2021-09-29

HUMAN SKIN DISEASE DETECTION AND CLASSIFICATION MODEL USING DEEP CONVOLUTIONAL NEURAL NETWORK

DINKIE, MINICHEL SHIFERAW

BAHIR DAR UNIVERSITY

BAHIR DAR INSTITUTE OF TECHNOLOGY

SCHOOL OF GRADUATE STUDIES

FACULTY OF COMPUTING

DEPARTMENT OF INFORMATION TECHNOLOGY

HUMAN SKIN DISEASE DETECTION AND CLASSIFICATION MODEL USING DEEP CONVOLUTIONAL NEURAL NETWORK

MSc Thesis

BY: -

DINKIE MINICHEL SHIFERAW

September 29, 2021

Bahir Dar, Ethiopia

HUMAN SKIN DISEASE DETECTION AND CLASSIFICATION MODEL USING DEEP CONVOLUTIONAL NEURAL NETWORK

MSc. Thesis

By: -

Dinkie Minichel Shiferaw

A thesis submitted to the school of Graduate Studies of Bahir Dar Institute of Technology, BDU in partial fulfillment of the requirements for the degree of Masters of Science in information technology in the faculty of computing.

Advisor Name: Dagnachew Melesew (Assistant Professor)

September 29, 2021

Bahir Dar, Ethiopia

DECLARATION

This is to certify that the thesis entitled "**Human Skin Disease Detection and Classification Model Using Deep Convolutional Neural Network**", submitted in partial fulfillment of the requirements for the degree of Master of Science in Information Technology under Faculty of Computing, Bahir Dar Institute of Technology, is a record of original work carried out by me and has never been submitted to this or any other institution to get any other degree or certificates. The assistance and help I received during the course of this investigation have been duly acknowledged.

Name of the student: - Dinkie Minichel Signature ____________

Date of submission: Oct 20, 2023

Place: Bahir Dar

BAHIR DAR UNIVERSITY
BAHIR DAR INSTITUTE OF TECHNOLOGY
SCHOOL OF GRADUATE STUDIES
FACULTY OF COMPUTING
THESIS APPROVAL SHEET

I hereby confirm that the changes required by the examiners have been carried out and incorporated in the final thesis.

Name of Student ***Dinkie Minichel Shiferaw*** Signature__________ Date Oct 20/2021

As members of the board of examiners, we examined this thesis entitled "***Human skin disease detection and classification model using deep Convolutional neural network***" by Dinkie Minichel. We hereby certify that the thesis is accepted for fulfilling the requirements for the award of the degree of Masters of Science in Information Technology

Board of Examiners:

Name of Advisor:	Signature	Date
Dagnachew Melisew (Assis. Prof)		NOV 24/2021
Name of External Examiner	Signature	Date
Fekade Getahun(Dr.)		20/10/21
Name of Internal Examiner:	Signature	Date
Birhani Hailu (Dr)		NOV 23/2021
Name of Chair Person:	Signature	Date
Seffi Gebeyehu (Assis.Prof)		Nov 23/21
Name of Chair Holder:	Signature	Date
Abdulkerim Mohammed (PhD)		NOV 24, 2021
Name of Faculty Dean:	Signature	Date
Asegahegn E.		Nov, 2021

Faculty Stamp

Bahir Dar University Bahir Dar Institute Of Technology Faculty Of Computing

ACKNOWLEDGMENT

First and foremost, I would like to thanks my God and Ever-Virgin, St. Marry, Mother of our Lord for giving me the strength, knowledge, ability and opportunity to undertake this thesis work.

I would like to express the greatest appreciation to my Advisor, Assistance Professor Mr, Dagnachew Melesew. I am deeply thankful to him for giving his insights and constructive comments, continuous guidance of this research.

I also thanks to Amhara Science, Technology and Information Communication Commission for providing me the chance, the required finance for under taking this study. I am also my thanks for staff members of Felege Hiwot Comprehensive Referral Hospital, Gamby General Teaching Hospital and Tibebe Giyon Specialized Hospital in Bahir dar city. Individually, I would like thanks to Dr Mulu Melkie and Dr. Samirawit Getachew for their continuous support throughout data collection time.

It is a pleasure to express my gratitude whole heartedly to my husband, Derebew Azanaw and my son Cherenet Derebew, for their praised for their understanding, patience, extraordinary love and encouragement during this thesis work.

I am very thankful to my mother, Kassaye Aragaw, my father, Minichel Shiferaw and my brother Birhanu Minichel for encouraging and supporting me in all my studies starting from early schools.

Finally, I am very grateful thanks to my brothers and the rest of all my families and friends who help me in one or the other way to a success in my academic work.

ABSTRACT

The skin is the biggest organ of the human body which protects our inner vital parts and organs from the outside environment. However, there are a number of diseases that affect the human skin such as, fungus, bacteria, allergies, enzyme, and viruses that have caused to skin abnormalities and need to be treated at earlier stages to avoid it from spreading. Identifying the disease type based on manual feature extractions or the symptoms is time consuming and requires extensive knowledge for perfect identification. The main objective of this study was developing human skin disease detection and classification model using deep convolutional neural network. The images of human diseased skin Tinea capitis, Tinea pedis Tinea corporis and Healthy skin have been collected from Tibebe Giyon Specialized Hospital, Felege Hiywot Comprehensive Referral Hospital, Gamby General Teaching Hospital using Techno CAMONX smartphones camera 14 mega pixel and DermNet.com image repository in jpg file format. In this study, a total of 4 classes and a total of 2,226 images are used. After collected the dataset, image augmentation, image preprocessing, thresholding segmentation, combined (CNN and BRISK) had been used as feature extraction techniques and SVM and SoftMax for classifier. And also, the researcher has applied Principal Component Analysis (PCA) to reduce the dimension of the combined features. The researcher has used MATLAB R2019a programming tools for overall coding mechanisms. From the experiment, the model achieved the testing accuracy of 88.9% and training accuracy of 98.44%.

Key words: *CNN, SVM, Local Feature Descriptor, Feature Extraction, PCA*

TABLE OF CONTENTS

LIST OF ABBREVATIONS

ADAM	Adaptive Moment Estimation
SGDM	Stochastic Gradient Descent with Momentum
AI	Artificial Intelligence
BIT	Bahir Dar Institute of Technology
CNN	Convolutional neural network
DL	Deep learning
DIP	Digital Image Processing
DNN	Deep Neural Network
GHE	Global Histogram Equalization
HSDDC	Human Skin Disease Detection and Classification
HEq	Histogram Equalization
ICT	Information communication Technology
LHE	Local Histogram Equalization
MLP	Multilayer Perceptron Classifier
PCA	Principal Component Analysis
SVM	Support Vector Machine
TANH	Hyperbolic Tangent
RELU	Rectified Linear Unit
ResNet	Residual neural network
UV	Ultraviolet radiation

LIST OF FIGURES

LIST OF TABLES

LIST OF ALGORITHMS

CHAPTER ONE

1. INTRODUCTION

1.1. Background

Skin is the biggest organ and most sensitive part of the human body which protects our inner vital parts and organs from the outside environment. Skin also helps in body temperature regulation. The skin contains of cells, pigmentation, blood vessels, and other components (V. Pugazhenthi, 2019). It is made up of 3 layers, namely, epidermis, dermis, and subcutaneous tissues. Epidermis, being the outermost skin layer, forms a waterproof and protective sheath around the body’s surface. The dermis, found beneath the epidermis, comprises of connective tissues and protects the body from stress and strain. A basement membrane tightly joins the dermis with the epidermis. The hypodermis, also called subcutaneous tissue, is not actually a part of the skin and lies below the dermis. It attaches the skin to the underlying bone and muscle and also supplies blood vessels and nerves to it (V. Pugazhenthi, 2019).

Skin perceives the outside condition and shields our inside organs and tissues from unsafe microscopic organisms, contamination and sun presentation. (Shuchi, et, al 2019). However, the Human skin infected by fungal infection, bacteria, allergy, or viruses, etc. Skin diseases are more common than other diseases. (Nawal Soliman, 2019).

Dermatological disorders are one of the most widespread diseases in the world. Despite being common its diagnosis is extremely difficult because of its complexities of skin tone, color, presence of hair (Sourav Kumar, 2018). The skin disease diagnosis includes series of pathological laboratory tests for the identification of the correct disease. For the past ten years these diseases have been the matter of concern as their sudden arrival and their complexities have increased the life risk. These Skin abnormalities are very infectious and need to be treated at earlier stages to avoid it from spreading (Sourav Kumar, 2018).

Common diseases that attack human skin are eczema, melanoma, vitiligo, mycosis, papillomas, impetigo, scabies, herpes, dermatitis, wart, psoriasis, acne, tinea capitis, tinea corporis, tinea pedis, etc (Haymanot, 2020).

Diseases like tinea corporis, tinea pedis, tinea capitis, Eczema and scabies are common in developing countries like Ethiopia. In western Ethiopia, more than 80 percent of randomly examined school children had at least one skin disease, which was usually caused by one, 2 of four conditions: scabies, capitis, tinea capitis, or pyoderma (Haymanot, 2020). Skin diseases are hazardous and often contagious, especially melanoma, eczema, and impetigo. These skin diseases can be cured if detected early (Jayashree, 2019).

So, in recent years there has been an increased interest in applying image processing techniques to the problem of skin disease identification. There are opportunities to improve skin lesion identification through the designing of a convenient lesion identification system. Many different approaches are used to classify skin lesion to its predefined classes using the features of lesion. (Haymanot, 2020).

There have been different research works about skin disease identification algorithms and techniques. The Authors (Manish Kumar and Rajiv Kumar, 2016) proposed an Intelligent System to Diagnosis the Skin Disease; such as Acne and psoriasis using KNN neural network. According to (Mrs. Jayashree et al, 2019) proposed mobile based skin disease; such as melanoma, eczema, and impetigo using Image Processing with Data Mining and Deep Learning. As referenced by (Padmavathi et al, 2020) have been studied Skin Diseases Prediction using CNN and KNN algorithms. The authors (Li-sheng, et al ,2018) Skin Disease Recognition Method Based on Image Color and Texture Features.

According to (Leelavathy et, al, 2020) have been proposed skin disease detection using hybrid architecture with image processing and machine learning techniques to predict type of disease. And also, (Samuel et al, 2019) proposed A Web-Based Skin Disease Diagnosis (atopic dermatitis, acne vulgaris, and scabies) Using Convolutional Neural Networks as a classifier. On the other hand, (Prem, et al, 2020) have been studied Skin Disease Detection techniques by using KNN as feature extraction and k-means as classifier.

In this study, we considered 4 classes of Human skin diseases including Tinea capitis, Tinea Corporis, Tinea pedis, and Healthy skin. The images are captured by Techno CAMONX smartphones camera 14 mega pixel and images are collected from Tibebe Giyon Specialized Hospital, Felege Hiywot Comprehensive Specialized Hospital and Gamby General Teaching Hospital in Bahir dar city and also DermNet.com image repositories. The proposed model was made by taking the input images followed by preprocessing techniques and segmentation of the diseased skin area. Then Feature extraction of the affected area using CNN and BRISK local feature and classification done by with the help of CNN and SVM.

1.2. Statement of the Problems

The human skin is the biggest organ of the human body and plays a huge part in a person's physical appearance. But the Skin infected by fungal infection, bacteria, allergy, or viruses, etc. (Nawal Soliman, 2019). Skin diseases are more common than other diseases. A skin disease may change texture or color of the skin. At large, skin diseases are chronic, infectious and sometimes may develop into skin cancer. Therefore, skin diseases must be diagnosed early to reduce their development and spreading (Nawal Soliman, 2019).

The advancement of lasers and Photonics based medical technology has made it possible to diagnose the skin diseases much more quickly and accurately. But the cost of such diagnosis is still limited and very expensive. So, image processing techniques help to build automated screening system for dermatology at an initial stage (Nawal Soliman, 2019). Machine learning plays an essential role in the medical field for the automation of many processes. It has been demonstrated that dermoscopy may actually lower the diagnostic accuracy in the hands of inexperienced dermatologists (Sumithra et al, 2015).

So that, the Researchers were, motivated towards the replacement of the above-mentioned manual methods with automated system, which was faster and more accurate.

Identification of skin disease has been done by different researchers. The authors (Sumithra et al, 2015) segmentation and classification of skin lesions are proposed, lesion areas are segmented using region growing method and Color and texture features are extracted to represent segmented lesion areas. The classification is performed with SVM, KNN and fusion of SVM and KNN Classifiers. The image processing and machine learning approach studied by the author (Nawal Soliman 2019). Discriminative Feature Learning for Skin Disease Classification using Deep CNN also studied by the Authors (Belal et al,2020).

To fill the gap mentioned in the above, we have to develop using a hybrid of high-level CNN features and local image features. In this research work, we develop Human skin disease detection and classification model using deep convolutional neural network. To this end, the research is conducted to answers the following research questions:

- Which image preprocessing techniques, are suitable for identifying human skin disease image?
- Which feature extraction techniques are best to identify human skin disease?
- Which classification techniques are more suitable for building a model for human skin disease detection and classification?
- To what extent the classifier identifies the human skin disease?

1.3. Objective of the study

1.3.1. General Objectives

The general objective of this thesis work is to developing human skin disease detection and classification model using deep convolutional neural network.

1.3.2. Specific objectives

To achieve the general objective, the following specific objectives are identified as follows.

- To collect, organize and prepare skin disease image dataset for training and testing.
- To select suitable image preprocessing techniques, feature extraction and classification techniques for the identification of human skin disease image.
- To finding best feature extraction techniques that is identifying Human skin disease.
- To develop a model for identify Human skin diseases.
- To evaluate the performance of the proposed model using test dataset.

1.4. Scope and limitations of the study

The basic objective of this study is to detection and classification of fungal Human skin diseases such as Tinea pedis, Tinea capitis, Tinea corporis and Healthy skin. The Skin Disease's symptoms which cannot be identified by naked eyes and not possible to capture by camera are not included in this study.

1.5. Significance of the study

The result of this study has the following advantages: -

- The model that can assist Dermatologists to detect and classify patients' disease type. This helps for experts to make the accurate decision.

- Patients spent much amount of time, labor and money in hospitals. Most of the patients are come from another place. So, patients can easily get the medication by the help of this model.
- In Ethiopia skin disease professionals are few in number and it has low service quality than compared with other developing country. Skin disease professionals will be used this skin disease identification model for easily diagnosis the patient. This reduces serious economic costs and save time.
- The models have great advantage for researchers, health policy makers and health professionals.

1.6. Thesis organization

This thesis work has five chapters that including the Introduction chapter. The thesis has organized as follows:

Chapter 2: In this chapter, a literature review on topics related to the objective of this research is presented. The areas that are directly related to this research are covered in detail.

Chapter 3: In this chapter, the methodology (methods and materials) that is used to implement the proposed detection and classification model is discussed briefly.

Chapter 4: In this chapter, the implementation of skin disease detection and classification model is presented and experimental results are discussed.

Chapter 5: Conclusion and recommendation of this thesis work is presented.

CHAPTER TWO

2. LITERATURE REVIEW AND RELATED WORK

2.1. Overview

Image processing, is a set of computational techniques for analyzing, enhancing, compressing, and reconstructing images. Its main components are importing, in which an image is captured through scanning or digital photography; analysis and manipulation of the image, accomplished using various specialized software applications; and output (e.g., to a printer or monitor). Image processing has extensive applications in many areas, including astronomy, medicine, industrial robotics, and remote sensing by satellites (Britannica, july 2021). Nowadays, image processing is among rapidly growing technologies and forms core research area within Engineering and computer science disciplines too.

In this chapter first, we describe the literature related to Human skin disease, are presented. Then, Different techniques of digital image processing (image acquisition, image preprocessing, and types of segmentation, feature extraction and classification), and deep learning are discussed in detail.

2.2. The Structure of Human Skin

Skin is the largest organ in the body that covers the entire external surface. It protects the internal organs from germs and thus helps prevent infections. It surrounded all other parts of human body (Haymanot, 2020). In terms of surface area skin is the second largest organ in the human body. The skin thickness varies over all parts of the body, between men and women and young and old. For the average adult human, the surface area is between 1.5 and 2.0 square meters. Averagely the skin holds 650 sweat glands, 20 blood vessels, 60,000 melanocytes, and more than 1,000 nerve endings (Haymanot, 2020).

