better way to learn non-linear characteristics of the output returned from convolution and pooling layers is a fully connected layer. Convolutionary networks which often not have completely linked layers and are called the Fully Convolutionary Network (FCN).

### 2.4.2 Application of Deep Learning in Cancer Detection

In various computer vision applications, Deep CNN is applied to solve different problems. Image processing methods can be used to improve clinical procedures and diagnostics by enhancing process quality and consistency, while minimizing manual monitoring by patients and experts. In the following, we see literature that uses CNN models via image processing for the detection of disease. Many writers use the techniques of image processing with the convergence of machine learning techniques to solve various health problems. A summary of the applications is discussed here.

Many cases in which the AI system correctly identified cancer were discovered by contrasting the errors of the AI system with errors from clinical readings, whereas the reader did not and vice versa. Most of the cases in which cancer was detected only by the AI system are invasive. The authors in [70] present a system of artificial intelligence (AI) that can exceed human experts in the prediction of breast cancer. A deep learning model was developed and tested using two large datasets from the UK for the detection of breast cancer in screening mammograms, totaling 25,658 images and 3,097 images from the USA. The authors used the CNN algorithm's deep learning model to classify the input images into two classes, namely healthy and unhealthy, and finally, the new method achieves 88 percent precision.

Another deep learning technique suggested by the authors in [71] uses a deep CNN to detect liver cancer using watershed transformation and techniques of the Gaussian mixture model. A total of 225 CT scan images of liver cancer were obtained from the imaging centers in this study and hospitals were collected for training and model research. The proposed algorithm is validated on the basis of a real-time clinical dataset collected from various scientifically developed patients. The key benefit of this automated detection process is that it uses a deep neural network classifier to achieve the best accuracy with negligible validation loss.

Finally, with a better precision of 99.38 percent, the model effectively classified 2 different groups, i.e. infected and uninfected with disease.

The author [72] developed a CNN model using deep learning to identify skin cancer diseases. The datasets are gathered from globally accessible datasets of 5,161 total images gathered for model training and research. Finally, with a higher accuracy of 95 percent, the model effectively classified 2 different groups, i.e. safe and unhealthy illnesses.

# Chapter 3: Related Work

## 3.1 Introduction

Discussing the applications of deep learning and image processing over health in the previous section, applications relating to the detection of cancer diseases are listed here in this section. Several authors have proposed and/or developed diagnostic and detection systems for cancer diseases using various techniques, including deep learning and image processing techniques.

## 3.2 Machine Learning

Kelwin Fernandes *et al.* in [73] used the approach of machine learning and developed an automated diagnostic model using the acetic acid method for the detection of cancer. In the last decade, the development of computer-aided diagnostic systems for the automated processing of digital colposcopes has attracted the attention of computer vision and machine learning groups, giving rise to a wide range of tasks and computational solutions. They used color, texture, edges, discrete wavelet transformation, spatial information as the features for static images and the changes in temporal ace to white response in the pre and post application of acetic acid for sequence-based recognition. They did only for cervical cancers detection but pre-cancer detection for cervical and other related cancers, the current stage and the type of HPV to lead cancers is not considered.

Sara Tous *et al.* in [12] debate on the global burden of HPV-related cancers, which appears to be a significant cause of cancer in both men and women. To determine the possible effect of HPV on the reduction of the burden of HPV-related disease and to help formulate guidelines on HPV prevention, the distribution of the HPV type and the burden of cancer cases due to the types of HPV included were estimated. By geographical area, sex, and age at diagnosis, they found little heterogeneity. Only cervical and oropharyngeal cancers have reported major geographic variations in HPV. A significantly higher contribution was observed for cervical cancer in Asia and Oceania, and a significantly higher contribution was observed for oropharyngeal cancer in Europe than in the Americas. As a result, one of the greatest public health goals and challenges has been identified to provide the majority of women around the world access to HPV identification and cervical screening. They suggests

that to increase decision making regarding to HPV related cancers detection techniques should be implemented.

## 3.3 Cervical Cancer Detection

R. Gorantla *et al.* [13] presents a deep learning approach to fully automated methodology called CervixNet for Cervical Cancer Screening using cervigrams is introduced in the work. In developing countries, routine monitoring for HPV in women has helped minimize the death rate. However, because of the shortage of affordable medical services, developed nations are also struggling to have low-cost solutions. In order to provide adequate care to save the patient from the clutches of excruciating death, early detection in cases of cervical and other associated cancers is crucial. They proposed a fully automatic methodology known as CervixNet for Cervical Cancer Screening using cervigrams with Accuracy of 86.77%. The key contribution of the paper is to implement the algorithm of Image Enhancement to enhance cervigrams and evaluate the HPV caused cancer detection model of the Hierarchical Convolutionary Mixture of Experts (HCME) in the detection problem of the cervix form. In addition, they use cross entropy to design a new loss function to improve the efficiency of the detection. Because of the limited datasets, the HCME model overcomes the issue of over fitting. They believe that this model will generalize well to other detection tasks, giving it broad applicability in the biomedical imaging field. Their methodology outperformed the latest state-of-the-art approaches when testing Intel and Mobile-ODT cervical cancer screenings on an open challenge database. The results show that CervixNet is robust for different noisy images and conditions for acquisition of images. They recommended that the creation and implementation of deep learning should be added over specified periods of time for future work. Also recommended is a theoretical guide to possible reviews. So, early detection of cancers and other HPV caused cancers are not considered.

Matthew Wilhelm *et al.* [14] are creating an automated model for HPV detection in images captured from the Linear Array HPV Genotyping Test of Roche Molecular Diagnostics. For 37 different forms, this test provides type-specific HPV genotype results with different levels of risk for cervical cancer growth. The algorithm was checked by 17 patients and found that only five of these possibilities were actually forms of HPV after testing the

method. The diagnosis of patient 17 was described as types 2, 3, 6, 10, and 22. They propose that the creation of a medical database to store the resulting data and to monitor the changes in the records of individuals over many years would be beneficial for researchers and medical experts. This database could be used to create clinical algorithms that would recommend behaviors based on prior patient results and the test data is too limited and not all types of images are included. There are more than 100 different types of HPV contributing to cervical, dental, vaginal and anal cancers, but only 37 types are detected and others HPV-related disease are not included it is only for cervical cancer with a small dataset.

Anish Simhal *et al.* [74] present significant morbidity and mortality worldwide was correlated with current cervical cancer. It is well known that high-risk HPV strains are one of the major causative agents for cervical cancer, and this form of malignancy is preventable. With the absence of HPV screening and low public awareness of the issue, the high incidence of cervical cancer with substantial mortality is proof of HPV infection abundance. A low-technology version of colposcopy is a visual inspection with acetic acid using the naked eye that either adds precision to human papillomavirus testing (when available) or acts as the primary screening instrument. The results indicate that the use of simple color and textural features from visual acetic acid inspection and visual inspection with iodine images of Lugol may provide impartial cervigram automation. This will enable for the automatic and expert detection of cervical pre-cancer at the point of therapy. Visual inspection with acetic acid (VIA) and visual inspection with Lugol's iodine (VILI) cervigrams using image processing algorithms to extract color and texture information. These characteristics are used to train and evaluate a support vector machine that can recognize cervigrams for cervical cancer diagnosis. The algorithm performs far better than the average diagnosis of a professional doctor. Because this algorithm was created using images captured with the Pocket Colposcope, it may be used to cervigrams captured with other colposcopes. The study's significance is that it developed an automated approach for cervical pre-cancer testing utilizing colposcopy images and machine learning, however it does not address other HPV-related cancers, and its accuracy is 80 percent.

Lavana Devi N and P. Thirumurugan [75] presents a deep learning approach was used and the discussion of automated screening is becoming more popular than manual screening

because the latter is incorrect. HPV infection is related to virtually all cases of cervical cancer (95% of cases) and a significant proportion of cases of anal cancer (88 percent of cases attributable to HPV). A varying proportion of vulvar, cervical and penis cancers are also causally associated with HPV. They compare the difference between the acetic acid test before and after it. The area of interest, i.e., has high gray scale intensity in the cervical region, high red color intensity and more focused areas. To segregate the irrelevant sections of the picture and the region of interest, the threshold method of segmentation is carried out. Features such as proportion of area of interest, coarseness of texture, and entropy are defined. These three characteristics are given as input to the model of fuzzy reasoning. The performance of the fuzzy method of reasoning gives the abnormality severity. The texture characteristics gave high specificity to the shift in temporal gray scale. The overall detection efficiency was enhanced by integrating both texture and temporal gray scale change. The detection results were eventually compared. Automated benign and malignant cervical cell detections minimize false negative and false positive cases, thus increasing the overall system's effectiveness. The advantage of the study is design an automated system for cervical cancer screening but other HPV-related cancers and pre-cancer detection is not considered.

## 3.4 Deep Learning

Mihalj Bakator and Dragica Radosav [76] cervical cancer is caused by HPV, which contributes to abnormal cell proliferation in the cervix area. Researchers employed a deep learning approach to solve the problem. Routine HPV testing in women has helped to reduce the death rate in underdeveloped nations. When it comes to deep learning and its application in medical diagnosis, there are two main approaches. The first is a detection method that entails reducing the number of possible outcomes (diagnostic) by mapping data to actual results. The second method involves combining physiological data, such as medical pictures, with data from other sources to identify and diagnose cancers or other disorders. Deep learning can also be used to assist in nutritional evaluation. When it comes to medical diagnosis, deep learning is certainly used in a variety of ways. In this situation, the theoretical context did not include a full discussion of how deep neural networks work. However, given the scope of the study and the intended audience (researchers whose

expertise does not include deep learning), a theoretical approach was not thought required. It's crucial to have a functional model in place for cancer diagnosis. They also explain deep learning strategies that can be used in the medical industry. Deep learning networks are used to make medical diagnoses in this case. There are features such as detection, segmentation, and prediction. Deep learning approaches can outperform other high-performing algorithms, according to the findings of the studies reviewed. As a result, it's acceptable to say that deep learning has a wide range of applications and will continue to do so.

**Table 3. 1:** *The Summery of Related Work*

| **Authors, Title and Publication Year** | **Method** | **Significant** | **Consequence** |
|---|---|---|---|
| Fernandes, K. et al. Automated Methods for the Decision Support of Cervical Cancer Screening using Digital Colposcopies. (2019) | Machine Learning | Automated diagnostic system for the detection of cancer. | Pre-cancer detection for cervical and other related cancers is not included. |
| Mercy Nyamewaa Asiedu, et al. Development of Algorithms for Automated Detection of Cervical Pre-Cancers With a Low-Cost, Point-of-Care, Pocket Colposcopies. (2017) | Machine Learning | Automated diagnosis of cervical pre-cancer. Accuracy is 81.3% | The main cause of cervical cancer is HPV but not considered. |
| Rohan Gorantla, et al. Cervical Cancer Diagnosis using CervixNet A Deep Learning Approach. (2019) | CNN | Automated method of CervixNet for cervical cancer screening is developed and its Accuracy is 86.77%. | Early detection and classification of cancers and the cause of cervical cancers are not included. |

| | | | |
|---|---|---|---|
| Matthew Wilhelm, et al.<br>Automated Detection of Human Papillomavirus: Via Analysis of Linear Array Images. (2010) | Digital Image Processing | Developed a system for the detection of HPV in images. | They used a very small dataset b/c the algorithm is tested by 17 patients, it's only for cervical cancer and not all HPV types included. |
| Lavanya Devi, et al.<br>Automated Detection of Cervical Cancer. (2019) | Deep Learning | Automated system for cervical cancer screening. | The cause and the phase of cervical cancer are not included. |

# Chapter4: Modeling Detection of HPV caused Cancer

The design of a model for detection of HPV caused cancer is the subject of this chapter. Details of the model, the principles and the algorithms used and developed, the evaluation metrics, and their related representation are presented in this chapter. In particular, descriptions of the HPV caused cancer detection model, how features are extracted and detection is carried out in the HPV caused cancer detection model and other pre-trained models are described.

## 4.1 Model Selection

A CNN deep learning algorithm is chosen based on different literature that has been carried out in computer vision, especially in the detection of diseases from images. CNNs represent an interesting approach for the processing of adaptive images. The algorithm is used for the extraction, detection, preparation, checking of features and for assessing the model's accuracy. CNN takes raw data without the need for separate pre-processing or feature extraction step.