A skin cell usually ranges from 25-40 micrometers squared, depending on a variety of factors. The average human skin cell is about 30 micrometers in diameter. The skin is divided into several layers. The epidermis is composed mainly of keratinocytes. Beneath the epidermis is the basement membrane (also known as the dermo-epidermal junction); this narrow, multilayered structure anchors the epidermis to the dermis. The layer below the dermis, the hypodermis, consists largely of fat (Lawton, 2019)

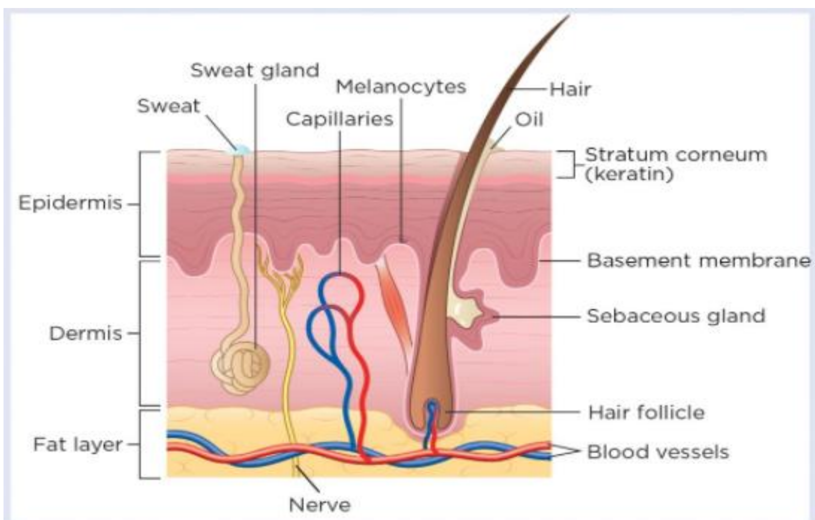


Figure 1: The structure of Human skin (Lawton, 2019)

Epidermis

The epidermis is the outer (top) layer of the skin, defined as a stratified squamous epithelium, primarily comprising keratinocytes in progressive stages of differentiation (Lawton, 2019). Keratinocytes produce the protein keratin and are the major building blocks (cells) of the epidermis. As the epidermis is a vascular (contains no blood vessels), it is entirely dependent on the underlying dermis for nutrient delivery and waste disposal through the basement membrane.

The prime function of the epidermis is to act as a physical and biological barrier to the external environment, preventing penetration by irritants and allergens. At the same time, it prevents the loss of water and maintains internal homeostasis (Lawton, 2019). The epidermis is composed of layers; most body parts have four layers, but those with the thickest skin have five. The layers are: Stratum corneum (horny layer), Stratum lucidum (only found in thick skin – that is, the palms of the hands, the soles of the feet and the digits), Stratum granulosum (granular layer), Stratum spinosum (prickle cell layer), and Stratum Basale (germinative layer) (Lawton, 2019).

Dermis

The dermis is the inner layer of the skin and is much thicker than the epidermis (1-5mm) (Lawton, 2019). Situated between the basement membrane zone and the subcutaneous layer, the primary role of the dermis is to sustain and support the epidermis. The main functions of the dermis are Protection, Cushioning the deeper structures from mechanical injury, providing nourishment to the epidermis and playing an important role in wound healing (Lawton, 2019).

Hypodermis

The hypodermis is the innermost (or deepest) and thickest layer of the skin. It is also known as the fat layer or subcutaneous tissue. It is mostly made up of fat. It lies between the dermis and muscles or bones and contains blood vessels that expand and contract to help keep your body at a constant temperature (Lawton, 2019). The hypodermis also protects your vital inner organs. Reduction of tissue in this layer causes your skin to sag.

2.3. Types of skin Disease

Fungal skin disease

Infections caused by pathogenic fungi and limited to the human hair, nails, epidermis, and mucosa are referred to as superficial fungal infections. Despite the fact that these infections rarely are dangerous or life threatening, detect at early stage is important because of their worldwide distribution, frequency, person-to-person transmission, and morbidity (Haiymanot, 2020). Dermatophytes are the most common agents of superficial fungal infections worldwide and widespread in the developing countries, especially in the tropical and subtropical countries like India and Sub-Saharan Africa, where the environmental temperature and relative humidity are high. Other factors such as increased urbanization including the use of occlusive footwear and tight fashioned clothes, has been linked to higher prevalence. Most common fungal infections are tinea corporis, Tinea capitis, tinea pedis, Tinea cruises, Pityriasis versicolor and etc (Haymanot, 2020). Among them we included the most frequent occurred fungal skin diseases those are Tinea capitis, tinea corporis, tinea pedis and Healthy skin.

Tinea refers to superficial infection with one of the three fungal generals called Microsporum, Epidermophyton, and Trichophyton collectively known as dermatophytes. These infections are among the most common diseases worldwide and cause serious chronic morbidity (Haymanot, 2020). Tinea infections are caused by dermatophytes and are classified by the involved site. The most common infections in children are tinea corporis and tinea capitis, whereas adolescents and adults are more likely to develop tinea cruris, tinea pedis, and tinea unguium.

Tinea Capitis

Tinea capitis is the most common superficial fungal infection that affects the skin of the scalp and the associated hair shafts. It's most common in young children. The symptoms can include localized bald patches that may appear scaly or red, associated scaling and itching associated tenderness or pain in the patches (seladi, 2020).

Figure 2: Tinea Capitis

Tinea Corporis

Tinea Corporis is a fungal infection that can affect your skin and scalp. It's caused by Dermatophytes. Tinea Corporis is also part of a group of fungi that grow on skin, particularly in damp and humid parts of your body. The Symptom usually starts as a reddish, itchy, scaly rash. Over time, patches of ringworm can spread and form red rings. And also patches that get blisters and start to ooze, bald patches on the scalp, patches that look like rings with a redder outside edge and thick, discolored, or cracked nails (if the infection is in the nails)

Figure 3: Tinea Corporis

Tinea Pedis

Tinea pedis is also known as Athlete's foot. It's a type of fungal infection that can affect the skin on your feet, as well as your hands and nails. The infection is caused by Dermatophytes, a group of fungi that can thrive in the warm and humid areas between your toes. It's particularly common among athletes and can spread from one person to another. You can also catch it from contaminated surfaces, like a public shower or locker room floors (Seladi, 2020).

Figure 4: Tinea pedis

2.4. Digital Image processing

Digital Image Processing means processing digital image by means of a digital computer. It is a use of computer algorithms, in order to get enhanced image either to extract some useful information. It is a type of signal processing in which input is an image and output may be image or characteristics/features associated with that image. In today's digital life, digital images are everywhere around us. An image is a visual representation of an object, a person, or a scene. (Tyagi, 2018). A digital image is a two-dimensional function f(x, y) that is a projection of a 3-dimesional scene into a 2-dimensional projection plane, where x, y represents the location of the picture element or pixel and contains the intensity value. When values of x, y and intensity are discrete, then the image is said to be a digital image. Mathematically, a digital image is a matrix representation of a two-dimensional image using a finite number of points cell elements, usually referred to as pixels (picture elements, or pixels). Each pixel is represented by numerical values: for grayscale images, a single value representing the intensity of the pixel (usually in a [0, 255] range) is enough; for color images, three values (representing the amount of red (R), green (G), and blue (B)) are stored. If an image has only two intensities, then the image is known as a binary image (Tyagi, 2018).

Most Digital image processing application for image classification follows similar steps such as: image acquisition, image preprocessing, segmentation, feature extraction and classification (Desalegn, 2020).

2.4.1. Image acquisition

Image acquisition is the first steps of digital image processing and the process of gathering relevant data. Image Acquisition is to transform an optical image (Real World Data) into an array of numerical data which could be later manipulated on a computer, before any video or image processing can commence an image must be captured by camera and converted into a manageable entity. (Vikas et al, 2017).

2.4.2. Image preprocessing

Image pre-processing is an essential step of detection in order to remove noises and enhance the quality of original image. The main purpose of this step is to improve the quality of image by removing unrelated and surplus parts in the back ground of image. Good selection of preprocessing techniques can greatly improve the accuracy of the system (Adel and Afsaneh, 2014). In this study, image resizing, noise removal filtering, image enhancement, grayscale conversion and thresholding segmentation is used as preprocessing techniques.

2.4.2.1. Image resizing

Image resizing is necessary when you need to increase or decrease the total number of pixels, whereas remapping can occur under a wider variety of scenarios: correcting for lens distortion, changing perspective, and rotating an image. Image Resizing is important for displaying visual media at different resolutions and aspect ratios. Image Resizing can be more effectively achieved with better understanding of image semantics (Dighe and Guru, 2014). There are a number of images resizing techniques

Nearest neighbor

Nearest neighbor it is a simplest interpolation. In this method each interpolated output pixel is assigned the value of the nearest sample point in the input image. This has the effect of simply making each pixel bigger (Shreyas, 2014).

Bilinear interpolation

Bilinear interpolation is used to know values at random position from the weighted average of the four closest pixels to the specified input coordinates, and assigns that value to the output coordinates. The two linear interpolations are performed in one direction and next linear interpolation is performed in the perpendicular direction. (Shreyas, 2014).

Bi-cubic interpolation

The bi-cubic interpolation is advancement over the cubic interpolation in two dimensional regular grids. The interpolated surface is smoother than corresponding surfaces obtained by above mentioned methods bilinear interpolation and nearest-neighbor interpolation. It uses polynomials, cubic or cubic convolution algorithm. (Shreyas, 2014). The Cubic Convolution Interpolation determines the grey level value from the weighted average of the 16 closest pixels to the specified input coordinates, and assigns that value to the output coordinates, the first four one-dimension. For Bi-cubic Interpolation (cubic convolution interpolation in two dimensions), the number of grid points needed to evaluate the interpolation function is 16, two grid points on either side of the point under consideration for both horizontal and perpendicular direction. (Shreyas, 2014).

Basic-splines (B-spline)

The nearest neighbor and bilinear interpolations compromise the quality of image over efficiency due to rectangular shape in the pass band and infinite side lobes. The B-spline interpolations smoothly connect polynomials with pieces. (Shreyas, 2014).

2.4.2.2. Noise Removal

Noise represents unwanted information which deteriorates image quality. Noise is a random variation of image intensity and visible as grains in the image. Noise means, pixels within the picture present different intensity values rather than correct pixel values. Noise originates from the physical nature of detection processes and has many specific forms and causes (Kaur, 2015).

Mean Filter: The mean filter is a simple spatial filter. Mean filter acts on an image by smoothing it. The mean filter is a simple sliding window spatial filter that replaces the center value in the window with the average of all the neighboring pixel values including it. This process is repeated for all pixel values in the image (Kaur, 2015).

Median Filter: Median filter is a simple and powerful non-linear filter which is based on order statics, whose response is based on the ranking of pixel values contained in the filter region. It is easy to implement method of smoothing images. The median filter also follows the moving window principle similar to the mean filter. A 3*3, 5*5, or 7*7 kernel of the pixels is scanned over pixel matrix of the entire image. In this filter, we do not replace the pixel value of the image with the mean of all neighboring pixel values; we replace it with the median value. Median filtering is done by, first sorting all the pixel values from the surround's neighborhood into numerical order and then replacing the pixel being considered with the middle pixel value (Kaur, 2015).

Wiener filter

The wiener filter method takes out noise from each pixel in an image using statistical approach. It performs an optimal transaction between inverse filtering and noise smoothing. It is a best filter in the process of noise smoothing to reduce the overall MSE. It tries to build an image by imposing a MSE constraint between the estimated and the original image (Maheshan, 2019).

2.4.2.3. Image enhancement

Image enhancement is a subfield of digital image processing. The purpose of image enhancement is to improve the contrast and sharpening the image to enable for further processing or analysis. It is the purpose of adjusting digital images so that the results are more suitable for display or further image analysis. Image enhancement improves the quality and the information content of original data before processing (Gan, 2020). The enhancement does not raise the inbuilt information content of the data other than it increases the dynamic range of the selected facial appearance as a result that they can be detected (Gan, 2020).

2.4.2.4. Gray Scale Conversion

Gray scale conversion it contains brightness information only and no color information. The gray scale data matrix values represent intensities images. (Enquahone,2017) The conversation of RGB image into gray scale image is done using rgb2gray () MATLAB function and it works by forming a weighted sum of R, G and B components as shown below equation 2.1(Enquahone,2017).

$$\text{Gray} = 0.2989 * R + 0.5870 * G + 0.1140 * B \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2.1)$$

2.4.3. Image Segmentation

In digital image processing, image segmentation is a most commonly used technique and analysis to partition an image into multiple regions, often based on the characteristics of the pixels in the image (Jeevitha et al, 2020). Image segmentation is the process of separating foreground from background, or clustering regions of pixels based on similarities in color or shape. (Jeevitha et al, 2020). During segmentation, the image usually is subdivided until the objects of interest were isolated from their background. There are currently many different ways of performing image segmentation, ranging from the simple thresholding method to advanced color image segmentation methods (Enquahone, 2017). These parts normally correspond to something that humans can easily separate and view as individual objects. Computers have no means of intelligently recognizing objects, and so many different methods have been developed in order to segment images (Enquahone, 2017).

Most of the image segmentation algorithms are based on two basic properties of intensity values: based on discontinuity and similarity (Jeevitha et al, 2020). The first approach is to segment the image based on abrupt change in intensity, such as edged in the images. The second approach is based on some pre-defined criteria upon which images are partitioned (Jeevitha et al, 2020). Segmentation can be classified as: Edge Detection, Threshold Based Image Segmentation, Region Based Image Segmentation, Feature Based Clustering Image Segmentation and Neural Network Based Image Segmentation. In this study, Global thresholding segmentation techniques are used.

2.4.4. Feature Extraction

As Kumar and Bhatia stated that the main goal of feature extraction is to obtain the most relevant information from the original data and represent that information in a lower dimensionality space. Feature extraction is describing the relevant shape information contained in a pattern so that the task of classifying the pattern is made easy by a formal procedure. In pattern recognition and in image processing, feature extraction is a special form of dimensionality reduction (Kumar and Bhatia, 2014). When the input data to an algorithm is too large to be processed and it is suspected to be redundant (much data, but not much information) then the input data will be transformed into a reduced representation set of features also called features vector. Transforming the input data into the set of features is called feature extraction. According to Awad, & Hassaballah, there are two types of features that are global and local features. Global features describe the image as a whole and can be interpreted as a particular property of the image involving all pixels; while, the local features aim to detect key points within the image and describe regions around these key points. After extracting the features and their descriptors from images, matching of common structures between images (i.e., features matching) is the next step for these applications (Awad, & Hassaballah, 2016). Local features refer to a pattern or distinct structure found in an image, such as a point, edge, or small image patch and we described as follows.

Scale Invariant Feature Transformation (SIFT) is both feature detector and feature descriptor. SIFT transforms an image into a large collection of local feature vectors, each of which is invariant to image translation, scaling, and rotation, and partially invariant to illumination changes and affine. It was proposed by Lowe and gives feature that are invariant to affine distortion, scale, illumination changes, noise, rotation and 3D viewpoint changes (Nikita, et al, 2016)

Speeded-Up Robust Feature (SURF) is a local invariant fast feature point detector and distinctive feature point descriptor. It is also used for object recognition, registration, 3D reconstruction, and classification. It is inspired by SIFT descriptor. It (standard version) is faster than SIFT. It uses Haar wavelet. SURF is better in terms of speed. (Nikita, et al 2016)

Histogram of Gradient (HOG) This descriptor used in image processing and computer vision for detection of object. In localized areas of images this technique counts appearances of gradient orientation.it is similar to SIFT descriptor, edge orientation histogram and shape context, but it is differed in terms of computed dense grid of evenly spaced cells. To improve accuracy, it uses overlying local contrast localization. (Nikita, et al 2016)

Binary Robust Invariant Scalable Key Points (BRISK) is a point-feature detector and descriptor. BRISK uses a circular symmetric pattern region shape and an aggregate of 60 point-pairs as line segments arranged in 4 concentric rings. BRISK takes point-pairs of both long segments and short segments, and this supports a measure of scale invariance, since coarse resolution may be mapped better by long segments and fine resolution may be mapped by short segments. The Algorithm is unique. (Nikita, et al, 2016). BRISK is faster than SIFT and SURF, while using less computational resource.

Binary robust independent elementary features (BRIEF): Is a binary feature point descriptor that employs simple tests using intensity difference to create binary feature vectors that effectively describe key points in a pair of image patches.

2.4.4.1. Conventional Neural Network

Convolutional Neural Network (CNN), also called ConvNet, is a type of Artificial Neural Network (ANN), which has deep feed-forward architecture and has amazing generalizing ability as compared to other networks with FC layers, it can learn highly abstracted features of objects especially spatial data and can identify them more efficiently (Ghosh, et. al, 2020). A deep CNN model consists of a finite set of processing layers that can learn various features of input data (e.g., image) with multiple level of abstraction. The initiatory layers learn and extract the high-level features (with lower abstraction), and the deeper layers learns and extracts the low-level features (with higher abstraction). (Ghosh, et. al, 2020).