The key benefit of using the CNN algorithm to detect cancers is that it is more robust and automatic than classical machine learning algorithms. There is a need to build various algorithms for different problems in the classical machine learning algorithm, because it uses more handmade algorithms. But, once we have developed an algorithm for the detection of cancers caused by HPV in CNN, it can be extended to other similar diseases. So, it is simpler to generalize and use as it has different but related problems. In our work, some of the key explanations that CNN would be used are:

- Previous research has demonstrated that CNN outperforms other detection methods and it is the state-of-the-art in computer vision.
- CNNs are better than other deep learning models for image-related tasks since they are designed to simulate human vision and understanding.
- Prior to prediction, most traditional approaches to machine learning need explicit extraction of the attributes used to examine the image.
- Most neural network algorithms only accept vectors (1D), however most real-world images are tensors (3D), hence the original image (input) must be flattened to a 1D

vector, which is time-consuming and expensive to compute, whereas CNN allows 3-dimensional images, and

- With the application of appropriate kernels, CNN can capture temporal and spatial dependencies.

## 4.2 Overview of the Model

Figure 4.1 depicts the HPV caused cancer detection model. The training dataset, preprocessing, feature extraction, proposed CNN model, model training components, model assessment, and detection model are all included in the model. The sub-tasks of the preprocessing phase are data cleaning, data augmentation, and image size reduction. During the network's training process, the CNN algorithm accepts the size of the normalized images in the training and testing datasets. More training data for the CNN model is generated using this strategy. The sections that follow provide a full description of each HPV-caused cancer detection model component.

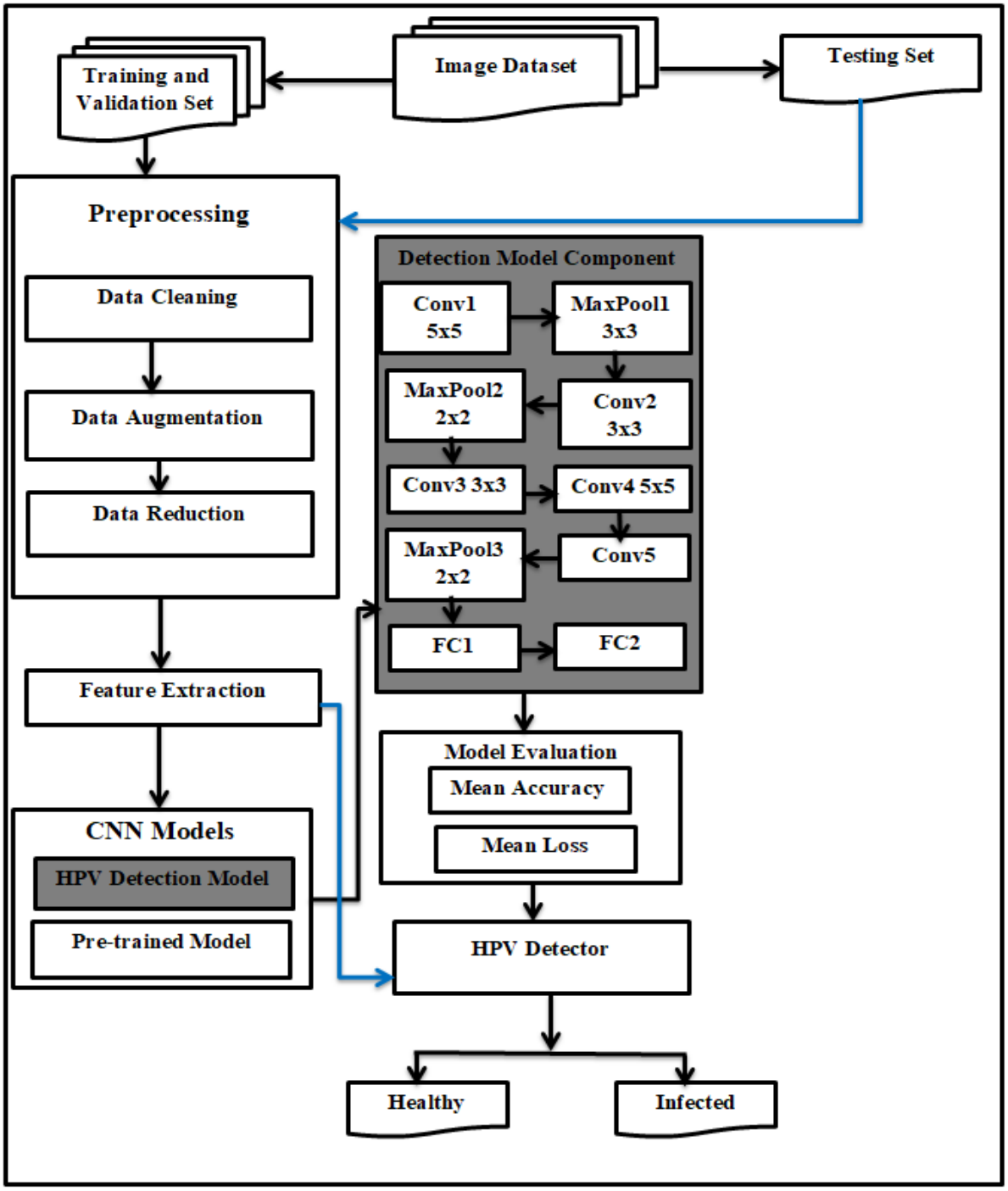


**Figure 4.1:** *The HPV caused cancer detection model*

The model calculates performance measures based on the training dataset and the original testing data. Within the CNN detection, a beneficial characteristic of each image is retrieved and done on the basis of the extracted function, a process known as model training. We access the model's performance during training by using the test dataset, which is essentially

utilized to generate the model's output. The model that performed better is saved and utilized as a predictive model after obtaining the model's outcomes. The testing procedure is then carried out by feeding the predictive model with unseen images as it is being prepared. Finally, the model provides detection, which is the probability that an image belongs to one of the classes learned during the training (in our case there are two classes the cancers caused by HPV i.e. called infected and cancers not caused by HPV i.e. called healthy).

### 4.2.1 Data Collection and Dataset Preparation

HPV related cancer images are collected from two hospitals in Addis Ababa: Betezata and St. Paul's Hospital. In addition to the images collected from these hospitals, other samples (images of healthy and contaminated HPV-related cancers) are gathered from various sources on the internet with the support of medical experts. We have collected 4,322 images for HPV caused oral cancer and 6,734 images for HPV caused cervical cancer which is a total of 11,056 images before augmentation for the two HPV caused cancers namely cervical and oral cancer the categories of infected and healthy. The model benefits from this by training with various imaging properties and conditions.

Manually categorized and labeled images in the training set and randomly chosen, unclassified and unlabeled image data in the testing set are prepared from the collected image data. The images in the test set vary from the images in the training set that are used. From the collected total images, a ratio of 70 percent and 30 percent are used for training and for testing respectively.

### 4.2.2 Data Preprocessing

#### 4.2.2.1 Data Cleaning

The key reason why we need data preprocessing comes from the fact that the detection algorithm can be ambiguous about redundant data and irrelevant characteristics and this can lead to unreliable and less generalizable models. Data obtained from various sources may usually not be in a machine-friendly format, may contain invalid and/or missed values, and may be too large for image processing. For both training and testing, machines need precise representations of the input data that are called features before feeding a deep learning algorithm. To prepare the data for preprocessing tasks in that format, such as data cleaning,

image enhancement for visual ease, noise removal for process consistency, image scaling and resizing or shrinking for rapid computation to reduce complexity, using filter functions and other essential techniques required to improve the quality of the input image.

#### 4.2.2.2 Data Augmentation

Data augmentation is the process of producing extra data from current training samples to increase the amount of training data points in a dataset. It enables the network to learn more complicated data features while avoiding overfitting [64].

In order to obtain more images for our dataset, we performed various data augmentation techniques on the original images in our research. We can execute data augmentation before feeding the data into our model (offline augmentation) or during the training process. Data augmentation is carried out during the training of the network in our study using Keras libraries. Any image that was presented to the network during the training and the algorithm is generated by the original image attached hereto as Annex A – Proposed CNN Model Code.

#### 4.2.2.3 Image Size Reduction

The original image generates any image that was provided to the network during training and testing.

The algorithm will take raw image pixels and learn the features on its own; however the images in the provided dataset come in various sizes. Because the model is trained on an ordinary computer with limited hardware resources such as CPU and memory, size normalization is performed in the dataset to produce a comparable size of all images for the CNN algorithm and to reduce the computational time of training. Finally, the algorithm resizes all of the images in the collection to 127 by 127 pixels attached hereto as Annex A – Proposed CNN Model Code.

```
Input: Image

Output: Resized Image

Start:

For each image in the dataset
```

```
   Resized_Image = resizing (image, target_size=127,127)
End for
Return Resized_Image
Stop
```

**Algorithm 4. 1:** *Image Size Reduction Algorithm*

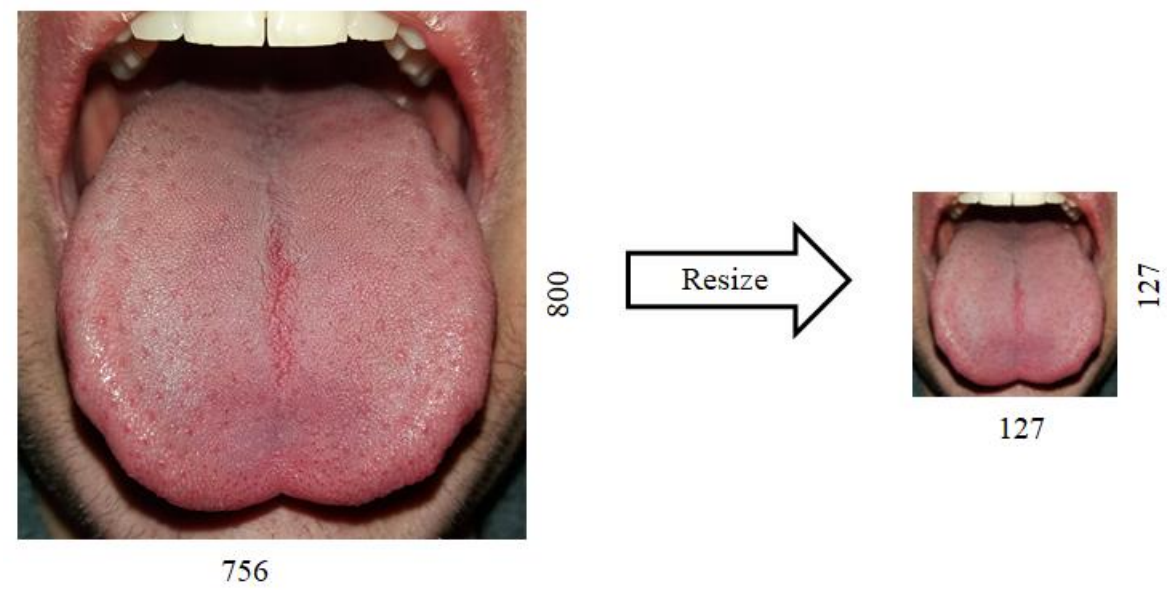


**Figure 4.2:** *Image Resized*

#### 4.2.2.4 Feature Extraction

The CNN method extracts crucial features that are needed to identify the images before the detection stage begins, hence feature extraction is the first step in most image detection problems. For the identification of cancers caused by HPV, one feature parameter is employed, which is a color feature, because the visual color difference determines whether or not the image in the traditional system is impacted by the disease or not by human vision. As a result, the HPV caused cancer detection model generates output (predefined classes) based on the color extraction of the input images that is learned during the training. During CNN training, the network learns which features to extract from the input image. CNN convolution layers extract characteristics, as explained in Section 2.4.1, and the major goal

of this layer is to extract features. The purpose of these layers is to extract local characteristics from the input image using a collection of filters or kernels that can be learned (Figure 4.2).

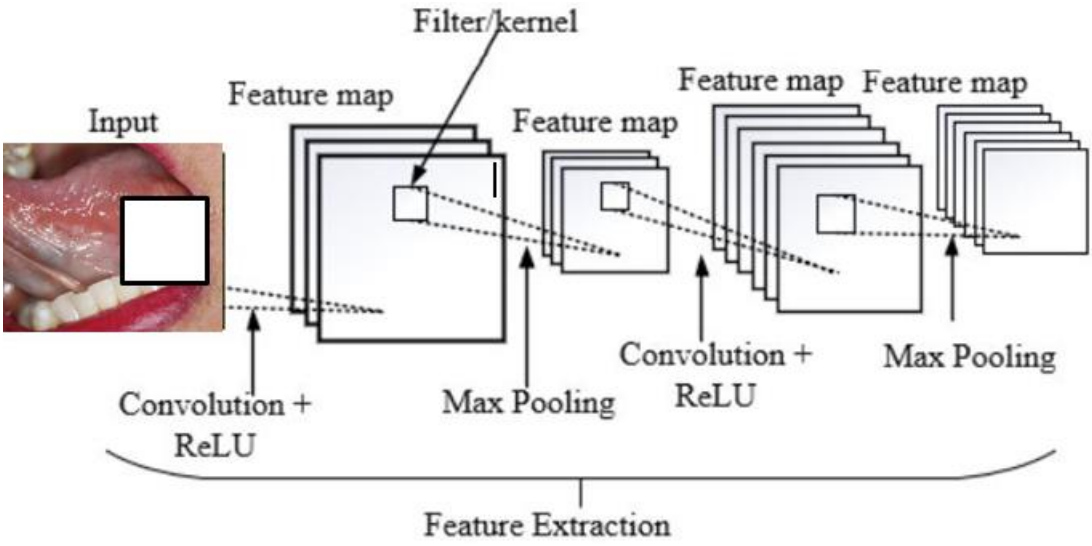


**Figure 4.3:** *In the HPV caused cancer detection model, Feature Extraction* [69]

As shown in Figure 4.3, the filters in the convolution layer slide from left to right across the incoming image entry to detect features. During feature extraction, the convolution layer accepts pixel values from the input image, multiplies and sums them with filter values (a set of weights), and then outputs the feature map. The extracted input image feature is the feature map, which comprises patterns that are utilized to distinguish the images presented. The map ($M_i$) of the feature is computed as:

$$\boldsymbol{M_i} = \sum_{\boldsymbol{k}}^{\boldsymbol{n}} \boldsymbol{W_{ik}} * \boldsymbol{X_K} + \boldsymbol{b} \qquad (1)$$

Where: $w_{ik}$ is an input image, $x_k$ is the input image's $k^{th}$ channel, and b is the bias term.

In this case, the features include the different color patterns of the image given. To add nonlinearity to the network, each value of the feature map is then passed through the activation functions. After nonlinearity, to decrease the feature map resolution and computational complexity of the network, the feature map is again fed into the pooling

layer. By cascading convolution layer, incorporating nonlinearity, and pooling layers, the method of extracting useful features in the input image consists of several similar steps.

### 4.2.3 HPV Caused Cancer Detection CNN Model

```
#HPV Caused Cancer Diseases Detection Model
model = Sequential()
model.add(Conv2D(32, (5, 5),strides=2, input_shape=(127,127,3)))
model.add(Activation('relu'))
model.add(MaxPooling2D(pool_size=(3, 3), strides=1))
model.add(Conv2D(32, (3, 3),strides=1))
model.add(Activation('relu'))
model. add(MaxPooling2D(pool_size=(2, 2), strides=1))
model.add(Conv2D(64, (3, 3), strides=1))
model.add(Activation('relu'))
model.add(Conv2D(64, (5, 5),strides=2))
model.add(Activation('relu'))
model.add(Conv2D(64, (3, 3), strides=1))
model.add(Activation('relu'))
model.add(MaxPooling2D(pool_size=(2, 2),strides=2))
model.add(Flatten()) #A this converts our 3D feature maps to 10 feature vectors
model.add(Dense(64))
model.add(Activation('relu'))
model.add(Dropout(0.5))
model.add(Dense(64))
model.add(Activation('relu'))
model.add(Dropout(0.5))

model.add(Dense(1))
model.add(Activation('sigmoid'))
```

**Algorithm 4. 2:** *HPV Caused Cancer Diseases Detection Model*

We measure the spatial size of the output volume in each layer by using the input image size (H1), receptive filter size (F), stride (S) and number of zero paddings (P) [67]. The following equation gives the exact volume size of production of all layers in the model proposed.

$$Output\ size\ (H2) = \frac{(H1-F+2P)}{S} + 1 \qquad (2)$$

Where: H1 is the input volume size, F is the size of the filter, P is the number of zero padding's, and S is the stride volume.

The initial spatial size of the input volume (H1) is 127 x 127 x 3 of a 3D image, and after some convolution operations it is changed. The initial size of the filter F and Stride S is 5 x 5

x 3 and 2 respectively, and after some convolution and pooling operations these sizes are changed. And there is no zero padding P in our network and the value of P throughout the model is always zero. The following table describes all the parameters of each layer in accordance with the equation (4.2) given above.

There are reciprocal limits to the spatial parameters. For example, if the input volume is H1 = 10, no zero-padding is used, P = 0, and the filter size is F = 3, then it will be difficult to use stride S = 2, since Equation (4.2) gives 4.5, which is not an integer, implying that the neurons do not fit neatly and symmetrically across the input. During the process of resizing images found in the dataset, we have considered this setting of parameters correct. If this structure is not taken into account, an exception will be thrown by libraries used to enforce the CNN model or the rest of the region will be zero pads or the picture will be cropped to make it fit.

**Input layer:** the input layer of our CNN model accepts RGB images with dimensions of 127 × 127 x 3 and two unique classes (diseased and healthy). This layer only transfers the input to the first convolution layer without any calculation. As a result, there are no learnable features in this layer, and the number of parameters is 0.

**Convolutional layer:** There are five convolutional layers in the HPV-caused cancer detection model. The model's first convolutional layer filters the 127 x 127 x 3 input picture with 32 kernels of size 5 x 5 x 3 and stride 2 pixels. Because (127 x 5)/2 + 1 = 62 and this layer has a depth of K = 32, its output volume is 62 x 62 x 32. The volume output product is 123,008. The total number of neurons in the layer is given here (first layer of conv). This volume's 62 62 x 32 neurons are coupled to a 5 x 5 x 3 region in the input volume. The number of parameters in the convolution layers is tracked by parameter sharing**.** If we look at the input, the first layer of convolution has 62 x 62 x 32 = 123,008 neurons, each with 5 x 5 x 3 = 75 weights and 1 bias. In the first layer of the model, this adds up to 123,008 x 75 = 9,225,600 parameters. This amount in our machine is obviously quite large and tough to accomplish. The concept of parameter sharing, which is one of the advantages of a neural network over a normal one, is introduced here. If we employ parameter sharing, and one function, say (x, y), is used at one particular area, it should be possible to compute at another location (x2, y2). In other words, a single 2-dimensional depth slice with a volume of 62 x

62 x 32 has 32 depth slices, each 5 x 5, with the same weight and bias. The first convolution layer in the HPV-caused cancer detection model includes only 32 distinct weight sets (one for each depth slice), for a total of 32 x 5 x 3 = 2,400 unique weights or 24,432 parameters, by adding 32 biases and the identical parameters to all 62 x 62 neurons in each depth slice. The output volume and properties of each learnable layer in the HPV-caused cancer detection model are described in Table 4.1.

The first convolutional layer's data (pooled output) is transmitted to the second convolutional layer, which filters it with 32 kernels of size 3 x 3 x 32. The third, fourth, and fifth convolutional layers are linked without interfering with the pooling layer. The third pooled convolution layer filters the outputs of the second pooled convolution layer using 64 kernels of size 3 x 3 x 64 as input. The fourth convolutional layer has 64 kernels of size 5 x 5 x 64, and the fifth convolutional layer contains 64 kernels of size 3 x 3 x 64. ReLU nonlinearity is used as an activation function in all of the convolutional layers in the HPV-caused cancer detection model. ReLU was chosen since it is faster than other nonlinearities [66].

**Pooling layer:** after the HPV caused cancer detection model's first, second, and third convolutional layers, three max-pooling layers remain. The initial max-pooling layer reduces the performance of the first convolutional layer with a filter of size 3 x 3 and stride 1. The second max-pooling layer takes the output of the second convolutional layer as an input and pools it using 2 x 2 step 1 filters. On the third max-pooling layer, there are size 2 x 2 filters with stride 2. The number of parameters in these layers is 0 because this layer has no learnable properties and just conducts sampling operations along the input volume spatial axis.

**Fully Connected (FC) layer:** The HPV-caused cancer detection model consists of three fully connected layers, including the output layer. The first two completely connected layers each have 64 neurons, whereas the final layer, the model output layer, only contains one neuron. After converting the 3D volume of data into a vector value, the first FC layer accepts the output of the third conv layer (Flattening). This layer computes both the class score and the number of neurons in the layer defined during model construction. It's the

same as regular NN, and as the name says, each neuron in this layer is linked to every number in the previous layer.

**Output layer:** The output layer is the model's last layer (the third FC layer), and it has sigmoid activation of one neuron. The model is intended to classify two groups using a technique known as binary detection.

Table 4.1: *HPV Caused Cancer Detection Model Summary of Parameters for Detection of HPV caused Cancer*

| **Layer** | | **Filter** | **Depth** | **Stride** | **Number of Parameters** | **Output Size** |
|---|---|---|---|---|---|---|
| Input | Image | - | - | - | 0 | 127 x 127 x 3 |
| 1 | conv2D + ReLU | 5 x 5 | 32 | 2 | 2,432 | 62 x 62 x 32 |
| | maxPool2D | 3 x 3 | - | 1 | 0 | 60 x 60 x 32 |
| 2 | conv2D + ReLU | 3 x 3 | 32 | 1 | 9,284 | 58 x 58 x 32 |
| | maxPool2D | 2 x 2 | - | 1 | 0 | 57 x 57 x 32 |
| 3 | conv2D + ReLU | 3 x 3 | 64 | 1 | 18,496 | 55 x 55 x 64 |
| 4 | conv2D + ReLU | 5 x 5 | 64 | 2 | 102,464 | 26 x 26 x 64 |
| 5 | conv2D + ReLU | 3 x 3 | 64 | 1 | 36,928 | 24 x 24 x 64 |
| | maxPool2D | 2 x 2 | - | 2 | 0 | 12 x 12 x 64 |
| | flatten | - | - | - | 0 | 9216 |
| 6 | FC + ReLu | - | 64 | - | 589,888 | 64 |
| | Dropout | - | - | - | 0 | 64 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 7 | FC + ReLu | - | 64 | - | 4,160 | 64 |
| | Dropout | - | - | - | 0 | 64 |
| Output | FC + Sigmoid | - | 1 | - | 65 | 2 |
| **Number of Total Parameters** | | | | | **763,681** | |

As shown in Table 4.1 and the algorithm attached hereto as Annex A – Proposed CNN Model Code, when compared to previous deep learning models such as AlexNet, which has 60 million parameters, VGG, which has 138 million, and InceptionV3, which has 24 million, the suggested model has 763,681 extremely modest parameters. Deep learning models are recognized for having a large number of parameters, therefore training them from scratch requires a lot of computational power and a lot of data. However, the suggested model can be trained and performs well with a small quantity of data and resources.

### 4.2.4 Training Components of HPV caused Cancer Detection Model

CNN model selection is very difficult components since most model are implemented in large-scale applications such as ILSVRC, which has millions of parameters with thousands of classes and requires high computational power. The model is implemented in our thesis in restricted hardware resources and is built for only two classes. The goal of a CNN model is to find an efficient model that can work on a small number of images with limited processing resources (CPU and GPU). To provide nonlinearity during network training, the proposed approach incorporates 5 convolution layers, 3 fully connected layers, ReLU as an activation feature in the hidden layers, and dropout after the first two fully connected layers to avoid over-fitting problems. We reduced the number of neurons, parameters, and filters in pre-trained CNN models due to a decrease in trained classes, hardware resources, and a huge number of images.