In CNN, the classification layer and the feature extraction layers learn together, that makes the output of the model more organized and makes the output more dependent to the extracted features. The implementation of a large network is more difficult by using other types of neural networks rather than using Convolutional Neural Networks. Nowadays CNN has been emerged as a mechanism for achieving promising result in various computer vision-based applications like image classification, object detection. Ghosh, et. al, 2020 stated that a CNN is composed of multiple building blocks known as layers of the architecture,

Building blocks of CNN

Convolutional Layer is the most important component of any CNN architecture. It contains a set of convolutional kernels (also called filters), which gets convolved with the input image (N-dimensional metrics) to generate an output feature map (Ghosh, et. al, 2020). A kernel can be described as a grid of discrete values or numbers, where each value is known as the weight of this kernel. During the starting of training process of an CNN model, all the weights of a kernel are assigned with random numbers (different approaches are also available there for initializing the weights). Then, with each training epoch, the weights are tuned and the kernel learned to extract meaningful features (Ghosh, et. al, 2020). The main target of using a filter (kernel) is to slide over an image and after that to get feature vectors as an output. It works by separating an image into smaller blocks and then convolving each block with a particular set of elements (by multiplying the filter elements with their corresponding blocks (receptive field) elements.In the first convolutional layer, simple feature as edges is filtered by the neurons. The neurons learn to gather the information to gain a bigger picture of the image in the following convolutional layers, hence a high-order feature detecting (Yu Han Liu, 2018).

Ghosh, et. al, 2020 stated that, In CNN the input is a multi-channeled image (e.g., for RGB image, it is 3 channeled and for Gray-Scale image, it is single channeled). Now, to understand the convolution operation, if we take a gray-scale image of 4 × 4 dimension and a 2 × 2 kernel with randomly initialized weights. Now, in convolution operation, we take the 2 × 2 kernel and slide it over all the complete 4×4 image horizontally as well as vertically and along the way we take the dot product between kernel and input image by multiplying the corresponding values of them and sum up all values to generate one scaler value in the output feature map. This process continues until the kernel can no longer slide further (Ghosh, et. al, 2020).

As Ghosh, et. al, 2020 clearly stated as shown in Fig. 5, where the 2 × 2 kernel (shown in light blue color) is multiplied with the same sized region (shown in yellow color) within the 4 × 4 input image and the resulting values are summed up to obtain a corresponding entry (shown in deep blue) in the output feature map at each convolution step.

After performing the complete convolution operation, the final output feature map, the convolution operation with no padding to the input image and with stride (i.e., the taken step size along the horizontal or vertical position) of 1 to the kernel is shown in Fig.5.

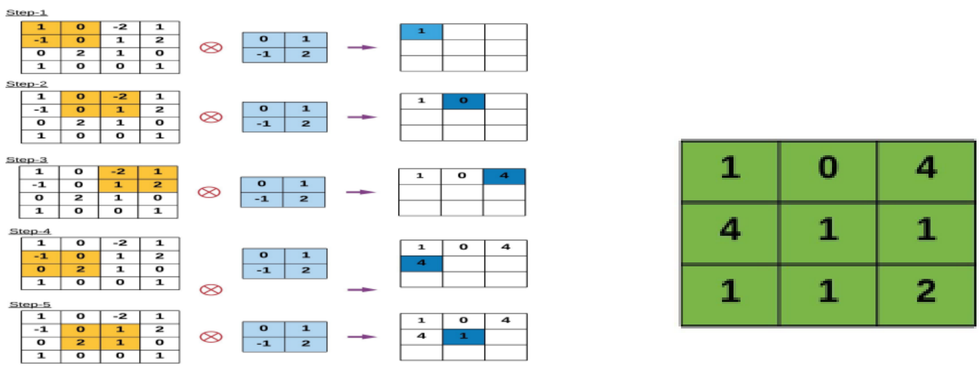


Figure 5: A. the first 5 steps of convolution operation B. The final feature maps

The padding is important to give border size information of the input image more importance, otherwise without using any padding the border side features are gets washed away too quickly. The padding is also used to increase the input image size; as a result, the output feature map size also gets increased.

Pooling layer

The pooling layers are used to sub-sample the feature maps (produced after convolution operations), i.e., it takes the larger size feature maps and shrinks them to lower sized feature maps. (Ghosh, et. al, 2020. While shrinking the feature maps it always preserves the most dominant features (or information) in each pool steps. The pooling operation is performed by specifying the pooled region size and the stride of the operation, similar to convolution operation. There are different types of pooling techniques are used in different pooling layers such as max pooling, min pooling, average pooling, gated pooling, tree pooling, etc. Max Pooling is the most popular and mostly used pooling technique (Ghosh, et. al, 2020).

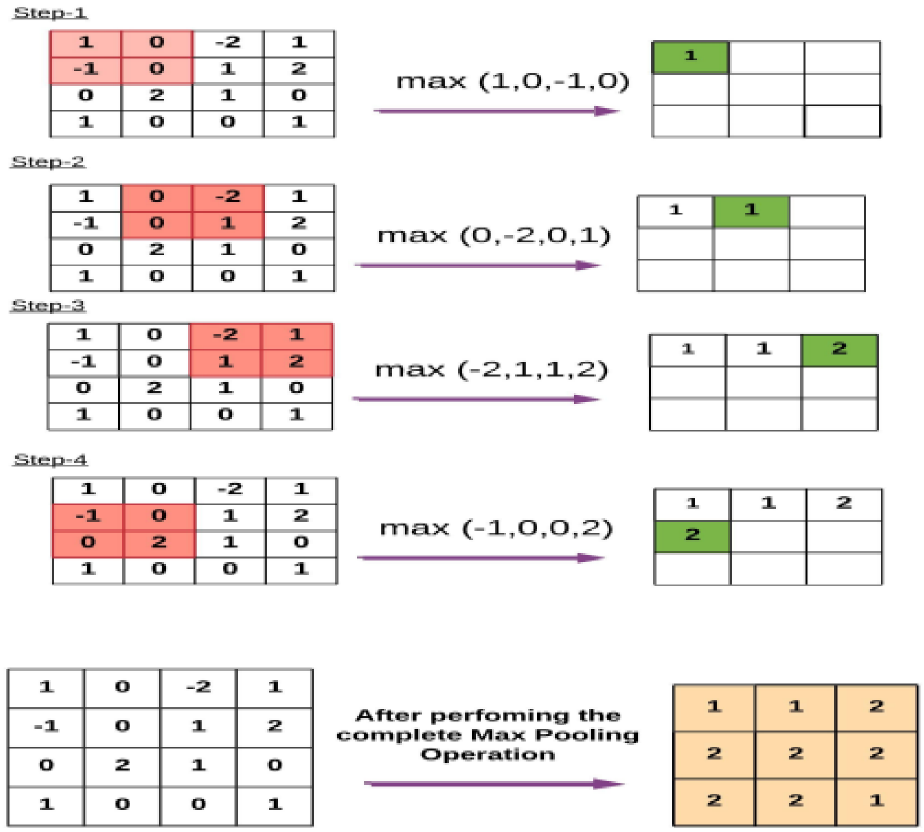


Figure 6: Initial steps as well as the final output of max-pooling operation

Where the size of the pooling region is 2 × 2 (shown in orange color, in the input feature map) and the stride is 1 and the corresponding computed value in the output feature map (shown in green) The formula to find the output feature map size after pooling operation as below:

$$h' = \left[\frac{h-f}{s}\right] \qquad (2.1)$$

$$w' = \left[\frac{w-f}{s}\right] \qquad (2.2)$$

Where h' denotes the height of the output feature map, w' denotes the width of the output feature map, *h* denotes the height of the input feature map, *w* denotes the width of the input feature map, f is the pooling region size and s denotes the stride of the pooling operation. (Ghosh, et. al, 2020).

Activation Functions (Non-Linearity)

The main task of any activation function in any neural network-based model is to map the input to the output, where the input value is obtained by calculating the weighted sum of neuron's 10 input and further adding bias with it (if there is a bias). (Ghosh, et. al, 2020) clearly stated that the activation function decides whether a neuron will fire or not for a given input by producing the corresponding output.

The non-linearity behavior of those layers enables the CNN model to learn more complex things and manage to map the inputs to outputs non-linearly. The important feature of an activation function is that it should be differentiable in order to enable error backpropagation to train the model. The most commonly used activation functions in deep neural networks (including CNN) are:

Sigmoid activation function takes real numbers as its input and bind the output in the range of [0,1]. The curve of the sigmoid function is of 'S' shaped (Ghosh, et. al, 2020). However, it has a vanishing gradient problem(Nwankpa et al.,2018). the mathematical equation of sigmoid is:

$$f(X)sigm = \frac{1}{1+e^{-x}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.2.3$$

The Tanh (hyperbolic tangent) activation function is another nonlinear activation function and used to bind the input values (real numbers) within the range of [-1, 1]

(Ghosh, et. al, 2020). it has a vanishing gradient problem(Nwankpa et al., 2018).The mathematical equation of Tanh is:

$$f(X)tanh = \frac{e^x - e^{-x}}{e^x + e^{-x}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ 2.4$$

The Rectifier Linear Unit (ReLU) is the most commonly used activation function in Convolutional Neural Networks. It is used to convert all the input values to positive numbers. The advantage of ReLU is that it requires very minimal computation load compared to others (Ghosh, et al., 2020). Because ReLU prevents vanishing gradient problem and whose range is either 0 or greater than 0(Nwankpa et al., 2018). The mathematical representation of ReLU is:

$$f(x) = max\ (0, x)\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.2.5$$

Fully Connected (FC) Layer is the last part (or layers) of every CNN architecture (used for classification) is consist of fully-connected layers, where each neuron inside a layer is connected with each neuron from its previous layer. The last layer of Fully-Connected layers is used as the output layer (classifier) of the CNN architecture (Ghosh, et al., 2020). The Fully-Connected Layers are type of feed-forward artificial neural network (ANN) and it follows the principle of traditional multi-layer perceptron neural network (MLP). The FC layers take input from the final convolutional or pooling layer, which is in the form of a set of metrics (feature maps) and those metrics are flattened to create a vector and this vector is then fed into the FC layer to generate the final output of CNN. (Ghosh, et al., 2020).

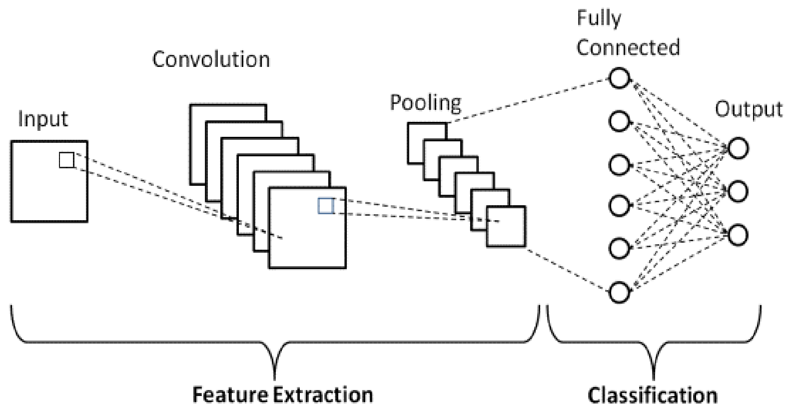


Figure 7: The architecture of Fully Connected Layers

In addition to activation layer, there is different optimizer techniques were developed by different authors. According to Yuan et al., 2021, there are two optimizer methods such as adaptive optimization methods (such as AdaGrad, RMSProp, Adam, and RAdam) and non-adaptive optimization methods (such as SGD and SGD with momentum) have recently been used in deep learning. The former has the characteristics of fast convergence speed but low convergence accuracy. The convergence speed of the latter is relatively slow, but the convergence accuracy is high. In this study we used ADAM and SGDAM optimizer methods are considered.

CNN Architectures

LeNet-5

The LeNet-5 is one of the earliest CNN architectures, which was designed for classifying the handwritten digits. It was introduced by LeCun et al. in 1998. The LeNet-5 has 5 weighted (trainable) layers, that is, three convolutional layer and two FC layers. Among them, each of first two convolution layer is followed by a max-pooling layer (to sub-sample the feature maps) and afterward, the last convolution layer is followed by two fully connected layers. The last layer of those fully connected layers is used as the classifier, which can classify 10 digits. A sigmoid function was used to include nonlinearity before applying the pooling operation.

AlexNet

The AlexNet model was proposed by (Krizhevesky et al, 2012) which is designed to classify Image Net data and enhances the learning capacity by creating deeper layers (a total of 11 layers) and applying different optimization parameters(Krishna & Kalluri, 2019). It comprises five convolution layer, three pooling layer, and three fully connected layers followed by the output classification layer. The input layer is the first layer that defines the input dimensions. The AlexNet used rectified linear unit (ReLU) non-linearity activation function after each convolutional and fully connected layer. (Ghosh, et al, 2020)

VGGNet

VGG is another CNN architecture, the fruitful use of deep CNNs in image detection problem has accelerated the CNN architectural design research works. Due to this, the authors Simonyan et al. proposed an effective CNN architecture design called Visual Geometry Group (VGGNet). Which is introduced by Simonyan and Zisserman in 2014. The authors introduced a total of 6 different CNN configurations, among them the VGGNet-16 (configuration D) and VGGNet-19 (configuration E) are the most successful ones. The main contribution of VGGNet is that it shows that the depth of a network is a critical component to achieve better detection or classification accuracy in CNNs. The

VGG architecture consists of some consecutive two convolutional layers with the ReLU activation function followed by a single max-pooling layer. It also used rectified linear unit (ReLU) non-linearity activation function after each convolutional and fully connected layer. The final output layer of the model is with a Softmax layer for classification. It is very deeper network compare to Alex Net (Ghosh, et al., 2020)

Google Net

The google Net architecture is CNN architecture; it uses network branches instead of using single line sequential architecture. The google Net was proposed by Szegedy et al. in 2014. The google Net has 22 weighted (learnable) layers; it used "Inception Module" as the basic building block of the network. The processing of this module happens in parallel in the network, and each (a simple basic) module consist of 1 × 1, 3 × 3 and 5 × 5 filtered convolution layers in parallel and then it combines their output feature maps, that can result in very high-dimensional feature output. Although the Google Net has 22 layers, but it has 12 times lesser parameters than AlexNet(Ghosh, et al, 2020). It has auxiliary classifiers, that is use to combat vanishing gradient problem. It also used rectified linear unit (ReLU) non-linearity activation function. It used an average pooling layer instead of the fully connected layers. The GoogLeNet used SGD learning algorithm with a fixed learning rate and with 0.9 as momentum factor. The Google Net was the winner of ILSVRC-2014 (Ghosh, et al., 2020).

Res Net

The Res Net's architecture uses residual mapping (H(x) = F(x) + x) instead of learning a direct mapping (H(x) = F(x)) and these blocks are called residual bocks. The complete

Res Net architecture is consisted of many residual bocks with 3 × 3 convolution layers. Although the ResNet (with 152 Layer) is 8 times deeper than VGGNets (22 layers), it has complexity lower than VGGNets (16/19) (Ghosh, et al, 2020). The ResNet used SGD learning algorithm with the min-batch size of 128, weight decay of 0.0001 and momentum factor of 0.9.

2.4.5. Classification

Classification is a procedure to classify images into several categories, based on their similarities. Classification algorithms play a major role in image processing techniques. It is used to classify the features that are extracted from the image into various classes based on different characteristics (Manjula, et al., 2017). The author Manjula, et al. ,2017 stated that there are several classification algorithms available for real world applications such as Naive- Bayes, Support Vector Machine, k-Nearest Neighbour algorithms, Shadow algorithms, Minimum Mean Distance (MMD), Neural Networks, Decision Trees, Hidden Markov Model, K-means clustering algorithms, Machine Learning are described as follows.

K - Nearest Neighbor Algorithm (KNN): KNN is one of the simplest statistical classification algorithms among various machine learning algorithms. The objects are clustered with a maximum number of votes given by the neighbors. The objects are assigned with most similar neighbors in the feature space (Manjula, et al., 2017)

Naive-Bayes Classifier: is a kind of machine learning technique which is based on the Bayes theorem with strong inter-relationship among features. Naive-Bayes is one type of classifier that models around detailed features of the problem. Each feature is independent of the other features. It is trained by the number of classification rules such as maximum posteriori to determine the Bayesian statistics (Manjula, et al., 2017)

Support Machine Vector (SVM): The classification and analysis of the data sets are frequently performed by the SVM Classifier. It is also one of the supervised learning models that are non-probabilistic and binary non-parametric classifier. SVMs are learning systems that use a hypothesis space of linear functions in a hyper space. SVM is trained with a Learning algorithm from optimization theory that implements a learning bias derived from statistical learning theory. The aim of Classification via SVM is to find a 'computationally efficient' way of learning good separating hyper planes in a hyperspace, where 'good' hyper planes mean ones optimizing the generalizing bounds and by 'computationally efficient' we mean algorithms able to deal with sample sizes of very high. (Maneela & Pushpendra, 2013). A kernel function is used to define the higher dimensional space. In SVM, different kernel function and parameter selection has important implications for algorithm. There are four categories of kernel functions in researches and applications.