## 4.3 Detection Using HPV Caused Cancer Detection Model

Detection into completely linked layers is done out in the suggested model. We have a total of three entirely connected levels, including the output layer, as we have seen in the HPV caused cancer detection model above (Figure 4.1). The major goal of these layers is to identify the input image using the information collected by the convolutional layers. The

first fully-connected layer accepts the convolution and pooling layers' performance. Before being sent into the fully linked layer, the outputs are merged and flattened to a single vector value. Each vector value reflects the probability that a given characteristic (in our dataset, color) belongs to a specific class. The flow of input values via the FC layer of the network is depicted in the figure below the algorithm Annex A – Proposed CNN Model Code.

If the image is infected, white hues indicate a high possibility of the diseased class. The data are multiplied by weights and sent via an activation function (ReLU) before being sent to the output layer, where each neuron expresses detection model (infected or healthy). To put it another way, the completely linked layer produces a single value by performing the input data dot product (the features produced from the convolution and pooling layers) and the weights.

## 4.4 Detection Using Pre-Trained Models

On a large-scale image detection assignment, CNN models are trained with a huge number of images, typically thousands of classes. Since millions of images and thousands of classes train the models, their ability to generalize an object is better. The properties gained by the models are then employed for many additional problems in the actual world, even though the difficulties are distinct from those of the original task. These models are known as pre-trained models, and the process of training and using them is referred to as transfer learning. Instead of building a huge CNN model from the ground up, anyone can use these pre-trained models to train their data. Our dataset is used to import and train pre-trained models of those models in our thesis. The convolution base, which includes convolution and pooling layers to extract meaningful features from the input picture, and the classifier, which classifies the input image based on the characteristics extracted from the convolution base, are the two primary pieces of most pre-trained CNN models. The knowledge base is the section of the convolution that contains the properties learned during training. Near the input layer, a convolution layer learns general features for all the images in the convolution base. Convolution layers near to the classifier, on the other hand, only have features that are unique to the given image. As a result, we may use convolution blocks to train other issues in the model's top (input) layer, and they generalize well.

There are two widely used methods for transfer learning on pre-trained models: one is to train the entire convolution base of the pre-trained model and only change the fully connected layer, and the other is to train a portion of the convolution base by freezing the pre-trained model's weights and modifying the fully connected layer. Fine-tuning some of the components of the convolution base is referred as training. We trained two pre-trained models, VGG16 and InceptionV3, on our dataset and compared the results to the suggested model.

## 4.5 Experimental Setup

In this thesis experiment, three options are investigated. The first two scenarios include the use of transfer learning to detect images, while the third involves the development of a new CNN-based model. During transfer learning, the well-known CNN models VGG16 and InceptionV3, which won the largest image detection competition, are used. Thousands of classes and millions of pictures are used to train the models. The HPV-caused cancer detection algorithm was created from the ground up, with millions of parameters reduced to just 763, 681 (Table 4.1). The total number of parameters is just the sum of all weights and biases.

## 4.6 Augmentation Parameters

The images are made by combining the augmentation settings in Table 4.2. Finally, the dataset is supplemented with various ways, yielding a sufficient number of images.

**Table 4.2:** *We Employed Augmentation Techniques*

| **Augmentation Parameter** | **Augmentation Factor** |
|---|---|
| Horizontal Flip | 1(True) |
| Shear Range | 0.3 |
| Range of Width Shift | 0.2 |
| Range of Height Shift | 0.2 |
| Range of Zoom | 0.2 |

## 4.7 Hyperparameter Settings

Hyperparameters are deep learning configurations that are defined before the training process begins and are external to the algorithm. There is no universal rule for selecting the optimum hyperparameters for a particular situation. To select the hyperparameters, numerous tests are carried out. The model's hyperparameters are listed in the Table:

- Optimization algorithms: the HPV caused cancer detection model is trained using the gradient descent optimization technique to reduce the error rate, and the weights are changed using the back propagation of the error algorithm. Gradient descent is by far the most popular and widely utilized optimization technique in deep learning studies. At the same time, each state-of-the-art deep learning library includes gradient descent optimization algorithms like Keras (used in our thesis). It adjusts the weight of the model and tweaks the parameters to minimize the loss function. The gradient descent is optimized using the Adaptive Moment Estimation (Adam) optimizer. Adam calculates the adaptive learning rate for each parameter and scales it using square gradients, as well as the moving gradient average.
- Learning rate: learning rate is used during weight update because back propagation was used to train the HPV caused cancer detection model. The amount of weight that needs to be altered during back propagation is set. The most difficult aspect of our experiment was determining the appropriate learning rate. In our experiment, we discovered that a learning rate with a low value takes longer to train than one with a higher one. However, a model with a lower learning rate is better than one with a higher learning rate. The experiment was carried out with learning rates of 0.001, 0.01, and 0.1. Then, even though it takes longer to train, learning rate 0.001 is the best for all tests.
- Loss function: the activation functions employed in the model's output layer (the last fully linked layer) and the sort of problem we're trying to answer both influence the loss function we utilize (whether regression or detection). In the suggested model, the sigmoid is used as an activation function in the last completely linked layer. A detection problem, specifically binary detection, is the type of task we're focusing on. As a loss function, we used binary cross-entropy loss in our model. Although other loss functions such as Categorical Cross-Entropy (CCE) and Mean Squared Error (MSE) exist, binary

cross-entropy is the suggested loss function for binary detection. It's ideal for models that calculate the difference between the actual and expected outputs, or output probabilities. Both BCE loss and CCE loss were used in the analysis.

- Activation function: experiments are carried out using two alternative activation functions: SoftMax and Sigmoid perform well in the suggested model, with Sigmoid outperforming SoftMax. Because it is the best option for a binary detection problem, the sigmoid activation function is utilized in the model's output layer.
- Number of epochs: the number of epochs is the number of pre-warned and backward iterations that the complete dataset goes through the model or network. In our experiment, we used several epochs ranging from 10 to 150 to train the model. When we utilize too small or too large epochs during training, we see a large gap between the training error and the validation error. After multiple experiments, the model becomes optimal at epoch thirty (30).
- Batch size: The batch size is the amount of data that we transport over the network all at once. We must divide the input into smaller batches since transmitting all of the data to the machine in a single epoch is too difficult. When preparing models, it is ideal to lower the machine's computing time. In our experiment, we used a batch size of 32 for model training.

# Chapter 5: Experiment and Evaluation

This chapter presents the development environments and the implementation of the HPV caused cancer detection process by using the proposed CNN algorithm specified in detail in the previous chapter. All the experimental data, such as the outcomes of and experiment and the discussion of these outcomes, are briefly provided in this chapter. Different graphs and tables display the outcomes of the experiments.

## 5.1 Development Environment and Tools

For the implementation of the HPV caused cancer detection model, many tools and techniques are used. All experiments have been carried out on a laptop with the following configuration for the implementation of our proposed system: Intel(R) Core i5 Processor 1.80GHz, 8GB RAM and Windows 10 operating system base. The software tools we used to implement the CNN algorithm are Python as a programming language with the anaconda environment libraries of TensorFlow and Keras. These tools meet all the conditions for consideration and are used in python, which is common to us.

**TensorFlow**

Tensorflow is a Google-created free and open source library that is today the most popular and fastest deep learning library [61]. TensorFlow's architecture allows for data preprocessing, model construction, model training, and model estimation. Tensors (n-dimensional arrays) represent all sorts of data in TensorFlow computations. During preparation, TensorFlow additionally employs a graphical framework to graphically describe the computational sequence.

**Anacoda**

Anacoda is used for model implementation and is a free and open source distribution of Python for applications related to data science and deep learning, aimed at simplifying the management and deployment of packages. It includes various IDEs such as Jupyter Notebook and Spyder that are used to compose the coding part. To implement the coding aspect, we used the Jupyter notebook. It is straightforward and runs on a web browser.

**Keras**

Keras is a python-written high-level neural network API operating on top of TensorFlow, Theano5, or Microsoft Cognitive Toolkit (CNTK). User-friendly, easily extendable with python, it is very easy to build a model, and most notably it includes pre-trained CNN models such as VGG16 and InceptionV3 that we use throughout the experiment. This makes simple, quick and supports both CNN and RNN or the mix of the two [61].

**MS-Visio**

MS-Visio is used for constructing the model proposed. This tool has been used to quickly build, interact and exchange data-linked diagrams with ready-made drawings and to simplify comprehensive data.

## 5.2 Model Evaluation

We need to know how the model generalizes after training our model for knowledge never seen before. This allows us to claim that the model is well classified with new data, or the model is only doing well for trained data (memorizing previously fed data) but not in new data (data not seen before). Model evaluation is the method of estimating the generalization accuracy of the model proposed with unseen data (in our case test data). Detection accuracy metrics are used in our thesis, which are recommended for detection problems and when all dataset groups have the same number of specimen samples. The accuracy is calculated by the division of number of correct predictions and the total number of predictions. In this process, the dataset is split into datasets for preparation, validation, and testing. We are able to feed the validation split to the model during the training to get performance metrics. The model returns training data accuracy and loss, and validation data accuracy and loss, which are training accuracy, validation accuracy, training loss, and validation loss. So, using these metrics, we can plot loss and accuracy graphs in relation to epochs. Finally, the test data (images not used in either the training or validation sets) is provided to the trained model to test the model's performance, then the model returns accuracy and loss of test data that is never seen during the training.

The color function of the image, as defined in detail in Section 4.4, was used to identify the input image. The action is intended for using the color function to categorize the image is

because we can tell whether it is healthy or unhealthy just by looking at it. Three different detection scenarios are used in the experiment to evaluate the detection results.

The first two scenarios are based on CNN models that have already been trained, whereas the third scenario is based on the HPV caused cancer detection model. Our tests have two primary steps, similar to most deep learning detection systems. The first is the training stage, and the second is the testing stage. During the training phase, data is continually presented to the classifier while weights are modified to get the desired result. To assess the detection algorithm's performance, the trained algorithm is applied to data that the classifier has never seen before (test data). The experimental effects are detailed below.

## 5.3 Pre-trained CNN Model

VGG and InceptionV3 are two pre-trained CNN models that have been fine-tuned and are often utilized in ImageNet pre-trained models. The VGG model is employed because of its simplicity, whereas the InceptionV3 model is chosen because of its rich characteristics. As a result, experiments are carried out in both a reasonably simple model and a complex model in order to obtain the detection accuracy of these models in our dataset. All of the experiments are run on the same dataset and in the same hyperparameter setting.

### 5.2.1 Detection of Cancers Caused by HPV by using VGG16 Pre-trained Model

The VGG model is distinguished by its simplicity, which is achieved by employing only 3 x 3 convolution layers that are pegged on one other as layer depth increases. The VGG model is offered in two variations: the VGG16 and the VGG19. Within the network, the VGG16 has 16 levels of weight while the VGG19 has 19 layers of weight. The model takes 224 x 224 RGB images as input and provides 1000 ImageNet dataset classes (14 million images divided into 1000 categories) [77, 78]. In a 3 x 3 receptive field, the input is conveyed via stacked convolution layers of the model with nonlinearity, which is ReLU. For all 3 to 3 convolutions, the model employs a step of 1 and spatial padding of 1. To limit the spatial size of the convolution layers' output, there is a max-pooling of window size 2 x 2 with step 1 for every three successive convolution layers. The VGG19 model contains 16 convolution layers, while the VGG16 model has 13 convolution layers with 5 max-pooling layers apiece [79, 80]. Finally, a stack of conv layers is followed by three entirely linked detection layers. The first two layers use a SoftMax activation function to achieve 4096 channel depth, and

the last layer has 1000 channel depth, which is equal to the number of classes in the ImageNet dataset.

In our experiment, we feed a down sampled RGB image with a size of 127 by 127 into the model, which is fine-tuned to provide two output groups in our dataset. The original VGG16 model includes 138 million parameters in total, which is enormous. Because the image of our model has a smaller spatial dimension and we only trained a subset of the model. We fine-tuned the VGG16 model, as previously stated, by simply using the network's conv foundation. We did a slew of experiments to discover the best pre-trained model by training the model's conv blocks. The model is trained by using the entire conv base of the network and only altering the totally linked layers, and the results demonstrate a considerable degree of overfitting. Overfitting occurs when the model weights are built with millions of images that are not from our dataset and thousands of classes, despite the fact that we sought to train the model with only 11,056 original images. As a result, we'll need to adjust some of the network weights and increase the amount of images using data augmentation approaches. We've decided to freeze portions of the model's layers (conv blocks) as a result of the improved data and perform a different experiment. We discovered that freezing the first three conv blocks was the best option, compared to freezing the first two and four conv blocks, using 66,336 images generated by augmentation procedures. To train the network, the hyperparameters specified in Table 4.3 above are employed. The experiment's performance has a mean accuracy of 99.4 percent for training and 99.4 percent for testing.