Linear kernel functions The SVM in the function is the hyper plane in sample space.

$$K(x, x') = xx' \quad \text{……………………………...………………2.6}$$

Polynomial kernel functions The SVM in the function is q-order polynomial classifier.

$$K(x, x_i) = [(xx_i) +]^q \quad \text{…………………………………………2.7}$$

RBF kernel function each general function Centre corresponds to a support vector. Output power value is automatically determined by the algorithm.

$$K(x, x_i) = exp\left[-\frac{|x-x_i|}{2\alpha^2}\ \right]^2 \text{..}2.8$$

Sigmoid kernel functions The SVM is multilayer perceptron containing a hidden layer. Hidden nodes are determined automatically by the algorithm**.**

$$K(x, x_i) = tanh(v(x, x_i) + c)\text{...................................}2.9$$

Artificial Neural Network is a kind of artificial intelligence that controls human mind's function. It is a non-parametric approach. Non-parametric approach has no assumption about the data and where correctness depends on the number of inputs and network. ANN learns from the environment, and stores the experiential knowledge. ANN is a collection of layers basically it has 2 layers i.e., input and output, but some system has hidden layers. Every layer has number of neurons. They connected with each other by a weighted link. (Maneela & Pushpendra, 2013).

Deep Learning is a class of machine learning techniques, where many layers of information processing stages in hierarchical supervised architectures are exploited for unsupervised feature learning and for pattern analysis/classification. The family of deep learning methods has been growing increasingly richer, encompassing those of neural networks, hierarchical probabilistic models, and a variety of unsupervised and supervised feature learning algorithms (Li Deng & Dong Yu, 2013). According to Li Deng & Dong Yu, Deep learning classification techniques are Deep belief network, Boltzmann machine, Restricted Boltzmann machine, deep neural network and Convolutional Neural Network described in details as follows.

Deep belief network (DBN): probabilistic generative models composed of multiple layers of stochastic, hidden variables. The top two layers have undirected, symmetric connections between them. The lower layers receive top-down, directed connections from the layer above. (Li Deng & Dong Yu, 2013).

Boltzmann machine (BM): a network of symmetrically connected, neuron-like units that make stochastic decisions about whether to be on or off. (Li Deng & Dong Yu,2013)

Restricted Boltzmann machine (RBM): a special type of BM consisting of a layer of visible units and a layer of hidden units with no visible-visible or hidden-hidden connections.

Deep neural network (DNN): a multilayer perceptron with many hidden layers, whose weights are fully connected and are often (although not always) initialized using either an unsupervised or a supervised pretraining technique. (Li Deng & Dong Yu,2013).

Convolutional Neural Network (CNN) is a well-known deep learning architecture inspired by the natural visual perception mechanism of the living creatures.

Yamashita, *et al.*, 2018 clearly stated that CNN is a type of deep learning model for processing data that has a grid pattern, such as images, which is inspired by the organization of animal visual cortex and designed to automatically and adaptively learn spatial hierarchies of features, from low- to high-level pattern. According to (Yamashita, *et al.*, 2018) CNN is a mathematical construct that is typically composed of three types of layers (or building blocks): convolution, pooling, and fully connected layers. The first two, convolution and pooling layers, perform feature extraction, whereas the third, a fully connected layer, maps the extracted features into final output, such as classification.

The advantage of CNNs over others classification algorithms (SVM, K-NN, ANN, and others) is that the CNNs learns the best features to represent the objects in the images and has a high generalization capacity, being able to precisely classify (Yamashita, *et al.*, 2018).

There are several differences to note between manual feature extraction and CNN. First, CNN does not require hand-crafted feature extraction. Second, CNN architectures do not necessarily require segmentation of tumors or organs by human experts. Third, CNN is far more data hungry because of its millions of learnable parameters to estimate, and, thus, is more computationally expensive.

2.5. Related works

The applications of imaging technology in different sectors have been studied and replacements of human with intelligent systems are carried out. Several researchers developed their model for diagnosis, detection, and classification of human skin disease. Different methods have been applied by many researchers in order to detect and classify human skin diseases from images. In this section we review of some approaches and we try to shows the gap we want to fill.

Leelavathy et, al, have proposed a Human Skin disease detection using computer vision and machine learning technique. The Authors used 2D Wavelet Transform algorithm for feature extraction and CNN for classification. Their method recognized three types of diseases such as Psoriasis, Lichen Planus, Pityriasis Rosea. The authors didn’t consider other fungal skin disease like tinea corporis, tinea pedis, tinea capitis and Healthy skin.

The Authors (Li-sheng, et al, 2018), presented Skin Disease Recognition Method Based on Image Color and Texture Features. The Authors used GLCM for feature extraction and SVM for classification. Their method recognized three types of diseases such as herpes, dermatitis and psoriasis. Their evaluation result is 85%, 90%, 95% of the recognition rate respectively. However, the authors have used traditional textures feature extraction method. The textures were traditional features which are less effective to generalize and distinguish important features between highly similar classes when used alone and not considered CNN. The authors didn't consider other fungal skin disease like tinea corporis, tinea pedis, tinea capitis and Healthy skin.

As referenced by (M. Kumar & R. Kumar, 2016) they have proposed an intelligent system human skin disease identification using the KNN algorithm. Their method recognized two human skin diseases such as acne and psoriasis. However, this model is not effective because the processing time and computational cost become high; KNN is a distance metric since it is a lazy learner. And also, they didn't consider other fungal skin disease like tinea corporis, tinea pedis, tinea capitis and Healthy skin.

As Referenced in (Carl et al, 2015) has been studied skin disease detection using Case-based Reasoning (CBR) with Image Processing techniques. They used the ABCD rule for feature extraction and Multilayer Perceptron Classifier (MLP) for classification. The detection rate of their method is 88%, 61%, 75%, 51%, 43%, and 34% for Eczema, Psoriasis, Acne, Skin Cancer, Scabies and Seborrhea Dermatitis respectively. From their result, only Eczema and Acne perform high accuracy. Though, the rest 4 types of diseases less perform than the two skin diseases.

In the work of (S.Malliga et al, 2020) has proposed Skin Disease Detection and Classification using Deep Learning Algorithms. The author used CNN both as feature extraction and classification. Their method detected 3 skin diseases such as Melanoma, Nevus, Sebborheic Keratosis classified. The authors using CNN algorithms, 70% accuracy is achieved. But the authors didn't consider other fungal skin disease like Tinea capitis, Tinea Corporis and Tinea Pedis and healthy skin.

The authors (Ubale &Paikrao, 2019) presented Skin Diseases Detection and Classification method using Different Color Phase Models. They have used KNN for classification and two-color phase models for feature extraction. The accuracy of this method is 81.6%. However, the authors have used traditional color feature extraction method. The colors were traditional features which are less effective to generalize and distinguish important features between highly similar classes when used alone and KNN very limited efficiency. And also, they didn't consider other fungal skin disease like tinea corporis, tinea pedis, tinea capitis and Healthy skin.

In the work of (Sumithra, et al, 2015) have been studied Segmentation and classification of skin lesions for skin diagnosis. They used SVM and KNN for classification purposes. Their method Recognized 5 different skin diseases such as, Melanoma, bullae, seborrheic keratosis, shingles and squamous cell. They compare the two algorithms by using individually and as fusion. The result of the proposed method is 46.71% and 34% of F-measure using SVM and KNN classifier respectively and with 61% of F-measure for a fusion of SVM and KNN. However, the textures are traditional features that are less effective to generalize and distinguish important features between highly similar classes when used alone and not considered CNN. They didn’t consider other fungal skin disease like tinea corporis, tinea pedis, tinea capitis and Healthy skin.

In the study of (JessicaVelasco, et al., 2019), have been studied A Smartphone-Based Skin Disease Classification Using MobileNet CNN. Their system recognized 7 types of skin disease such as; Acne, Eczema, Pityriasis rosea, Psoriasis, Tinea Corporis, Varicella (chickenpox), and Vitiligo. They used oversampling techniques and data augmentation to increase the performance of the model. They have got 94% accuracy. But they didn’t consider other fungal skin disease like tinea pedis, tinea capitis and Healthy skin.

2.6. Summary

In this chapter, we explained the general science of Human skin structure and different techniques of digital image processing. And also, we have presented the most recent related works for automating skin disease detection systems to show and explain the performance of these works. Additionally, we have tried to show how the gaps of these related works have to been filled out. In this thesis work, we have designed a combined of BRISK local feature and CNN feature extraction applied to detect and classify Tinea Capitis, Tinea pedis, Tinea Corporis and Healthy skin. And fill the gaps in the previous related works.

Table 1: related work summary

No	Author	Title	Techniques	Disease types	Result	Limitation
1	Li-sheng, et	Skin Disease Recognition	GLCM for feature	herpes, dermatit	90%	The authors have used traditional textures feature

	al .2018	Method Based on Image Color and Texture Features	extraction and SVM for classification	is and psoriasis		extraction method. The textures were traditional features which are less effective to generalize and distinguish important features between highly similar classes when used alone and not considered CNN
2	M. Kumar & R. Kumar, 2016	AN INTELLIGENT SYSTEM TO DIAGNOSIS THE SKIN DISEASE	Skin image detection, image processing, and image recognition and image classification using KNN	acne and psoriasis	58%	This model is not effective because the processing time and computational cost become high; KNN is a distance metric since it is a lazy learner.
3	Ubale &Paikrao, 2019	Detection and Classification of Skin Diseases using Different	Color phase as feature extraction and KNN as classifier	Acne,Melanoma,Mycosis,Papillomas,psoriasis,Viti	81.6%	They used traditional feature extraction methods colors. The colors have been traditional features that were less efficient to generalize and distinguish important features between

		Color Phase Models		lligo and wart		highly similar classes.
4	Sumithra, et al, 2015	Segmentation and Classification of Skin Lesions for Disease Diagnosis	color and texture features as feature extraction and SVM&KNN as classifier	Melanoma, Bullae, Seborrheic keratosis, Shingles, Squamous	61%	The Authors have used traditional classification methods. The textures were traditional features that are less effective to generalize and distinguish important features between highly similar classes when used alone and not considered CNN
5	S.Malliga et al, 2020	Skin Disease Detection and Classification using Deep Learning Algorithms	CNN as feature extraction and Classification	Melanoma, Nevus, Sebborheic Keratosis	70%	Using CNN as a classifier increased the computational time of the model. Even though increasing the dataset and the number of hidden layers in the CNN model improves the performance of the CNN, significantly increases the training time of the CNN.

CHAPTER THREE

3. METHODOLOGY (MATERIALS AND METHODS)

3.1. Overview

In this chapter, we briefly described the proposed Human skin disease Detection and classification model. The model passes through different steps; such as image acquisition, it passes through preprocessing, image segmentation, feature learning and classification of Human skin disease to identify the given image based on the disease type. And also General descriptions of the proposed Human skin disease architecture, which methods and materials are used to achieve the general objectives of this thesis work are described in detail.

3.2. Human skin Disease Detection and classification model Architecture

The proposed model architecture of Human skin disease identification is depicted in figure 8. The proposed model contains the following major responsibilities of image processing. These are image acquisition, preprocessing, segmentation, feature extraction, and classification. The proposed model has two phases it includes the training phase and testing phase. The training phase uses training dataset collected from diseased skin and healthy skin images and applied to trained model. The testing phase to evaluation of the trained model and uses previously unseen diseased skin and Healthy skin pass through the same steps as in the training phases.

The proposed model starts by acquisition of Human skin disease image dataset for preprocessing. In the preprocessing step: resizing the image size by bi- cubic interpolation, noise removal by median filtering, color enhancement by histogram equalization, image augmentation, Segmentation by global thresholding and Grayscale conversation are applied. Then, feature extraction is next, using CNN, which is used to extract the high-level features and local feature descriptor, which is used to extract low level features of the Human skin disease image. Then, in the training phase, the CNN features and local features are combined together and feed in to the classifier to train the model. Where as in the testing phase, CNN features and local features combined and feed in to the previously trained model. The general description of the proposed model is as shown in fig. 8.

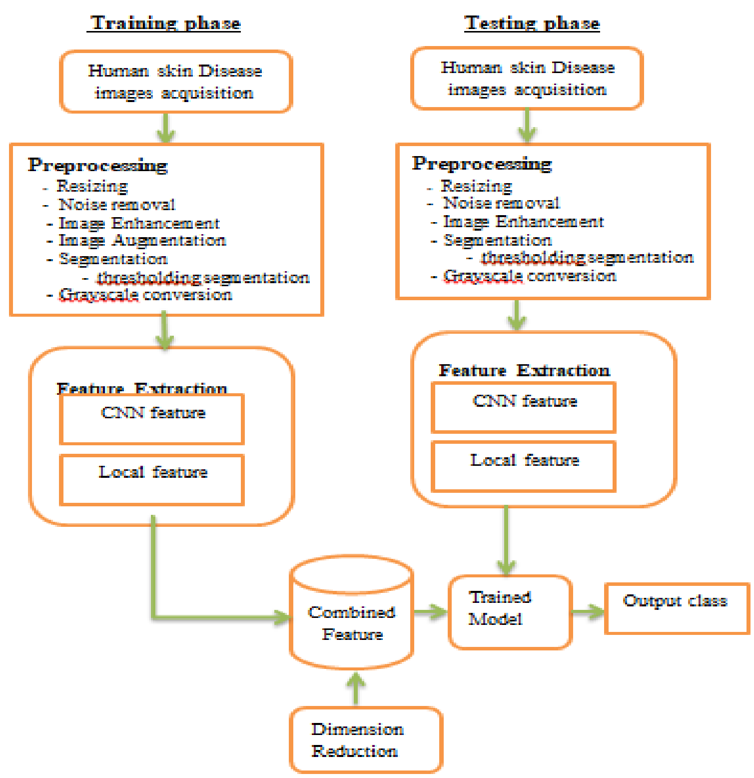


Figure 8: The proposed model architecture

3.2.1. Image Acquisition

The image acquisition setting was cross-sectional methods (Images have taken during patients coming for services to Hospital. So, in this thesis work, we used Techno CAMON X smartphone camera with 14 mega pixels to capturing human skin disease in the form of JPG (Joint Photographic Experts Group) file format to acquired images. The researcher has captured human skin diseases images from March 23, 2021 up to April 30, 2021 in three Hospitals, Bahir Dar City. When we were taking images, the camera was fixed on a stand which reduces the movement of hand and capturing uniform images of skin disease. According to work done by (Abrham, et al, 2016), we acquire image on the distance of 130 mm from the human skin disease. We collected diseased skin (tinea capitis, tinea corporis and tinea pedis) and Healthy skin from Tibebe Giyon Specialized Hospital, Felege Hiywet Comprehensive Specialized Hospital, Gamby General Teaching Hospital in Bahir Dar city and DermNet.com image repositories. The Researcher has an ethical approval Permission to conduct the study from Bahir Dar university college of medicine and health science institutional Review Board Bahir Dar, Ethiopia. Moreover, the researcher clearly presented the objective of images taken from patients and then asked oral assent and consent before the pictures have taken. During the researcher collected the images, the Specialized Doctors of the Hospitals were guide very well in the occasion of identifying and labeling the human skin diseases. So, we have collected a total of 954 original images from four classes and to maximize the collected dataset we used data augmentation techniques.

3.2.2. Image Preprocessing

Image preprocessing is the second step of digital image processing. Image pre-processing is an essential step of detection in order to remove noises and enhance the quality of original image. The main purpose of this step is to improve the quality of image by removing unrelated and surplus parts in the back ground of image. Good selection of preprocessing techniques can greatly improve the accuracy of the system (Adel and Afsaneh, 2014). In this study, we used different preprocessing techniques such as image resizing, histogram equalization (HE), noise removal, segmentation; gray scale conversion and contrast enhancement are used as an image preprocessing technique. The techniques of preprocessing this study are described as follows.

3.2.2.1. Image resizing

The collected images have different in size, which is width is ranging between 200 and 2000 pixels and length 200 to 3000pixels. Image resize is basic step in image preprocess because of the data set must have similar size that is preferable for the proposed model and the large pixel's size consumes too much computational time and cost. We resized the original dataset images into 224X224, 256x256 and 300X300 pixels. When we resize an image from larger to smaller and smaller to larger will happen information loss. So to remove such types of problems we used an interpolation technique to make the resized image contain all information of the original image without information loss. There are many interpolation techniques from those we used bicubic interpolation techniques. To resize the input image to 224 x 224,256x256 and 300X300 bi-cubic interpolation technique is used because according to (Shreyas, 2014), bi-cubic interpolation is smoother and blur is not formed even when the image is interpolated many times.

Step 1: read original image

Step 2: Image = Original image

Step 3: Image = imresize (Image, [224x224, 256x256, and 300x300], Cubic)

Step 4: Image = image to array (Image)

Step 5: Return resized Image

Algorithm 1 :Image resizing algorithm

Figure 9: Tinea capitis A. Original 251x201pix image and B. 224x224 pix resized image

3.2.2.2. Grayscale conversion

RGB is 3 channel image color formats; it must be change in to 1 channel to apply threshold segmentation. Pixel value as 0 signifies black color, 1 signifies the white color and intermediate values are represented by the shades of gray. In this thesis work, after resized the dataset the conversation of RGB image into gray scale image is done using rgb2gray () MATLAB function and it works by forming a weighted sum of R, G and B components as shown below equation 2.1(Enquahone,2017).

$$Gray = 0.2989 * R + 0.5870 * G + 0.1140 * B \quad (2.1).$$

Step 1**:** read Resized image

Step 2: Image = Resized image

Step 3: Image = rgb2gray (Image)

Step 4: Image = image to array (Image)

Step 5: Return gray Image

Algorithm 2: RGB to Gray Conversation algorithm

Figure 10: Tinea capitis A. Original RGB image and B. Gray scale image

3.2.2.3. Contrast enhancement: Histogram equalization

It is a common technique for enhancing the appearance of images. The aim of image enhancement is to improve the interpretability or perception of information in images for human viewers, or to provide `better' input for other automated image processing techniques. In this study Histogram equalization is considered as contrast enhancement; Histogram equalization is a mechanism that is used to distribute the intensity value on the histogram in order to increase areas of lower contrast to higher contrast.

3.2.2.4. Noise removal: median filter

According to (Kaur, 2015) Noise represents unwanted information which deteriorates image quality. Noise is a random variation of image intensity and visible as grains in the image. Noise means, pixels within the picture present different intensity values rather than correct pixel values. In this thesis we used median filter as a noise removal. Median filter is a simple and powerful non-linear filter which is based on order statics, whose response is based on the ranking of pixel values contained in the filter region. It is easy to implement method of smoothing images. Median filtering is done by, first sorting all the pixel values from the surround's neighborhood into numerical order and then replacing the pixel being considered with the middle pixel value (Kaur, 2015).

Step 1 Select a two-dimensional template **T** of size 3 × 3.
Step 2: Assuming that the pixel is being processed on a point with coordinates **(x, y**).
Step 3: Arrange the pixels, within the template, in ascending order.
Step 4: Compute **Tmed** as the median of the pixel values in template **T**.
Step 5: Replace coordinates **(x, y)** by **Tmed**
Step 6: Repeat steps 1 to 5 until all the pixels in the entire image are processed

Algorithm 3: Noise removal using Median Filter

3.2.2.5. Image augmentation

Data augmentation is an effective and universal technique for improving generalization performance of deep neural networks (Ju Xu.et al, 2020). It is an effective technique for the reduction of over fitting, which is a process of modifying the dataset through random transformations, such as flipping, shifting, translation, cropping, random rotating, random zooming and random erasing, so that the model will not see the same inputs during the training iterations (Shorten & Khoshgoftaar, 2019). In this thesis work, Random Rotation, Random Zoom, Horizontal shift and Vertical Shift had been used as a technique for image augmentation.

3.2.3. Segmentation

Image segmentation is a most commonly used technique and analysis to partition an image into multiple regions, often based on the characteristics of the pixels in the image. Image segmentation is the process of separating foreground from background, or clustering regions of pixels based on similarities in color or shape. (Jeevitha et al, 2020). During segmentation, the image usually is subdivided until the objects of interest were isolated from their background. There are many different ways of performing image segmentation among them we used Global thresholding segmentation.

Global Thresholding: is one of the ways to categorize pixels of an image into group of objects. Image thresholding is done by understanding the intensity distribution among image pixels. A distinct intensity value, T, is chosen to act as a testing parameter for thresholding operation (Sirisha, Ravi & Khan, Am., 2017). The thresholding operation simply performs the relational operation on the pixel values; if the test pixel possesses an intensity value greater than the specified threshold value then it is grouped to represent the object else it represents background. Where T is the selected threshold value, f (x, y) test image and g (x, y) are the resultant threshold image. Thresholding operation is effectively represented by equation.

$G\ (x, y) = \{1, iff\ (x, y) \geq T\ otherwise\ 0$ …………………………………… (3.1)

(Jeevitha et al, 2020). Different thresholding techniques proposed by different researchers are OSTU thresholding, P-tile method, Histogram dependent technique, Edge Maximization technique, Mean method, and visual technique.

Figure 11: Tinea capitis A. Original image and B. Threshold segmented

3.2.4. Feature Extraction

Feature extraction is the process of representing a raw image in a reduced form and retrieving meaningful information from an image that is used to identification of images to predefined classes. In this thesis, CNN and BRISK local feature descriptors are considered to use the double advantage of those techniques. Here, CNN extracts the representing feature from the enhanced RGB image. And also, local feature descriptors are extract features from contrast enhanced or gray scale images.

3.2.4.1. CNN for feature extraction

Convolutional neural network is a deep learning method that can be used for image feature extraction and image classification process. Feature extraction contains a number of convolution layers followed by an activation function and pooling layer. The classifier usually consists of fully connected layers (Lorentzon, 2017). In feature extraction, CNN provides global features called texture, color and shape that extracts automatically. In CNN there are different layers that are combined with each other to perform the specified task. From those layers the first layer is convolutional layer, it contains of a series of filters or learnable kernels which aim at extracting global features from the input, and each kernel is used to calculate a feature map. The first convolutional layer extracts low-level meaningful features such as edges, corners, textures and lines. Second convolutional layer extracts higher-level features, but the highest-level features are extracted in the last convolution layer. The input to the first convolution layer is segmented 224 x 224 x 3 image. We have used 3 convolution layers. The convolution operation requires four parameters, such as number of filters that are used to control the depth of the output volume. In our proposed model, we used 3,8,16 and 32 filters; we go down to the fully connected layer, the numbers of features are to be extracted (filters) are increased. The second parameter is the receptive field size (kernel). Kernel size refers to the size of the filter, which revolves around the feature map. When we overlay the convolution kernel on top of the input image, we can compute the product between the numbers at the same location in the kernel and the input, and we get a single number by summing these products together. We have used 3 x 3 kernel size. The third parameter is stride size, which controls how the filter convolves around the feature map. It is the rate at which the kernel passes over the input image. We have used stride sizes of one. The last parameter is padding that gives the option to make input data wider. We have to use "zero" padding; hence the size of the output is equal to the size of the input.

Pooling layers: are used to sub-sample the feature maps (produced after convolution operations), i.e., it takes the larger size feature maps and shrinks them to lower sized feature maps. While shrinking the feature maps it always preserves the most dominant features (or information) in each pool steps. The pooling operation is performed by specifying the pooled region size and the stride of the operation, similar to convolution operation. There are different types of pooling techniques are used in different pooling layers such as max pooling, min pooling, average pooling, gated pooling, tree pooling, etc. Max Pooling is the most popular and mostly used pooling technique (Ghosh, et al.,2020). Hence, max pooling is considered for this study as a pooling layer this is due to average pooling method smooth out the image and hence the sharp features may not be identified when this pooling method is used which addresses the maximum output within a rectangular neighborhood. Max pooling outputs only the maximum number in each kernel.

Activation function, are mathematical equations that determine the output of a neural network. The function is attached to each neuron in the network, and determines whether it should be activated or not, based on whether each neuron's input is relevant for the model's prediction. Hyperbolic tangent (TANH) and Rectified linear unit (ReLU) activation function are used as an activation layer in our model. This helps to identify suitable activation function towards to Human skin disease identification in both computational time and accuracy.

Batch Normalization is used to reduce the “internal covariance shift” of the activation layers. The internal covariance shift can be explained as the change in the distribution of activations in each layer. Due to continuous weight updating during training, the “internal covariance shift” can may become very high (it may happen when the training data samples are taken from several different distributions e.g., day-light images and night-vision images) and with this high” internal covariance shift” the model takes more time to converge and training time will increase. To solve this problem Batch Normalization operation is implemented as a layer in CNN architecture. On top of a regularizing effect, batch normalization also helps the convolutional network a resistance to vanishing gradient during training. This important feature makes the proposed model to decrease training time and result in better performance. (Ghosh, et al., 2020). Therefore, in this study, we have employed batch normalization layer in order to achieved better result.

To prevent a model from over fitting dropout is used. The dropout layer is the input elements with a certain probability are inactivating or released out such that the separate neurons are talented to learn the features that are less dependent on its surroundings. We have applied dropout with p = 0.1 (probability of every neuron to be dropped) after final fully connected layer where the final fully connected layer is followed by the classifier.

Fully Connected (FC) Layer is the last layers of every CNN architecture (used for classification) are consisted of fully-connected layers, where each neuron inside a layer is connected with each neuron from its previous layer. The last layer of Fully-Connected layers is used as the output layer (classifier) of the CNN architecture. The FC layers take input from the final convolutional or pooling layer, which is in the form of a set of metrics (feature maps) and those metrics are flattened to create a vector and this vector is then fed into the FC layer to generate the final output of CNN. The final fully connected layer typically has the same number of output nodes as the number of classes. We have used four fully connected layers. each followed by activation function ReLU and TanH an optimizers SGDM and ADAM, and finally learning rate value to compute the final output probabilities for each class before applying to the classifier, the last layer holds four nodes (equal to the number of classes), which are directly applied to the classifier.

3.2.4.2. Local feature descriptor for Feature Extraction

There are different local feature detector and descriptor in digital image processing, the most common local feature detector and descriptor are SIFT, SURF, HOG, BRISK and BRIFF (Nikita, et al, 2016). The local features are aim to detect key points within the image and describe regions around these key points. In this study we consider BRISK as local feature descriptor. Because BRISK is a point-feature detector and descriptor and takes point-pairs of both long segments and short segments, and this supports a measure of scale invariance, since coarse resolution may be mapped better by long segments and fine resolution may be mapped by short segments (Nikita, et al, 2016). BRISK is faster than SIFT and SURF, while using less computational resource.

3.2.4.3. PCA for Dimensional Reduction

Dimensionality reduction (DR) of image features plays an important role in image retrieval and classification tasks. In this study, Principal Component Analysis used as dimensionality reduction techniques. PCA is a technique which uses sophisticated underlying mathematical principles to transform a number of possibly correlated variables into a smaller number of variables called principal components. It is one of the most important results from applied linear algebra. The advantage of PCA is finding the patterns in the data and compressing data by reducing the number of dimensions without loss of information. The mathematical concepts that are used for PCA are Standard Deviation, Variance, Co–variance and Eigenvectors (Shereena & David, 2015).

3.2.5. Image Classification

Image classification is an essential part of image processing, to classify the given input raw image into their predefined class. In image classification the image is classified in terms of class that is already predefined based on training. For this purpose, there are so many classification algorithms as discussed in literature review. So, in our study we selected SoftMax and SVM as an identification purpose of the human skin disease.

3.2.5.1. Soft Max as a Classifier

In Digital image processing, CNN uses as Feature Extraction and Image classification purpose. So, CNN uses Softmax activation function as classifier in output dense layer for classifying the input images based on their features extracted in the final fully-connected layer of the CNN.

3.2.5.2. SVM as a Classifier

SVM is a non-parametric and binary classifier that differentiates and divides the classes by determining the boundaries in feature space and maximizes the margin between the classes. The model works by creating a hyperplane that separates data points from one class from those from the other class, with a margin as high as possible. The margin is the maximal width of the slab parallel to the hyperplane that has no interior data points. The support vectors which give the model its name are the data points closest to the hyperplane and therefore determine the margin.

3.3. Summary

In this chapter, the design of the convolutional neural network for human skin disease detection and classification is discussed thoroughly. The components of the model such as image acquisition, preprocessing, segmentation; feature extraction and classification are described in detail. It includes the procedure for acquiring the images algorithm to identify human skin disease through the image processing application.

CHAPTER FOUR

4. EXPERIMENTAL RESULTS AND DISCUSSION

4.1. Overview

In this chapter, we describe the dataset used in the thesis work, tools used in the implementation part and parameters are described. Detail implementation procedure, dataset preparation and experimental results at the training and test phases are presented. And also, described the effect of image preprocessing, activation function and learning rate and SVM classifier.

4.2. Dataset preparation

We have prepared the dataset for training and testing, we used Techno CAMON X smartphone camera with 14 mega pixels to acquire images of Human skin disease (tinea capitis, tinea corporis and tinea pedis) and Healthy skin collected from Tibebe Giyon Specialized Hospital, Felege Hiywet Comprehensive Specialized Hospital, Gamby General Teaching Hospital in Bahir Dar city and DermNet.com image repositories. During collecting an image, the Specialized Doctors of the Hospitals are guided wonderfully in the occasion of identifying and labeling the Human skin disease images.

Table 2: Number of collected dataset

No	Skin Disease type	Image sources				original No. of collected images	After augmented image
		Felege Hiwot	Tibebe Giyon	Gamby	DermNet		
1	Tinea Capitis	18	60	42	279	**398**	**931**
2	Tinea pedis	6	6	3	342	**355**	**833**
3	Tinea Corporis	6	0	0	99	**105**	**238**
4	Healthy skin	23	56	17	0	**96**	**224**
Total		**53**	**122**	**62**	**720**	**954**	**2,226**

4.3. Experimental Setup

In this work, we have used MATLAB R2019a tools with supported libraries for image processing applications. Experiments are done by Lenovo laptop computer on window 10 operating system with Intel(R) core (TM) i7-7500U CPU @ 2.70GHz 2.90GHz, 8 GB RAM and 1TB Hard disk.

Based on the above experimental setup we have done different experiments using 2,226 datasets images for SoftMax classifier and SVM classifier.

We have used 70% of the dataset for training and 30% of the dataset is for testing for SVM classifier in BRISK feature, CNN feature as well as combined feature. We have used accuracy, confusion matrix and time for measuring the performance of the proposed model.

4.4. Experiments

In this study, we clearly show the experiments of observing the effects of image size (224,256 and 300), the effect of optimizer for CNN, the effects of activation function, 0.01 learning rate, fully connected CNN and SVM (with two kernel function) and finally, the dimension reduction forms using PCA.

4.4.1. Using 224X224 preprocessed image, ADAM optimizer with RELU activation function

The first experiment is conducted using optimizer parameter (adaptive moment estimation) ADAM, was tested during the training phase. As clearly shown in figure 12 below, the model obtains training accuracy of 97.66% and time elapsed of 18min59 sec using RELU activation function.

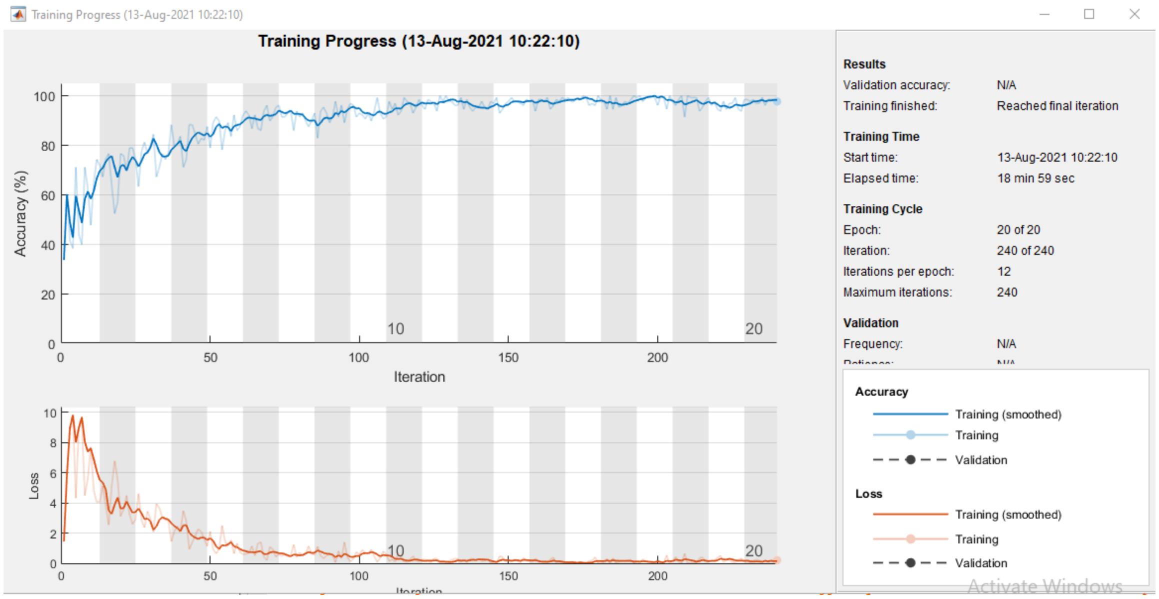


Figure 12: ADAM optimizer with RELU activation function.