### 5.2.2 Result Analysis of VGG16

The two charts below show the detection accuracy and loss of the pre-trained VGG16 model with respect to epochs that we performed experiments by making some changes to the original pre-trained model in order to classify the model well in our dataset using detection accuracy metrics such as training and validation accuracy, training loss, and validation loss. The training precision is approximately 98 percent in the first epoch, progressively increasing to 99 percent in the fourth epoch. The model's training accuracy is higher between epochs 4 and 30, with 99 percent accuracy. The dataset improves in accuracy over the first few epochs. As seen in the figures below, the validation accuracy line is virtually in step with the training accuracy line, and the validation loss line is also nearly in step with the

training loss. Despite the fact that the validation accuracy and loss of validation lines are not linear, the model overfits. To put it another way, validation loss is falling rather than increasing, while validation accuracy is stable.

In the table below, the results of the pre-trained VGG16 model experiment are shown separately, with detection precision metrics reported as a percentage for train, validation, and test data.

**Table 5.1:** *Pre-Trained Model VGG16's Mean Accuracy and Loss*

| Metrics | Mean Accuracy | | | Mean Loss | | |
|---|---|---|---|---|---|---|
| | Training | Validation | Testing | Training | Validation | Testing |
| Value | 99.4% | 99.68% | 99.44% | 28.90% | 27.18% | 9.27% |

### 5.2.3 Detection of Cancers Caused by HPV by using InceptionV3

Inception is a Google GoogLeNet-developed deep CNN computer vision model that takes its name from the popular internet meme "We Need to Go Deeper". This approach provides a network that is deeper and uses fewer computational resources (a large number of layers and a large number of neurons in each layer). When we say deeper network, we mean that as the number of layers increases, the network is more likely to overfit. One thing to keep in mind is that as the number of neurons in each layer grows, so does the computational resource requirement. The original model addresses this issue by substituting a sparsely linked network (filters of various sizes on the same layer) for the FC layer, particularly in the convolution layer, and this technique allows us to keep processing costs low while increasing network depth.

This model is trained on the ImageNet dataset and delivers a final output of 1000 groups by accepting a size of 299 × 299 x 3 images as input. It contains a total of 42 layers, and it is computationally faster than the VGG model, which has only 16 and 19 layers. In our experiment, we used our size 127 x 127 RGB picture dataset to train the pre-trained inception model, and we only adjusted the performance to two classes without using any

fine-tuning techniques in the conv base. The network was given a total of 66, 336 images, which resulted in a positive performance with no overfitting issues.

### 5.2.4 Result Analysis of InceptionV3

The result analysis of Inception V3 pre-trained model training accuracy in the first epoch is approximately 94 percent, and the validation accuracy is around 99 percent. The training's accuracy improves almost immediately, as evidenced by the increasing value at epoch 2. The accuracy of validation improves linearly and does not drop at the same time that the loss of validation falls linearly without growing, and the difference between the accuracy and loss of training and validation is not significant. As a result, there is an overfitting issue in the model when we exercise with our dataset.

The findings obtained from the pre-trained InceptionV3 model experiment are described separately in the following table using the detection accuracy metrics as a percentage for the training data, validation data, and test data.

**Table 5.2:** *Pre-Trained Model Mean Accuracy and Loss Of Inceptionv3*

| Metrics | Mean Accuracy | | | Mean Loss | | |
|---|---|---|---|---|---|---|
| | Training | Validation | Test | Training | Validation | Test |
| Value | 98.93% | 99.68% | 99.44% | 28.93% | 27.49% | 10.12% |

## 5.4 HPV Caused Cancer Detection CNN Model

In this thesis, a CNN model that can run on a modest amount of hardware and generate promising results was built and enhanced. As stated previously in section 4.8. The model contains eight layers in total, five convolutions and three dense layers. Like the other pre-trained models we've used in this thesis, it receives 127 by 127 color photographs and outputs two classes of output. Following the implementation of data augmentation in the dataset, the HPV caused cancer detection model is trained utilizing a total of 66,336 images. The HPV caused cancer detection model runs several experiments by varying the ratio of training and testing datasets, using different learning speeds, comparing the dataset before and after augmentation, and finally applying various activation functions.

### 5.4.1 Scenario 1: Modifying the Training and Testing Dataset Ratio

The results of the HPV caused cancer detection model's experiment with varied training and test split ratios are carried out, with detection accuracy metrics expressed in percentages for train, validation, and test data, respectively.

We carried out with four training and testing ratio of 6:4, 7:3, 8:2 and 9:1. The HPV caused cancer diseases detection model performs well with varying training and testing data set ratios. Among the other three experiments, using 80 percent for training and 20 percent for testing is better or ideal and its result is 99.7%. The 8:2 ratios suggest that 80 percent of the total dataset is used for training and 20 percent is used for research. Furthermore, the training data is derived from the validation data. The validation data is 40 percent of the training data in 6:4 ratios (not the entire dataset), and the validation data is 30 percent of the training data in 7:3 ratios, and so on.

### 5.4.2 Scenario 2: Learning Rate Changing

The HPV-caused cancer disease detection model experiment with different learning rates, using detection accuracy metrics as a percent for the train data, validation data, and test data individually. As evidenced by the following findings, offering higher learning rates is less accurate than offering lower learning rates. As a result, the HPV-caused cancer detection model considers a learning rate of 0.001 to be ideal and the result is 99.4%.

### 5.4.3 Scenario 3: Using Different Activation Function

The findings of the HPV-caused cancer diseases detection model experiment with independent detection accuracy metrics for training data, validation data, and test data in the form of percentages. In the HPV-caused cancer diseases detection model, the sigmoid activation function is regarded as the best and is suggested for binary groups.

### 5.4.4 Scenario 4: With and Without Dataset Augmentation

The findings of the HPV-caused cancer diseases detection model experiment with independent detection accuracy metrics for training data, validation data, and test data in the form of percentages. In the HPV-caused cancer diseases detection model, the dataset with augmentation technic is better rather than of without dataset augmentation.

### 5.4.5 Result Analysis for the HPV Caused Cancer Detection Model

As we can see in the following graph (Figure 5.6), the value of the training accuracy line gets 96 percent at the beginning of the training and the value of the validation accuracy line is about 99 percent, so the values of training accuracy get higher up to epoch 6. Training precision lines pass 99 percent after epoch 6 and increase very slowly and the validation precision lines stay the same as the first epoch, and it does not increase because the optimizer has found a local minimum for the loss. When we see the training loss curves in Figure 5.7, the plot from the first epoch to epoch 30 decreases linearly from 22 to 4. The training loss lines cross 4 after epoch 10, which decreases considerably from the beginning of the training and does not pass 4, which is the smallest value in the training. The plot is not decreasing linearly when the validation loss curves remains constant in Figure 5.7.

Finally, we can see that the accuracy of validation is in line with the accuracy of training and the loss of validation is in sync with the loss of training. The curves show that there is no over fitting in the HPV caused cancer detection model, because the accuracy of validation is increasing very slowly not decreasing and the loss of validation is not increasing, and most significantly, there is not much training difference and accuracy of validation and there is also not much gap between training and loss of validation. Therefore, we can assume that the generalization potential of our model has been much stronger because the loss of the training set was just marginally greater relative to the validation loss the algorithm attached hereto as Annex A – Proposed CNN Model Code.

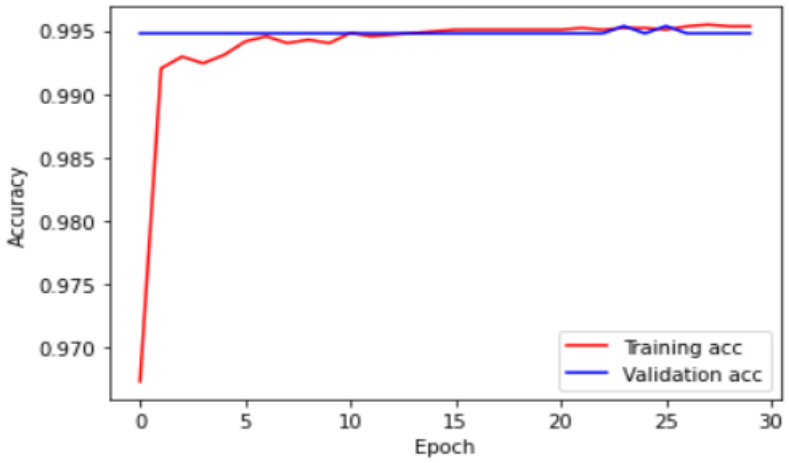


**Figure 5.1:** *HPV Caused Cancer Detection Model Training and Validation Accuracy*

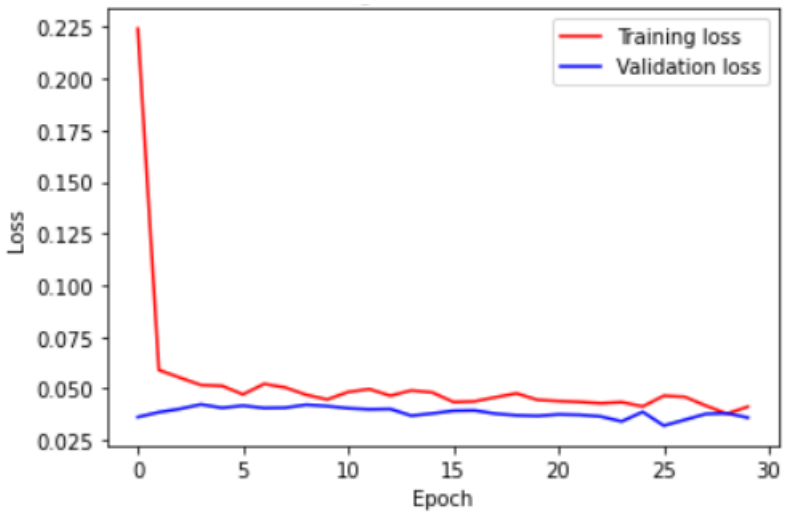


**Figure 5.2:** *HPV Caused Cancer Detection Model Training and Validation Los*

The findings of the HPV-caused cancer detection model experiment are reported in the table below, with the method attached as Annex A – Proposed CNN Model Code, using the detection accuracy metrics individually in the form of percentage for training data, validation data, and test data.

**Table 5.3:** *The Accuracy and Loss of the HPV-Caused Cancer Detection Model*

| **Metrics** | **Accuracy** | | | **Loss** | | |
|---|---|---|---|---|---|---|
| | Training | Validation | Testing | Training | Validation | Testing |
| Value | 99.3% | 99.4% | 99.4% | 5.2% | 3.8% | 3.8% |

## 5.5 Discussion

Our HPV caused cancer detection model has been proven by medical experts Betezatha hospital for HPV-related cervical cancer and ENT specialist from St. Paul's hospital for HPV-related oral cancer. The experiments are carried out using three distinct CNN models, which are two pre-trained models, VGG and InceptionV3, and the suggested model, as given in the preceding sections. All of the experiments are carried out using the same hardware and system specification. The amount of images in the dataset used to train models does not vary depending on the depth of the models or the number of parameters. All models were trained using the same set of 66, 336 pictures. All of the models were tested on an

unknown dataset and yielded positive results during model training. When the performance of the HPV caused cancer detection models is compared to the two pre-trained models and models discussed in our related works section, we discover that the HPV induced cancer detection model surpasses the others.

The mean percentage of training accuracy for VGG16, InceptionV3, and the suggested CNN model is 99.4, 98.9, and 99.3, respectively, as shown in the graph below. These findings show that the models outperform on the training dataset. The mean validation accuracy for the VGG16, InceptionV3, and HPV-caused cancer detection models is 99.4, 99.6, 99.6, and 99.4, respectively. When we compute the difference between mean training accuracy and mean validation accuracy for each of the three experiments, we find that the difference is very small, and nearly all mean training accuracy and mean validation accuracy are the same in the HPV caused cancer detection model. This means that there is no overfitting in the models, and we can conclude that the generalization potential of the suggested model is high. We achieved 28.9, 28.9, and 5.2 for the mean training loss that is utilized for the three experiments to assess the discrepancy between the expected value and the actual value: VGG16, InceptionV3, and the HPV caused cancer detection model, respectively. When we quantify the difference between mean losses of training and mean validation, the mean loss of validation is 27.1, 27.4, and 3.8, which is about the same as the mean loss of training.

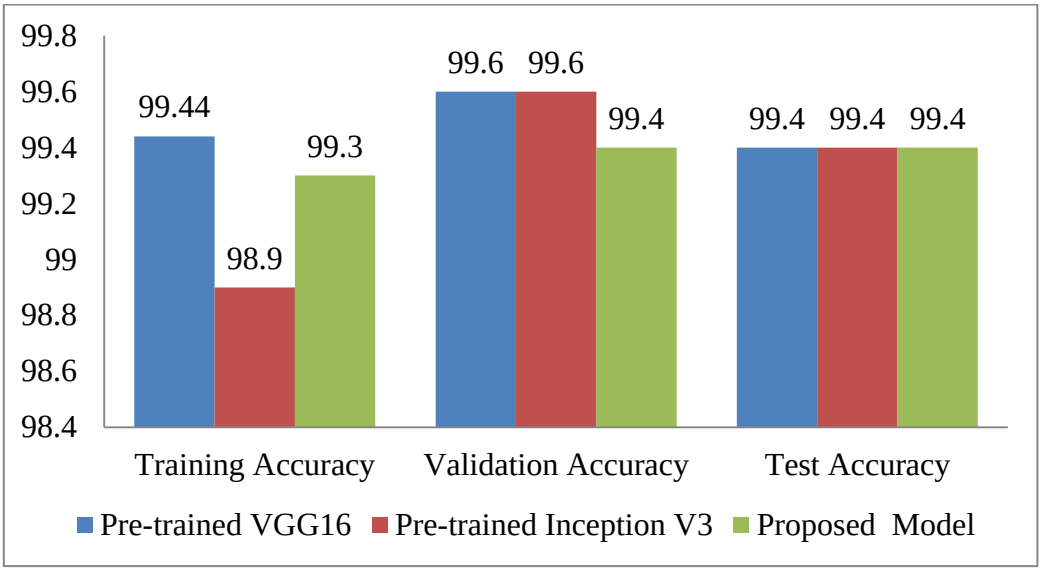


**Figure 5.3:** *The Three Experiments Mean Accuracy*

We have obtained a promising result by testing the models with unseen data. The test precision of VGG16 is 99.44 percent, InceptionV3 is 99.44 percent, and the HPV caused cancer detection model's test accuracy is 99.44 percent. The test results show that the HPV caused cancer detection model can effectively identify the given image as infected or safe.

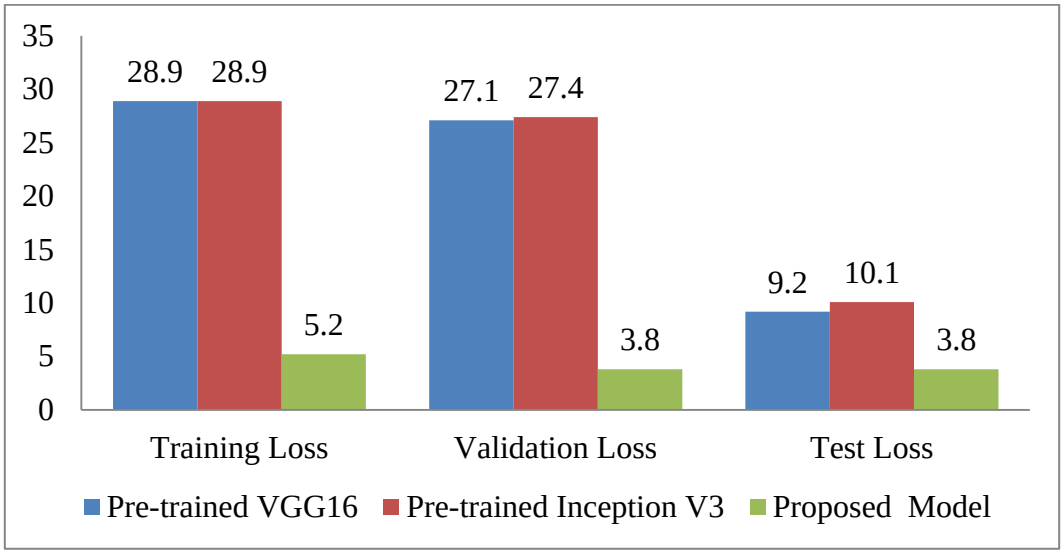


**Figure 5.4:** *The Three Experiments Mean Loss*

Figure 5.9 shows that all of the tests failed, and the HPV caused cancer detection model's value is lower than the two pre-trained models, VGG and InceptionV3. The VGG16 test loss is 9.2, the InceptionV3 test loss is 10.1, and the suggested model test loss is 3.8. All of which are acceptable for the HPV caused cancer detection model. As a result, the HPV caused cancer detection model performs well in both the training and testing datasets. The main reason the HPV caused cancer detection model performs better is due to the dataset we used to train the model, and the second main reason is that our HPV caused cancer detection model uses smaller filters in the network's convolution layer. Using a smaller convolution aids in the description of extremely small features utilized to distinguish between the input image and the possibility of missing a crucial feature. Most deep learning algorithms are trained on high-performance computing machines with a faster GPU, a massive number of images (in millions), and tens of millions of parameters, particularly in computer vision for image detection problems. However, we can train and achieve better outcomes by using compact networks with fewer parameters, less hardware use, and less data. More reliable findings can be produced if the data set photographs are captured in a stable environmental

setting, which is a stable distance from an object to the camera, correct illumination, and proper focus. The model's accuracy would also be improved by preprocessing the images by removing noise and undesirable characteristics.

# Chapter 6: Conclusion and Future Work

## 6.1 Conclusion

Cancer diseases caused by HPV are the most common killer infectious diseases in the world. HPV is a significant cause of the highest risk factor for cervical cancer, oral cancer and other associated cancers. As one of the main killer diseases in the world, there is an immediate need to diagnose and detect the disease for treatment at an early stage. To address this need, we proposed HPV caused cancer detection model and implemented a deep learning approach using the CNN algorithm in order to detect the diseases early on. To detect HPV caused cancers, we have presented a CNN model. The HPV caused cancer detection CNN model can then be used to detect HPV caused cancers of cervical and oral.

During the experiment, with the aid of medical exercises, we used digital images directly obtained from hospitals and the internet from Kaggle. The two pre-trained models, namely the VGG16 and InceptionV3, and the HPV caused cancer detection model, were trained. All of the models were able to find a successful detection result after many experiments. The VGG16 model has 99.4 percent training or detection accuracy and 99.4 percent testing accuracy, the InceptionV3 pre-trained model has 98.9 percent training or detection accuracy and 99.4 percent testing accuracy, and the HPV caused cancer detection model has 99.3 percent training or detection accuracy and 99.4 percent testing accuracy. The VGG16 model has a training loss of 28.9% and a testing loss of 9.2%, the InceptionV3 pre-trained model has a training or detection loss of 28.9% and a testing loss of 10.1%, and our HPV caused cancer detection model has a training loss of 5.2 percent and a testing loss of 3.8 percent.

Our findings show that the HPV caused cancer detection CNN model can significantly support accurate detection of HPV-caused cancers with low computational power and small images, which is far less than sufficient for deep learning algorithms because most deep learning algorithms are trained on millions of high computational resource images. To that end, the results of the experiment allow us to work with our model and test more HPV-related cancers.

## 6.2 Contribution

The contribution of this work to the scientific community and the general population is the design and development of a new CNN model that uses deep learning to better detect cervical and oral HPV-caused malignancies. We conducted numerous trials with pre-trained models and the suggested model to achieve these goals, and we were successful.

The contributions of this thesis work are that:

- We developed a new deep learning CNN model that can detect HPV-related diseases with a minimum hardware and software requirements.
- We tested our new model on HPV-related diseases detection and have achieved a better result.
- We have shown that the choice of CNN algorithms can directly affect a HPV caused cancer diseases detection model's accuracy and loss. We have compared three CNN models namely VGG16, Inception V3 and HPV caused cancer detection models and found that our CNN model performs better on those models for HPV caused cancer diseases detection.

## 6.3 Future Work

As a future work, various issues can be dealt with in order to enhance the model:

- Include all HPV type related cancers and expand the dataset.
- Detect and distinguish all types of high-risk and low-risk HPV caused cancer diseases.

## References


[1] Angela A. Cleveland, Julia W. Gargano, Ina U. Park, Marie R. Griffin, Linda M. Niccolai, Melissa Powell, Nancy M. Bennett, Kayla Saadeh, Manideepthi Pemmaraju, Kyle Higgins, Sara Ehlers, Mary Scahill, Michelle L. Johnson Jones, Troy Querec, Lauri E. Markowitz and Elizabeth R. Unger, "Cervical adenocarcinoma in situ: Human papillomavirus types and incidence trends in five states," *International Journal of Cancer*, April 2019.

[2] Maria Demarco, Olivia Carter-Pokras, Noorie Hyun, Philip E. Castle, Xin He, Cher Dallal, Jie Chen, Julia C. Gage, Brian Befano, Barbara Fetterman, Thomas Lorey, Nancy Poitras, Tina R. Raine-Bennett, Nicolas Wentzensen and Mark Schiffman, "Validation of a Human Papillomavirus (HPV) DNA Cervical Screening Test That Provides Expanded HPV Typing", *Journal of Clinical Microbiology*, Vol. 56, pp. 10-17, February 2018.

[3] S. Alizon, C.L. Murall and I.G. Bravo, "Why human papillomavirus acute infections matter," Viruses, October 2017.

[4] HyungJae Lee, Mihye Choi, Minkyung Jo, Eun Young Park, Sang-Hyun Hwang and Youngnam Cho, "Assessment of clinical performance of an ultrasensitive nanowire assay for detecting human papillomavirus DNA in urine," *Gynecologic Oncology*, November 2019.

[5] WHO, Human papillomavirus (HPV) and cervical cancer, Available on: https://www.who.int/news-room/fact-sheets/detail/human-papillomavirus-(hpv)-and-cervical-cancer, [Accessed January 2020], [Updated January 2019]

[6] Yousif Mohamed, Y. Abdallah and Tariq Alqahtani, "Research in Medical Imaging Using Image Processing Techniques," *Medical Imaging - Principles and Applications*, June 2019.

[7] D. Saranyaraj and M. Manikandan, "Medical Image Processing to Detect Breast Cancer - A Cognitive-Based Investigation," *2017 4th International Conference on Signal Processing, Communications and Networking*, March 17.

[8] Ferlay J, Ervik M, Lam F, Colombet M, Mery L, Piñeros M, Znaor A, Soerjomataram I and Bray F. “Global Cancer Observatory: Cancer Today,” *International Agency for Research on Cancer*, 2018.

[9] Darrell M. West and John R. Allen, How artificial intelligence is transforming the world, April, 2018.

[10] Ethiopia HPV Information Center, Human Papillomavirus and Related Cancers, Fact Sheet 2018, Available on: https://hpvcentre.net/statistics/reports/ETH_FS.pdf, [Accessed February 2020], [Updated June 2019].

[11] Chee Kai Chan, Gulzhanat Aimagambetova , Talshyn Ukybassova, Kuralay Kongrtay, and Azliyati Azizan, “Human Papillomavirus Infection and Cervical Cancer: Epidemiology, Screening, and Vaccination Review of Current Perspectives,” *Journal of Oncology*, October 2019.

[12] Silvia de Sanjo, Beatriz Serrano, Sara Tous, Maria Alejo, Belen Lloveras,Beatriz Quiros, Omar Clavero, August Vidal, Carla Ferrandiz-Pulido, MiquelAngel Pavon, Dana Holzinger, Gordana Halec, Massimo Tommasino, WimQuint, Michael Pawlita, Nubia Munoz and Francesc Xavier Bosch, “Burden of Human Papillomavirus (HPV)-RelatedCancers Attributable to HPVs 6/11/16/18/31/33/45/52and 58,” *JNCI Cancer Spectrum*, 2019.