Table 3: ADAM optimizer with RELU activation function.

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:04	33.59%	1.4496	0.0100
5	50	00:05:19	78.91%	2.1193	0.0100
9	100	00:09:11	92.97%	0.6045	0.0100
13	150	00:12:58	97.66%	0.1953	0.0100
17	200	00:16:27	96.09%	0.2521	0.0100
20	240	00:18:59	97.66%	0.2107	0.0100

4.4.2. Using 224X224 preprocessed image ADAM optimizer, TANH activation function

In this Experiment, we conduct TANH activation function. As compared to RELU activation function of CNN with the same ADAM optimizer and learning rate value 0.01, TANH activation function has less performance than RELU 94.53% accuracy and elapsed time 18 min 38 sec achieved as shown below figure 13 and table 4.

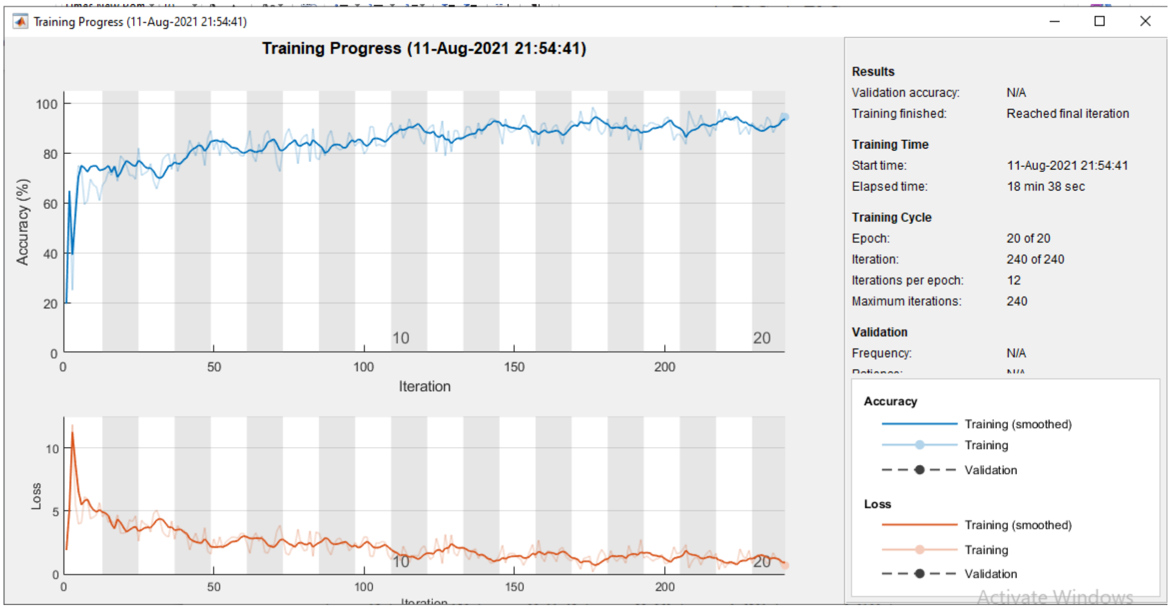


Figure 13: ADAM optimizer with TANH activation function

Table 4: ADAM optimizer with TANH activation function

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:06	19.53%	1.9086	0.0100
5	50	00:03:48	81.25%	2.7859	0.0100
9	100	00:07:41	79.69%	2.9565	0.0100
13	150	00:11:43	89.06%	1.6201	0.0100
17	200	00:15:36	93.75%	0.8418	0.0100
20	240	00:18:38	94.53%	0.6939	0.0100

4.4.3. Using 224X224 preprocessed image, SGDM optimizer, with RELU activation function

In this experiment, we conduct another optimizer parameter which is stochastic gradient descent with momentum (SGDM) optimizer of CNN with RELU activation function as shown in figure 14 and table 5 the model obtains accuracy of 96.88% and time elapsed of 16 min 24 sec. As we can see the training result, the computational time is reduced using SGDM optimizer of CNN as compared to ADAM but the accuracy is decreased by 0.78%.

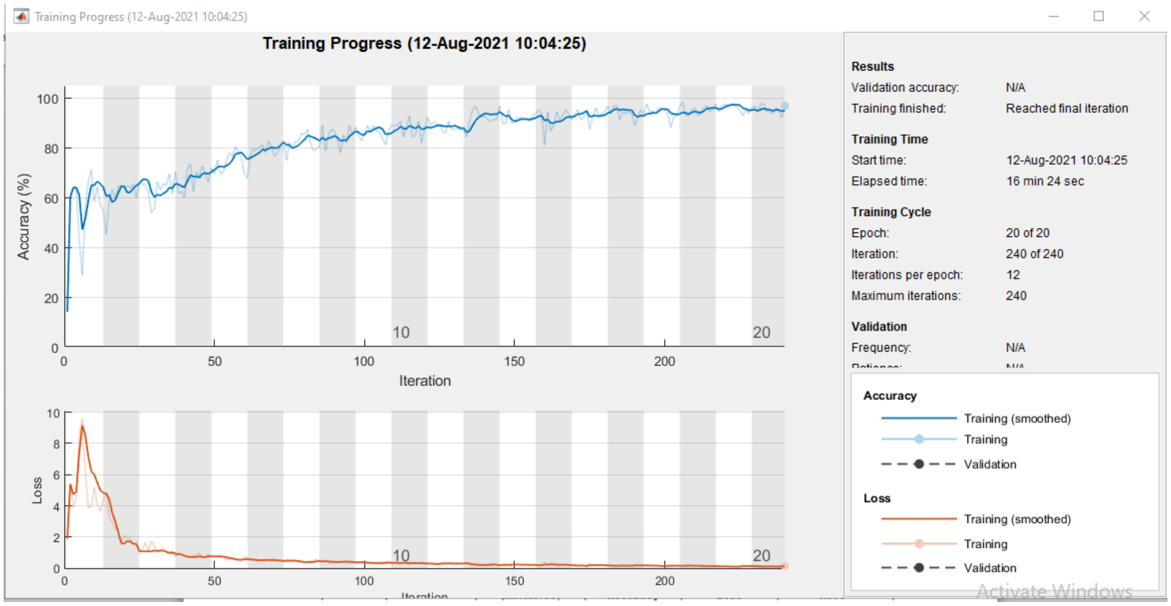


Figure 14: SGDM optimizer with RELU activation

Table 5: SGDM optimizer with RELU activation

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:07	14.06%	1.8811	0.0100
5	50	00:03:43	69.53%	0.7608	0.0100
9	100	00:07:01	85.94%	0.3230	0.0100
13	150	00:10:28	90.63%	0.2804	0.0100
17	200	00:13:43	95.31%	0.1358	0.0100
20	240	00:16:24	96.88%	0.1497	0.0100

4.4.4. Using 224X224 preprocessed image, SGDM optimizer with TANH activation function

In this experiment we conduct stochastic gradient descent with momentum (SGDM) optimizer of CNN with TANH activation function as shown in figure 15 and table 6. the proposed model obtains the accuracy of 92.19% and time elapsed of 19 min 04 sec. As we can see in the training result, the computational time is increase but the accuracy is decreased using TANH activation function of CNN as compared to RELU.

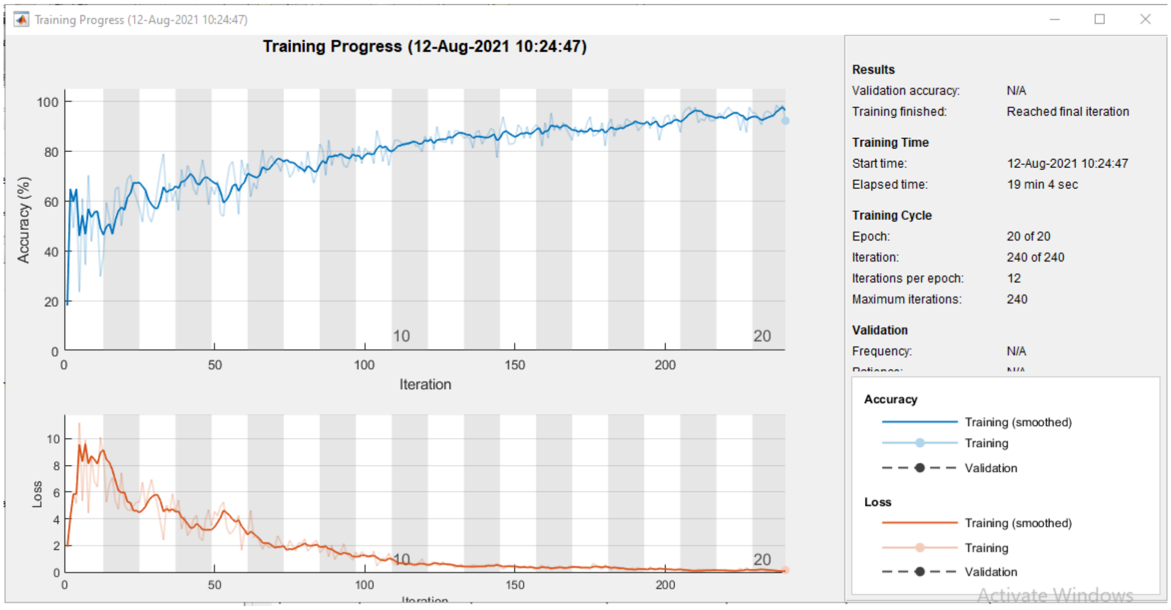


Figure 15: SGDM optimizer with TANH activation

Table 6: SGDM optimizer with TANH activation

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:05	17.97%	1.9042	0.0100
5	50	00:04:10	62.50%	4.5291	0.0100
9	100	00:08:08	81.25%	1.3236	0.0100
13	150	00:12:05	83.59%	0.4731	0.0100
17	200	00:16:00	93.75%	0.2042	0.0100
20	240	00:19:04	92.19%	0.1572	0.0100

4.4.5. Using 300X300 preprocessed image, ADAM optimizer with RELU activation function

In this experiment, after getting the best accuracy of ADAM optimizer with RELU activation function using 224X224 image size of the proposed model. We increased the image size as shown in figure 16 and table 7 the training accuracy became 96.88% with an elapsed time of 27 min 48 sec. As we can see the training result, when we increase the image size with same optimizer, activation function and learning rate. but the computational time is increase and the accuracy is decreased.

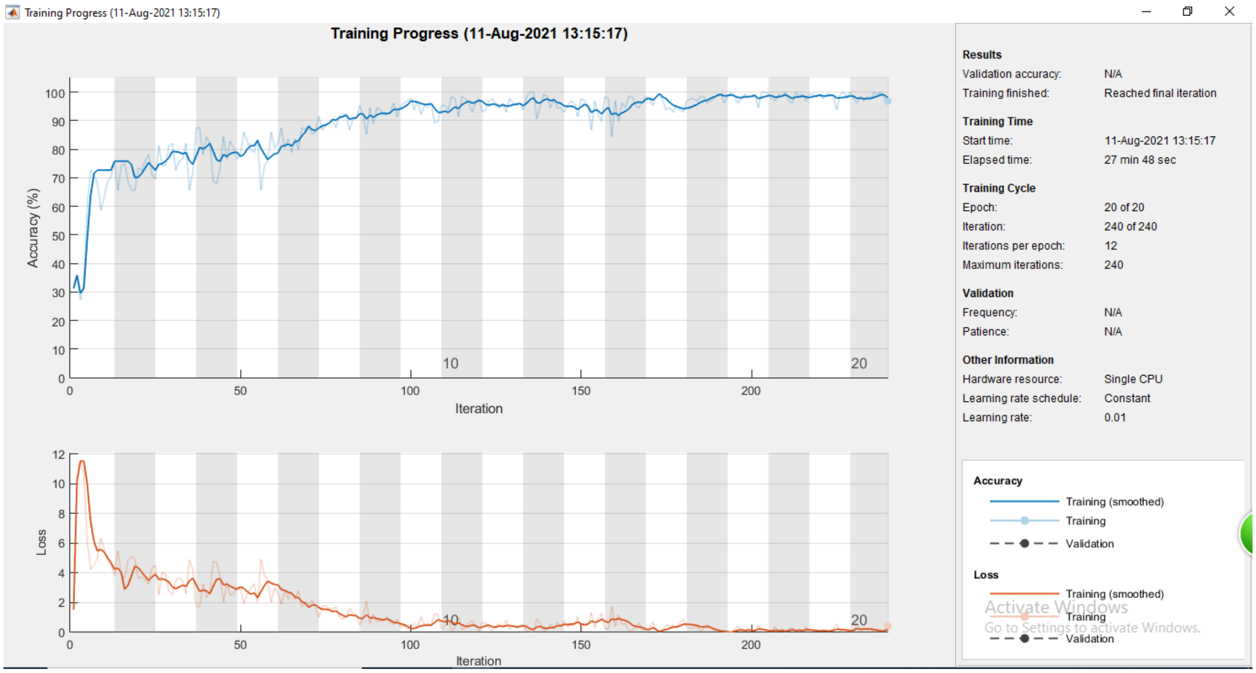


Figure 16: ADAM optimizer RELU activation

Table 7: ADAM optimizer RELU activation

Training on single CPU.
Initializing input data normalization.

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:45	31.25%	1.5302	0.0100
5	50	00:06:36	78.13%	2.9157	0.0100
9	100	00:12:04	97.66%	0.3642	0.0100
13	150	00:17:55	97.66%	0.1891	0.0100
17	200	00:23:17	98.44%	0.1436	0.0100
20	240	00:27:48	96.88%	0.4249	0.0100

4.4.6. Using 300X300preprocessed image, ADAM optimizer with TANH activation function

In this experiment, after getting the accuracy of ADAM optimizer with TANH activation function using 224X224 image size of the proposed model. When we increased the image size as shown in figure 17 and table 8 the training accuracy became 92.91% with an elapsed time of 37 min 2 sec.

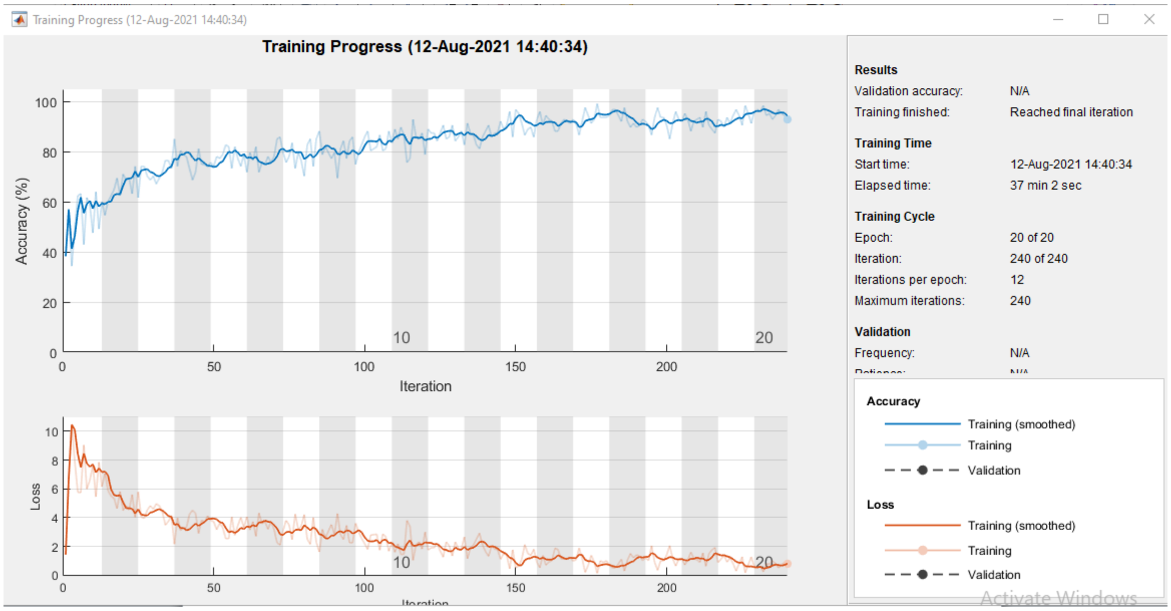


Figure 17: ADAM optimizer TANH activation

Table 8: ADAM optimizer TANH activation

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:10	38.28%	1.4142	0.0100
5	50	00:07:23	74.22%	3.7226	0.0100
9	100	00:14:47	85.94%	2.1852	0.0100
13	150	00:22:28	96.88%	0.4982	0.0100
17	200	00:30:56	90.63%	1.3850	0.0100
20	240	00:37:02	92.97%	0.7848	0.0100

4.4.7. Using 300X300preprocessed image, SGDM optimizer with RELU activation function

In this experiment, after getting the accuracy of SGDM optimizer with RELU activation function using our proposed model. We increased the image size as shown in figure 18 and table 9 the training accuracy became 94.53% with an elapsed time of 44 min 56 sec.