[13] Rohan Gorantla, Rajeev Kumar Singh, Rohan Pandey and Mayank Jain, “Cervical Cancer Diagnosis using CervixNet A Deep Learning Approach,” 2019 *IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE)*, 2019.

[14] Matthew Wilhelm, Brian Nutter, Rodney Long and Sameer Antani, “Automated Detection of Human Papillomavirus: Via Analysis of Linear Array Images,” *2010 IEEE Southwest Symposium on Image Analysis & Interpretation*, Vol. 10, pp. 205- 208, 2010.

[15] Zilong Hu , Jinshan Tang , Ziming Wang , Kai Zhang , Lin Zhang and Qingling Sun, “Deep Learning for Image-based Cancer Detection and Diagnosis”, May 2018.

[16] R.Chtihrakkannan, P.Kavitha, T.Mangayarkarasi and R.Karthikeyan, “Breast Cancer Detection using Machine Learning,” *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, Vol. 8, pp. 3123 – 3126, September 2019.

[17] Siegel, R. L., Miller, K. D., & Jemal, A., "Cancer statistics," CA: A Cancer Journal for Clinicians, 2020.

[18] Tohid Mahmoudi , Miguel de la Guardia and Behzad Baradaran, "Lateral flow assays towards point-of-care cancer detection: A review of current progress and future trends," *TrAC Trends in Analytical Chemistry*, Vol. 125, February 2020.

[19] Yao Lu, Jia-Yu Li and Yu-Ting Su, "A Review of Breast Cancer Detection in Medical Images," School of Electrical and Information Engineering, Tianjin University, Tianjin, China, 2018.

[20] Prenitha Lobo and Sunitha Guruprasad, "Detection and Segmentation Techniques for Detection of Lung Cancer from CT Images," *Proceedings of the International Conference on Inventive Research in Computing Applications*, 2018.

[21] Patrice Monkam, Shouliang Qi, He Ma, Weiming Gao, Yudong Yao and Wei Qian, "Detection and Detection of Pulmonary Nodules Using Convolutional Neural Networks: A Survey", IEEE Access, Vol. 7, pp. 78075 – 78091, June 2019.

[22] SanaUllah Khan , Naveed Islam , Zahoor Jan , Ikram Ud Din , Joel J. P. and C Rodrigues, "A novel deep learning based model for the detection and detection of breast cancer using transfer learning," Pattern Recognition Letters, March 2019.

[23] Chiao, J.-Y., Chen, K.-Y., Liao, K. Y.-K., Hsieh, P.-H., Zhang, G., & Huang, T.-C., "Detection and detection the breast tumors using mask R-CNN on sonograms," Medicine, March 2019.

[24] Tanzila Saba, Sana Ullah Khan, Naveed Islam, Naveed Abbas, Amjad Rehman, Nadeem Javaid and Adeel Anjum, "Cloud-based decision support system for the detection and detection of malignant cells in breast cancer using breast cytology images," *Microscopy Research and Technique*, August 2018.

[25] S. Jane Henley, Cheryll C. Thomas, Jacqueline M, Lauri E. Markowitz, Mona Saraiya, "Human Papillomavirus–Attributable Cancers United States, 2012–2016, Morbidity and Mortality Weekly Report, Vol. 68, pp. 724 – 728, August 2019.

[26] Kalyani Sonawane, Ryan Suk, Elizabeth Y. Chiao, Jagpreet Chhatwal, Peihua Qiu, Timothy Wilkin, Alan G. Nyitray, Andrew G. Sikora, and Ashish A. Deshmukh, "Oral Human

Papillomavirus Infection: Differences in Prevalence Between Sexes and Concordance With Genital Human Papillomavirus Infection," NHANES 2011 to 2014, Annals of Internal Medicine, October 2017.

[27] Chee Kai Chan, Gulzhanat Aimagambetova, Talshyn Ukybassova, Kuralay Kongrtay and Azliyati Azizan, "Human Papillomavirus Infection and Cervical Cancer: Epidemiology, Screening, and Vaccination: Review of Current Perspectives," *Journal of Oncology*, Vol. 2019, pp. 1 – 11, October 2019

[28] Kazuhiro Kobayashi, Kenji Hisamatsu, Natsuko Suzui, Akira Hara, Hiroyuki Tomita and Tatsuhiko Miyazaki, "A Review of HPV-Related Head and Neck Cancer," *Journal of Clinical Medicine*, August 2018.

[29] N.A. Parmin, Uda Hashim, Subash C.B. Gopinath, S. Nadzirah , Zulida Rejali , Amilia Afzan and M.N.A. Uda, "Human Papillomavirus E6 biosensing: Current progression on early detection strategies for cervical Cancer," *International Journal of Biological Macromolecules*, Vol. 126, pp. 877 – 890, 2019.

[30] F. D'andrea, G. F. Pellicanò, E. Venanzi Rullo, F. D'aleo, A. Facciolà, C. Micali1, M. Coco, G. Visalli, I. Picerno, F. Condorelli, M. R. Pinzone, B. Cacopardo, G. Nunnari and M. Ceccarelli, "Cervical Cancer In Women Living With HIV," *World Cancer Research Journal*, 2019.

[31] Mahmood Rasool, Sara Zahid, Arif Malik, Irshad Begum, Hani Choudhry, Shakee Ahmed Ansari , Siew Hua Gan , Mohammad Amjad Kamal, Muhammad Asif, Fawzi Faisal Bokhari, Nawal Helmi, Mustafa Zeyadi, Mohammed Hussein Al-Qahtani and Mohammad Sarwar Jamal, "The human papillomavirus, cervical cancer and screening strategies: an update," *Biomedical Research*, Vol. 30, pp. 16 – 22, August, 2018.

[32] Megan J. Huchko, Easter Olwanda, Yujung Choi and James G. Kahn, "HPV-based cervical cancer screening in low-resource settings: Maximizing the efficiency of community-based strategies in rural Kenya," *Clinical Article Gynecology*, Vol. 148, pp. 386 – 391, January 2020.

[33] Abdulaziz Hakeem and Frank Alfred Catalanotto, “The role of dental professionals in managing HPV infection and oral cancer,” *Journal of Cancer Prevention & Current Research*, Vol. 10, pp. 82 – 88, August 2019.

[34] Alejandro Ismael Lorenzo-Pouso, Pilar Gándara-Vila, Cristina Bangal, Mercedes Gallas, Mario Pérez-Sayáns, Abel García, Ellen M. Daley & Iria Gasamáns, “Human Papillomavirus-Related Oral Cancer: Knowledge and Awareness Among Spanish Dental Students,” *Journal of Cancer Education*, May 2018.

[35] Malik Sallaml, Esraa Al-Fraihat, Deema Dababseh, Alaa’ Yaseen1, Duaa Taim, Seraj Zabadi, Ahmad A. Hamdan, Yazan Hassona, Azmi Mahafzah and Gülşen Özkaya Şahin, “Dental students’ awareness and attitudes toward HPV-related oral cancer: a cross sectional study at the University of Jordan,” BMC Oral Health, Vol. 19, pp. 1 – 11, 2019.

[36] A. Thirumal Raj, Shankargouda Patil, Archana A. Gupta , Chandini Rajkumar and Kamran H. Awan, “Reviewing the role of human papillomavirus in oral cancer using the Bradford Hill criteria of causation,” Disease-a-Month, 2018.

[37] Faber, M. T., Frederiksen, K., Palefsky, J., & Kjaer, S. K., “Risk of anal cancer following benign anal disease and anal cancer precursor lesions: A Danish nationwide cohort study,” Cancer Epidemiology Biomarkers & Prevention, 2019.

[38] Joanna Krzowska-Firych, Georgia Lucas, Christiana Lucas, Nicholas Lucas and Łukasz Pietrzyk, “An overview of Human Papillomavirus (HPV) as an etiological factor of the anal cancer,” *Journal of Infection and Public Health*, 2018.

[39] Chia-Ching J. Wang and Joel M. Palefsky, “HPV-Associated Anal Cancer in the HIV/AIDS Patient,” Cancer Treatment and Research, 2019.

[40] Jessica S. Wells, Lisa Flowers, Sudeshna Paul, Minh Ly Nguyen, Anjali Sharma & Marcia Holstad, “Knowledge of Anal Cancer, Anal Cancer Screening, and HPV in HIV-Positive and High-Risk HIV-Negative Women,” *Journal of Cancer Education*, March 2019.

[41] Mette T. Faber, Kirsten Frederiksen, Joel M. Palefsky and Susanne K. Kjaer, “Risk of Anal Cancer Following Benign Anal Disease and Anal Cancer Precursor Lesions: A Danish Nationwide Cohort Study,” *Cancer Epidemiology, Biomarkers & Prevention*, pp. 185 – 192, October 2019

[42] Lysandra Voltaggio, W. Glenn McCluggage, Jeffrey S. Iding, Brock Martin, Teri A. Longacre and Brigitte M. Ronnett, “A novel group of HPV-related adenocarcinomas of the lower anogenital tract (vagina, vulva, and anorectum) in women and men resembling HPV-related endocervical adenocarcinomas,” Modern Pathology, October 2019.

[43] Jiafeng Pan, Kimberley Kavanagh, Kate Cuschieri, Kevin G. Pollock, Duncan C. Gilbert, David Millan, Sarah Bell, Sheila V. Graham, Alistair R.W. Williams, Margaret E. Cruickshank, Tim Palmer and Katie Wakeham, “Increased risk of HPV-associated genital cancers in men and women as a consequence of pre-invasive disease,*” International Journal of Cancer*, Vol. 145, pp. 427 – 434, 2019.

[44] Mario Preti, John Charles Rotondo, Dana Holzinger, Leonardo Micheletti, Niccolò Gallio, Sandrine McKay-Chopin, Christine Carreira, Sebastiana Silvana Privitera, Reiko Watanabe, Ruediger Ridder, Michael Pawlita, Chiara Benedetto , Massimo Tommasino and Tarik Gheit, “Role of human papillomavirus infection in the etiology of vulvar cancer in Italian women,” Infectious Agents and Cancer, 2020.

[45] Caroline Measso do Bonfim, Letícia Figueiredo Monteleoni, Marília de Freitas Calmon, Natália Maria Cândido, Paola Jocelan Scarin Provazzi, Vanesca de Souza Lino, Tatiana Rabachini, Laura Sichero, Luisa Lina Villa, Silvana Maria Quintana, Patrícia Pereira dos Santos Melli, Fernando Lucas Primo, Camila Fernanda Amantino, Antonio Claudio Tedesco, Enrique Boccardo & Paula Rahal, “Antiviral activity of curcumin-nanoemulsion associated with photodynamic therapy in vulvar cell lines transducing different variants of HPV-16, Artificial Cells, Nano medicine, and Biotechnology,” *An International Journal*, Vol. 48, pp. 515 – 524, February 2020.

[46] Freddie Bray, Mathieu Laversanne, Elisabete Weiderpass and Marc Arbyn, “Geographic and temporal variations in the incidence of vulvar and vaginal cancers,” September 2018.

[47] Yong-Bo Yu, Yong-Hua Wang, Xue-Cheng Yang, Yang Zhao, Mei-Lan Wang, Ye Lian and, Hai-Tao Niu, “The relationship between human papillomavirus and penile cancer over the past decade: a systematic review and meta-analysis, *Asian Journal of Andrology*,” Vol. 21, pp. 375 – 380, May 2019.

[48] Boris Schlenker and Peter Schneede, “The Role of Human Papilloma Virus in Penile Cancer Prevention and New Therapeutic Agents,” *European Association of Urology*, Vol. 5, pp. 41 – 45, September 2018.

[49] Matthew J Rewhorn, Je Song Shin, Jane Hendry, Alastair McKay, Ross Vint, Hing Y Leung, Robert N Meddings, David S Hendry and Michael Fraser, “Rare male cancers: Effect of social deprivation on a cohort of penile cancer patients,” Journal of Clinical Urology, June 2020.

[50] WebMD, Human papillomavirus (HPV), Available on: https://www.webmd.com/sexual-conditions/hpv-genital-warts/hpv-virus-information-about-human-papillomavirus#1, [Accessed May 2020], [Updated January 2020]

[51] Jiayao Lei, Alexander Ploner, Camilla Lagheden, Carina Eklund, Sara Nordqvist Kleppe, Bengt Andrae, K. Miriam Elfstro, Joakim Dillner, Pa¨r Sparen and Karin Sundstrom, “High-risk human papillomavirus status and prognosis in invasive cervical cancer: A nationwide cohort study,” *PLOS Medicine*, October 2018.