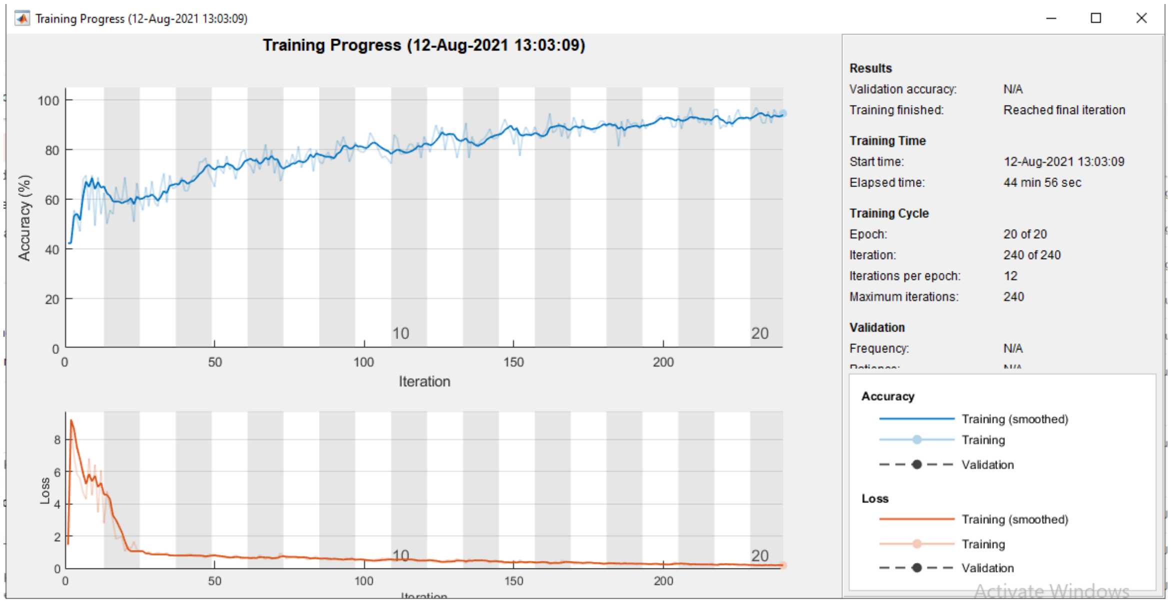


Figure 18: SGDM optimizer RELU activation

Table 9: SGDM optimizer RELU activation

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:14	42.19%	1.4532	0.0100
5	50	00:09:32	64.84%	0.8543	0.0100
9	100	00:18:49	78.91%	0.4730	0.0100
13	150	00:27:05	82.03%	0.4955	0.0100
17	200	00:36:44	92.19%	0.2590	0.0100
20	240	00:44:56	94.53%	0.1832	0.0100

4.4.8. Using 300X300 preprocessed image SGDM optimizer, TANH activation function

In this experiment, after getting the accuracy of SGDM optimizer with TANH activation function using 224X224 of the proposed model. When we increased the image size as shown in figure 19 and table 10 the training accuracy became 88.28% with an elapsed time of 33 min 20 sec.

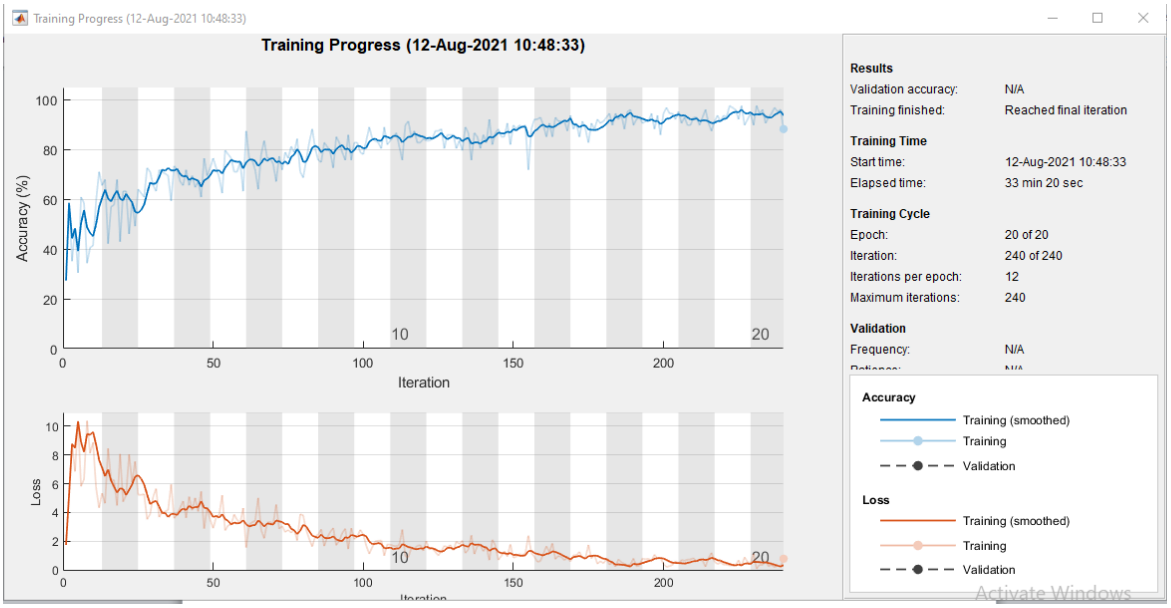


Figure 19: SGDM optimizer TANH activation

Table 10: SGDM optimizer TANH activation

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:08	27.34%	1.7429	0.0100
5	50	00:07:21	76.56%	2.9091	0.0100
9	100	00:14:23	78.13%	2.4717	0.0100
13	150	00:21:16	89.84%	0.8531	0.0100
17	200	00:27:57	94.53%	0.4161	0.0100
20	240	00:33:20	88.28%	0.7867	0.0100

4.4.9. Using 256X256 preprocessed image ADAM optimizer, TANH activation function

In this experiment, after getting the accuracy of ADAM optimizer with TANH activation function using 224X224 of the proposed model. We increased the image size as shown in figure 20 and table 11 the training accuracy became 93.75% with an elapsed time of 25 min 41 sec.

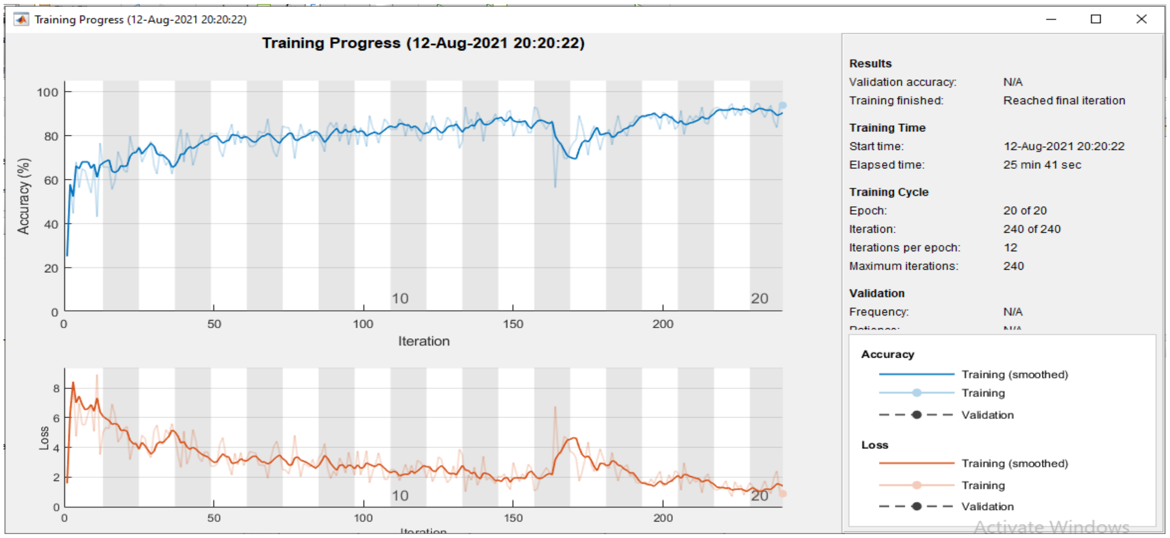


Figure 20: ADAM optimizer TANH activation

Table 11: ADAM optimizer TANH activation

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:11	25.00%	1.5776	0.0100
5	50	00:05:22	80.47%	2.7278	0.0100
9	100	00:10:44	80.47%	2.8431	0.0100
13	150	00:16:04	79.69%	2.9796	0.0100
17	200	00:21:25	90.63%	1.3941	0.0100
20	240	00:25:41	93.75%	0.8798	0.0100

4.4.10. Using 256X256 preprocessed image SGDM optimizer, TANH activation function

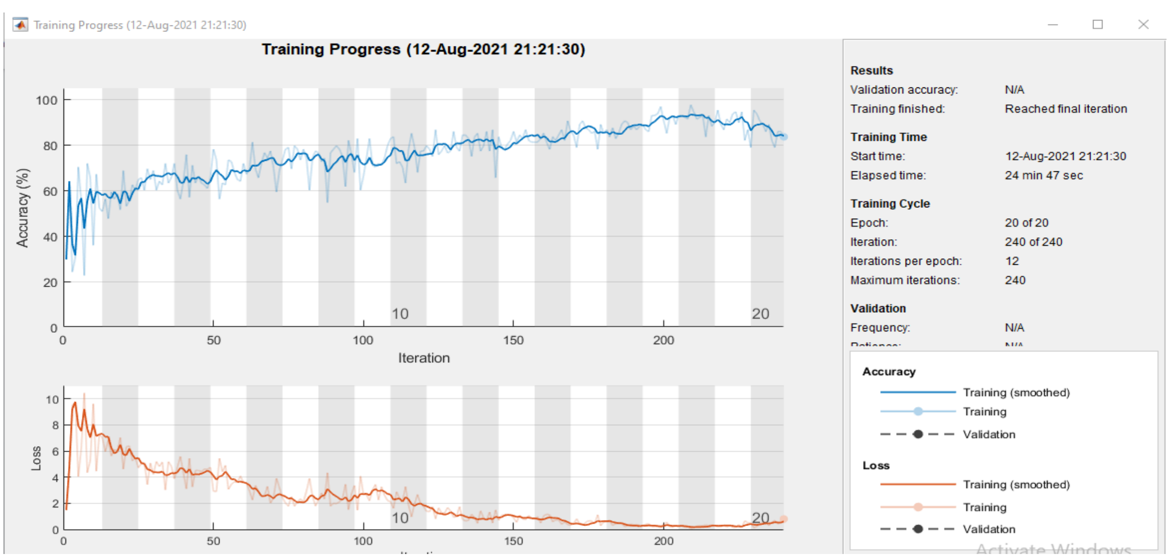


Figure 21: SGDM optimizer TANH activation

Table 12: SGDM optimizer TANH activation

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:07	29.69%	1.4872	0.0100
5	50	00:05:27	78.13%	2.9297	0.0100
9	100	00:10:24	66.41%	2.6710	0.0100
13	150	00:15:27	85.94%	0.8835	0.0100
17	200	00:20:39	90.63%	0.4237	0.0100
20	240	00:24:47	83.59%	0.8045	0.0100

4.4.11. Using 256X256 preprocessed image SGDM optimizer, RELU activation function

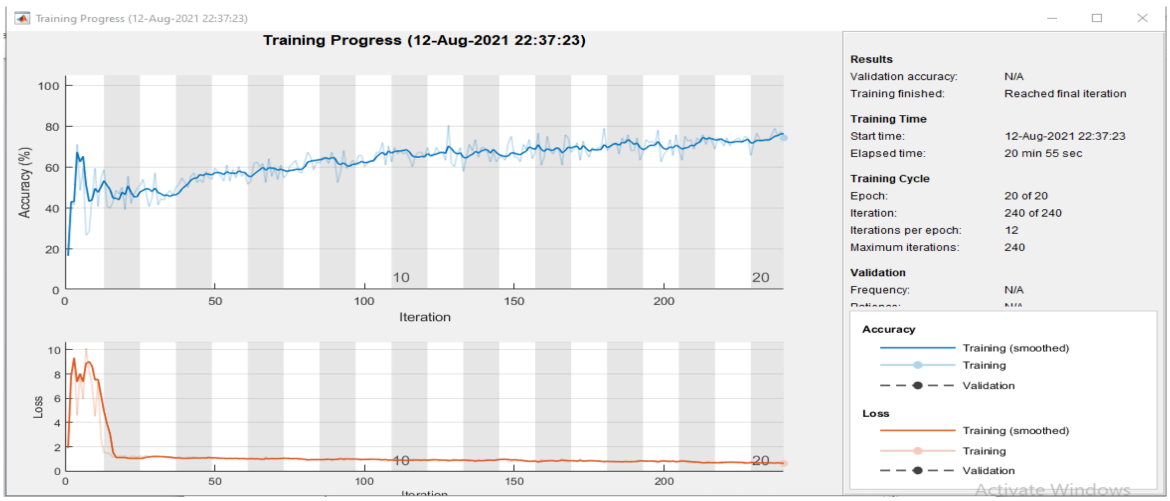


Figure 22: SGDM optimizer RELU activation

Table 13: SGDM optimizer RELU activation

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:06	16.41%	1.8838	0.0100
5	50	00:04:24	59.38%	1.1067	0.0100
9	100	00:08:47	60.16%	0.9926	0.0100
13	150	00:13:07	64.84%	0.9766	0.0100
17	200	00:17:28	66.41%	0.8960	0.0100
20	240	00:20:55	74.22%	0.6199	0.0100

We have been conducting an experiment for three image sizes, two activation function two optimizers and 0.01 learning rate. The training accuracy difference between image size of 224 x 224 pxl, ADAM, RELU, 256 x 256 pxl, ADAM, RELU and 300 x 300 pxl ADAM, RELU was only 0.78. But the elapsed time is much higher in 256X256,300 x

300 pxl image size. So, Adam optimizer, RELU activation function and 224 x 224 pxl image sizes have been used as an input for the remaining experiment CNN model.

4.4.12. Segmentation and BRISK feature

After finding Image size, optimizer, activation function and learning rate for the proposed model, the next task is conducting the effects of segmentation in fully connected CNN. Here we observe that by using segmentation techniques, the elapsed/training time was reduced to 14 min 28sec and the accuracy is 97.66% as shown in figure 23 and table 14 below.

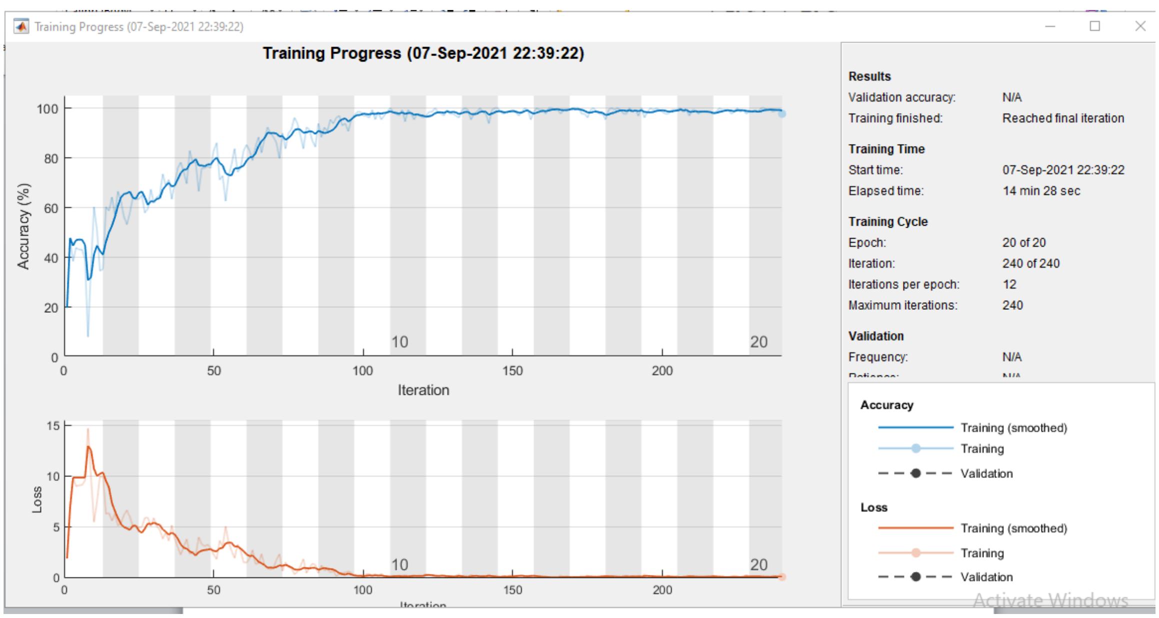


Figure 23: Segmentation and BRISK feature

Table 14: Segmentation and BRISK feature

Initializing input data normalization.