[52] J. Schmidhuber, "Deep learning in neural networks: An overview," Neural networks, vol. 61, pp. 85-117, 2015.

[53] Dinu A.J, Ganesan R, Felix Joseph and Balaji V, “A study on Deep Machine Learning Algorithms for diagnosis of diseases," *International Journal of Applied Engineering Research*, vol. 12, pp. 6338-6346, 2017.

[54] G. Ian, B. Yoshua and C. Aaron, Deep Learning, MIT Press, 2016.

[55] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," science, vol. 313, no. 5786, pp. 504-507, 2006.

[56] Y. LeCun, Y. Bengio and G. Hinton, "Deep learning," nature, vol. 521, no. 7553, pp. 436-444, 2015.

[57] Y. Bengio, A. Courville and P. Vincent, "Representation learning: A review and new perspectives," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 8, pp. 1798-1828, 2013.

[58] A. Graves, A.-r. Mohamed and G. Hinton, "Speech recognition with deep recurrent neural networks," *2013 IEEE international conference on acoustics, speech and signal processing*, pp. 6645-6649, 2013.

[59] F. A. Gers, J. Schmidhuber and F. Cummins, "Learning to forget: Continual prediction with LSTM," 1999.

[60] G. E. Hinton, "Deep belief networks," Scholarpedia, vol. 4, no. 5, p. 5947, 2009.

[61] P. Josh and G. Adam, Deep Learning a Practitioner's Approach, Sebastopol: O'Reilly Media, 2017.

[62] Ker, J., Wang, L., Rao, J., & Lim, T., "Deep Learning Applications in Medical Image Analysis," IEEE Access, Vol. 6, pp. 9375–9389. 2018.

[63] B. Christopher M, Pattern Recognition and Machine Learning, New York: Springer-Verlag, 2006.

[64] C. Francois, Deep Learning with Python, New York: Manning Publications, 2017.

[65] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," *European conference on computer vision*, pp. 818-833, 2014.

[66] A. Krizhevsky, I. Sutskever and G. E. Hinton, "Imagenet detection with deep convolutional neural networks," Advances in neural information processing systems, pp. 1097-1105, 2012.

[67] F. Li, J. Justin and Y. Serena, "CS231n: Convolutional Neural Networks for Visual Recognition," Stanford University, spring 2018. [Online]. Available: http://cs231n.stanford.edu/index.html. [Accessed 20 May 2020].

[68] A. Ng, "Convolutional Neural Networks," corsera, [Online]. Available: https://www.coursera.org/learn/convolutional-neural-networks/lecture/hELHk/poolinglayers. [Accessed 10 June 2019].

[69] Charlotte Pelletier, Geoffrey I. Webb and Franc¸ois Petitjean, "Temporal Convolutional Neural Network for the Detection of Satellite Image Time Series," *IEEE for possible publication*, November 2018.

[70] Scott Mayer McKinney, Marcin Sieniek, Varun Godbole, Jonathan Godwin, Natasha Antropova, Hutan Ashrafian, Trevor Back, Mary Chesus, Greg C. Corrado, Ara Darzi,

Mozziyar Etemadi, Florencia Garcia-Vicente, Fiona J. Gilbert, Mark Halling-Brown, Demis Hassabis, Sunny Jansen, Alan Karthikesalingam, Christopher J. Kelly, Dominic King, Joseph R. Ledsam, David Melnick, Hormuz Mostofi, Lily Peng, Joshua Jay Reicher, Bernardino Romera-Paredes, Richard Sidebottom, Mustafa Suleyman, Daniel Tse, Kenneth C. Young, Jeffrey De Fauw & Shravya Shetty, "International evaluation of an AI system for breast cancer screening," *Nature Research*, Vol. 577, pp. 89 – 114, January 2020.

[71] Amita Das, U. Rajendra Acharya, Soumya S. Panda and Sukanta Sabut, "Deep learning based liver cancer detection using watershed transform and Gaussian mixture model techniques," *Cognitive Systems Research*, December 2018.

[72] A. Dascalu and E.O. David, "Skin cancer detection by deep learning and sound analysis algorithms: A prospective clinical study of an elementary dermoscope," *EBioMedicine*, Vol. 43, pp. 107 – 113, May 2019.

[73] Fernandes, K., Cardoso, J. S., & Fernandes, J., "Automated Methods for the Decision Support of Cervical Cancer Screening Using Digital Colposcopies", IEEE, to be published, doi: 10.1109/ACCESS.2018.2839338

[74] Mercy Nyamewaa Asiedu, Anish Simhal, Usamah Chaudhary , Jenna L. Mueller, Christopher T. Lam, John W. Schmitt, Gino Venegas, Guillermo Sapiro, and Nimmi Ramanujam, "Development of Algorithms for Automated Detection of Cervical PreCancers With a Low-Cost, Point-of-Care, Pocket Colposcope," *IEEE Transactions On Biomedical Engineering*, Vol. 66, pp. 2306-2318, August 2017.

[75] Lavanya Devi. N and P.Thirumurugan, "Automated Detection of Cervical Cancer," *International Journal of Innovative Technology and Exploring Engineering (IJITEE),* Vol. 8, pp. 2278-3057, August 2019.

[76] Mihalj Bakator and Dragica Radosav, "Deep Learning and Medical Diagnosis: A Review of Literature," *Multimodal Technologies and Interact*, August 2018.

[77] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, pp. 1-14, 2014.

[78] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich, "Going deeper with convolutions," *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1-9, 2015.

[79] R. Miikkulainen, J. Liang, E. Meyerson, A. Rawal, D. Fink, O. Francon, B. Raju, H. Shahrzad, A. Navruzyan and N. Duffy, "Evolving deep neural networks," in Artificial Intelligence in the Age of Neural Networks and Brain Computing, *Elsevier*, 2019, pp. 293-312.

[80] B. McMahan and M. Streeter, "Delay-tolerant algorithms for asynchronous distributed online learning," *Advances in Neural Information Processing Systems*, pp. 2915-2923, 2014.

## Annex A: The Proposed CNN Model Code

```
#Importing all appropriate libraries for model development, generation of detection reports and other
import sys
import os
import zipfile
from keras.preprocessing.image import ImageDataGenerator
from keras import optimizers
from keras.layers.convolutional import Conv2D
from keras.layers.convolutional import MaxPooling2D
from keras.layers.core import Activation
from keras.layers.core import Flatten
from keras.layers.core import Dense
from keras.layers.core import Dropout
from keras import callbacks
import matplotlib.pyplot as plt
import matplotlib.image as mpimg
%matplotlib inline
from keras.models import Sequential

# ARGV is a list of arguments on the command line. The number of the command-line arguments is len(sys. argv).
# argc stands for argument count, while argv stands for argument vector
DEV = False
```

```
argvs = sys.argv
argc = len(argvs)
if argc > 1 and (argvs[1] == "--development" or argvs[1] ==
"-d"):
    DEV = True
    if DEV:
        epochs = 2
        if not DEV:
            epochs = 200
train_data_path = r'D:\HPV30\Dataset\Train'
validation_data_path = r'D:\HPV30\Dataset\Val'
test_data_path = 'D:\HPV30\Dataset\Test'
img_width, img_height = 127, 127
batch_size = 32

# HPV caused cancer detection model development or creation

model = Sequential()
model.add(Conv2D(32, (5, 5),strides=2,
input_shape=(127,127,3)))
model.add(Activation('relu'))
model.add(MaxPooling2D(pool_size=(3, 3), strides=1))
model.add(Conv2D(32, (3, 3),strides=1))
model.add(Activation('relu'))
model. add(MaxPooling2D(pool_size=(2, 2), strides=1))
```

```
model.add(Conv2D(64, (3, 3), strides=1))
model.add(Activation('relu'))
model.add(Conv2D(64, (5, 5),strides=2))
model.add(Activation('relu'))
model.add(Conv2D(64, (3, 3), strides=1))
model.add(Activation('relu'))
model.add(MaxPooling2D(pool_size=(2, 2),strides=2))
model.add(Flatten())
model.add(Dense(64))
model.add(Activation('relu'))
model.add(Dropout(0.5))
model.add(Dense(64))
model.add(Activation('relu'))
model.add(Dropout(0.5))

model.add(Dense(2))
model.add(Activation('sigmoid'))

# defines the loss function, the optimizer and the metrics

model.compile(loss='binary_crossentropy',
              optimizer=optimizers.Adam(lr=0.001),
              metrics=['accuracy'])
model.summary()
```

```
# Image Data Generator for Augmentation
train_datagen = ImageDataGenerator(
     rescale = 1./255,
     shear_range = 0.3,
     width_shift_range = 0.2,
     height_shift_range = 0.2,
     zoom_range = 0.2,
     horizontal_flip = True
)

img_height = 127
img_width = 127
batch_size = 32
train_data_path = r'D:\HPV30\Dataset\Train'
validation_data_path = r'D:\HPV30\Dataset\Val'
test_data_path = r'D:\HPV30\Dataset\Test'
test_datagen = ImageDataGenerator(rescale = 1./255)
train_generator = train_datagen.flow_from_directory(
      directory = train_data_path,
      target_size = (img_height, img_width),
      batch_size = batch_size,
      class_mode='binary'
)
validation_generator = test_datagen.flow_from_directory(
      directory = validation_data_path,
```

```
        target_size = (img_height, img_width),
        batch_size = batch_size,
        class_mode = 'binary'
 )
test_generator = test_datagen.flow_from_directory(
    directory = test_data_path,
target_size = (img_height, img_width),
batch_size = batch_size,
shuffle = False,
class_mode = 'binary'
)

#Train the Sample Images
nb_train_samples = len(train_generator.filenames)
nb_classes = len(train_generator.class_indices)
nb_validation_samples = len(validation_generator.filenames)
nb_test_samples = len(test_generator.filenames)

#History
history = model.fit(
train_generator,
steps_per_epoch = int(nb_train_samples/batch_size),
epochs = 30,
validation_data = validation_generator,
validation_steps = int(nb_validation_samples/batch_size))
```

```
#Lets plot the train and val. curve
#get the details form the history object
import matplotlib.pyplot as plt
%matplotlib inline
acc = history.history['accuracy']
val_acc = history.history['val_accuracy']
loss = history.history['loss']
val_loss = history.history['val_loss']
epochs = range(1, len(acc) + 1)

#Graphing our training and validation
acc = history.history['accuracy']
val_acc = history.history['val_accuracy']
loss = history.history['loss']
val_loss = history.history['val_loss']
epochs = range(len(acc))
plt.plot(epochs, acc, 'r', label='Training acc')
plt.plot(epochs, val_acc, 'b', label='Validation acc')
plt.title('Training and validation acc')
plt.ylabel('Accuracy')
plt.xlabel('Epoch')
plt.legend()
plt.figure()
plt.plot(epochs, loss, 'r', label='Training loss')
```

```
plt.plot(epochs, val_loss, 'b', label='Validation loss')

plt.title('Training and validation loss')

plt.ylabel('Loss')

plt.xlabel('Epoch')

plt.legend()

plt.show()


#Meann Accuracy and Mean Loss of the HPV caused cancer
detection model

import numpy as np

print("Training Accuracy:",
np.mean(history.history['accuracy']))

print("Validation Accuracy: ",
np.mean(history.history['val_accuracy']))

print("Training Loss:", np.mean(history.history['loss']))

print("Validation Loss, ",
np.mean(history.history['val_loss']))


#Test Loss and Test Accuracy of the HPV caused cancer
detection model

#predict against unseen data

test_eval =
model.evaluate_generator(test_generator,nb_validation_samples
//32, verbose=1)

print('Test loss:', test_eval[0])

print('Test accuracy:', test_eval[1])
```

**Declaration**

I, the undersigned, declare that this thesis is my original work and has not been presented for a degree in any other university, and that all source of materials used for the thesis have been duly acknowledged.

**Declared by:**

**Name: Solomon Gezahegn Temesgen**

**Signature: _____**

**Date: _____**

**Confirmed by advisor:**

**Name: _____**

**Signature: _____**

**Date: _____**