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:04	19.53%	1.8555	0.0100
5	50	00:03:08	82.81%	2.2474	0.0100
9	100	00:06:16	96.88%	0.2555	0.0100
13	150	00:09:11	98.44%	0.1607	0.0100
17	200	00:12:07	98.44%	0.0372	0.0100
20	240	00:14:28	97.66%	0.0339	0.0100

4.4.13. Segmentation and BRISK feature with PCA

In this experiment we conducted the effects of PCA towards the computational time of CNN. Here the training accuracy of the proposed model is increased to 98.44% and training time is decreased from 14 min 28 sec to 12 min 42 sec after applying principal component analysis as shown in figure 24 and table 15 below.

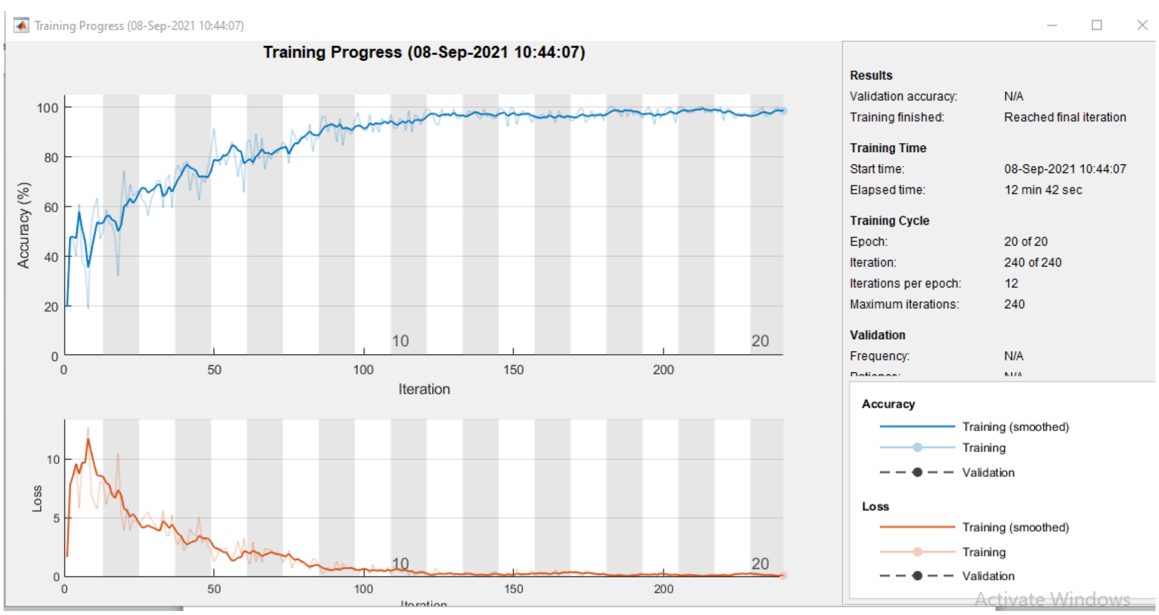


Figure 24: BRISK feature with PCA

Table 15: BRISK feature with PCA

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:04	19.53%	1.6445	0.0100
5	50	00:02:37	91.41%	1.2350	0.0100
9	100	00:05:22	90.63%	0.6987	0.0100
13	150	00:08:01	99.22%	0.0331	0.0100
17	200	00:10:36	97.66%	0.1191	0.0100
20	240	00:12:42	98.44%	0.0537	0.0100

4.5. Model Performance Evaluation

After conducting the training of the proposed model, the second phase is measuring/testing the performance of the proposed model using the testing dataset. Our proposed model contains ADAM optimizer, RELU activation function and learning rate 0.01 because the default learning rate is 0.01. Under this experiment, we have conducted three testing mechanism i.e., fully connected CNN as a classifier and feature extraction, CNN as feature extraction and support vector machine (SVM) as a classifier. Whereas considering SVM we have tested polynomial and Gaussian kernel function of SVM in order to get a better result. Therefore, in fully connected CNN achieved an accuracy of 84.4%. As shown in figure 25 and table 16

Table 16: Confusion Matrix of fully connected CNN

	Healthy skin	Tinea Capitis	Tinea Corporis	Tinea pedis
Healthy skin	55	13	1	6
Tinea Capitis	6	263	4	8
Tinea Corporis	3	2	29	20
Tinea pedis	3	1	37	216

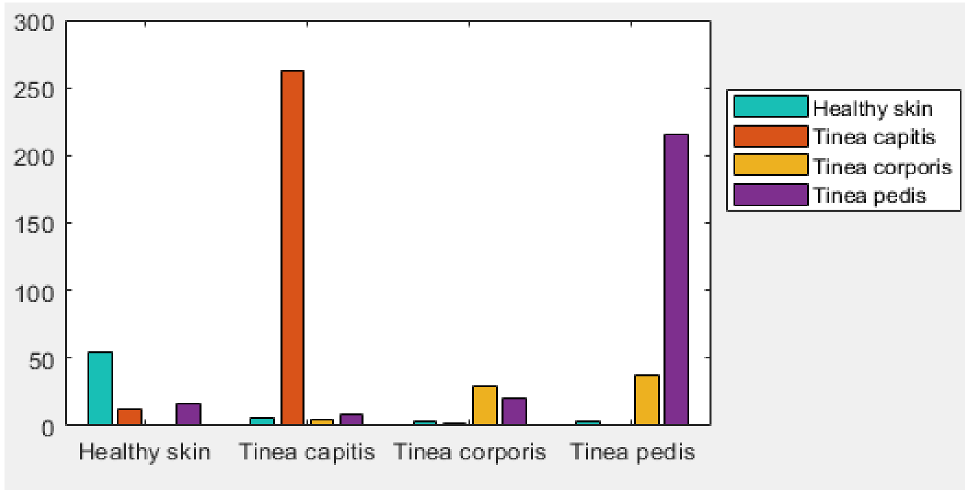


Figure 25: Confusion Matrix of fully connected CNN

The next experiment is replacing Softmax classifier with SVM. Once we get feature vectors using CNN, we can apply classification techniques to classify to their predefined class. In this case, we have selected the multiclass SVM classifier. We have tested Polynomial kernel function of SVM achieved an accuracy of 83.1% as show fig 26 and table 17 below.

Table 17: CNN combined with SVM polynomial

	Healthy skin	Tinea Capitis	Tinea Corporis	Tinea pedis
Healthy skin	48	13	3	3
Tinea Capitis	13	254	0	12
Tinea Corporis	1	6	34	30
Tinea pedis	6	10	16	218

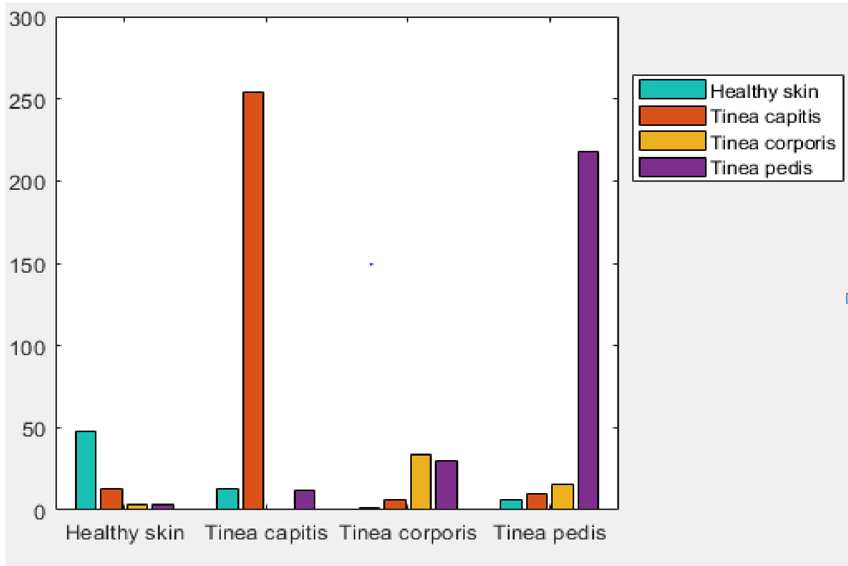


Figure 26: CNN combined with SVM polynomial

To increase the performance of SVM, we have tested another Gaussian kernel function of SVM has the achieved an accuracy of 88.9% as shown in figure 27and table 18 below.

Table 18: CNN combined with SVM Gaussian

	Healthy skin	Tinea Capitis	Tinea Corporis	Tinea pedis
Healthy skin	50	11	2	5
Tinea Capitis	8	258	4	19
Tinea Corporis	1	2	65	3
Tinea pedis	5	10	15	220

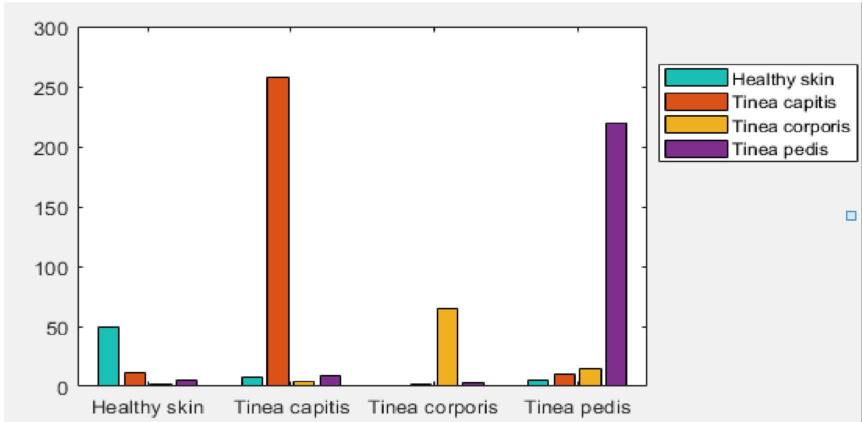


Figure 27: CNN combined with SVM Gaussian

4.6. Comparison of proposed Model with CNN architecture

Comparison is done with CNN architecture using the same dataset and parameter with architectural difference in the model. As we described in the literature review there are different types of CNN architectures such as LeNet, Alex Net, ResNet, VGGNet, (VGG 16 and 19) and Google Net. From those models we have to select Alex Net and to evaluate with the proposed model. The result achieved 85.94% of accuracy at epoch 20 as shown in the figure 28.

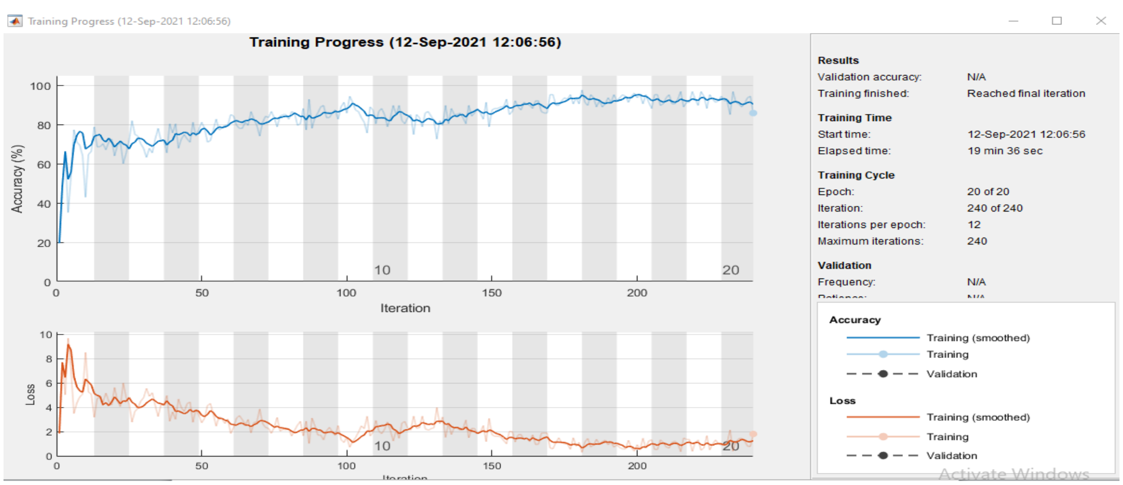


Figure 28: Training result of AlexNet

Table 19: Training result of AlexNet

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:05	19.53%	1.8489	0.0100
5	50	00:04:12	81.25%	2.7329	0.0100
9	100	00:08:09	85.94%	1.5460	0.0100
13	150	00:12:15	87.50%	1.7421	0.0100
17	200	00:16:24	93.75%	0.9251	0.0100
20	240	00:19:36	85.94%	1.8104	0.0100

4.7. Discussion of the Results

Diagnosing, classification and detection of human skin diseases are done by different researchers. In this research work, we develop human skin disease detection and classification model using deep convolutional neural network. The human skin diseases are Tinea capitis,tinea corporis Tinea pedis and Healthy skin collected from Felege Hiwot comprehensive referral Hospital, Tibebe Giyon Specialized Hospital , Gamby General Teaching Hospital in Bahir Dar city and DermNet.com image repositories. We have seen collection of required datasets and labeled by the Doctor, appropriate preprocessing techniques that should be applied to enhance the performance of the training accuracy and the capacity of convolutional neural network for feature extraction in automatic manner and classification. The classifier is separately trained using preprocessed data with Relu and Tanh activation function, ADAM and SGDM optimizer and learning rate 0.01 using different image size. In the first experiment, the model was trained using preprocessed image, 224x224 image size, ADAM optimizer, RELU activation function and default learning rate individually. Then, using preprocessed image, 256x256 image size, ADAM optimizer, RELU activation function and default learning rate individually. Then , using preprocessed image, 300x300 image size, ADAM optimizer, RELU activation function and default learning rate individually. The training accuracy difference between image size of 224 x 224 pxl, ADAM, RELU, 256 x 256 pxl, ADAM, RELU and 300 x 300 pxl ADAM, RELU was only 0.78. But the elapsed time is much higher in 256X256, 300 x 300 pxl image size. So, Adam optimizer, RELU activation function and 224 x 224 pxl image sizes 97.66% achieved training accuracy. Using Global thresholding segmentation techniques we have got 97.66% and elapsed time was decreased. The other thing is enhancing the features of CNN by embedding Binary robust invariant scalable key points (BRISK). Combined feature vector takes more time so to reduce the computational time we have applied principal component analysis to reduce the dimension of the combined feature vector and we have got 98.44% training accuracy. After we extracted the features, the classification model classifies using SVM and SoftMax classifier and we have got 88.9% testing accuracy. Lastly we have to compare proposed model with Alexnet CNN architecture. From the experiment we have got 85.94% training accuracy.

4. 8. Summary

In this chapter, we described required datasets and labeled by the Doctor, appropriate preprocessing techniques that should be applied to enhance the performance of the training accuracy and the capacity of convolutional neural network for feature extraction and classification using ADAM and SGDM optimizer, TANH and RELU activation function and learning rate of 0.01, an experimental evaluation of the proposed model for the identification/classification of Human skin diseases. The experimentation done comparison of the image size, optimizer, activation function and applying the application of PCA dimension reduction in the two features (CNN and BRISK feature) are discussed. Finally, the result of CNN, BRISK, and the combined features are discussed in detail.

CHAPTER FIVE

5. CONCLUSION AND RECOMMENDATIONS

5.1. Conclusion

Skin is the biggest organ and most sensitive part of the human body which protects our inner vital parts and organs from the outside environment. Skin also helps in body temperature regulation. However, the Skin may be infected by fungal infection, bacteria, allergy, or viruses, etc. Skin diseases are more common than other diseases. A skin disease may change texture or color of the skin. At large, skin diseases are chronic, infectious and sometimes may develop into skin cancer. Therefore, developing a model that identifies the disease of skin to reduce their development and spreading in early stage.

In this thesis, we developed a model to detect and classify human skin diseases which is Tinea capitis, Tinea Corporis, Tinea pedis and Healthy skin. In this study, we follow experimental research, which involves data set preparation for training and testing the Human skin disease identification model, experimentation and evaluation. First we collected diseased skin and Healthy skin images from Tibebe Giyon Specialized Hospital, Felege Hiywet Comprehensive Specialized Hospital, Gamby General Teaching Hospital in Bahir Dar city and DermNet.com image repositories. We have prepared a total of 2,226 images. During collecting an image, the Specialized Doctors of the Hospitals are guided very well in the occasion of identifying and labeling the Human skin disease images. Then images are pre-processed using resizing by bi-cubic interpolation, noise removed by Median filter, contrast enhancement by Histogram equalization and global thresholding segmentation. After preprocessed the images, we proposed new feature extraction techniques using CNN, BRISK and the combination of the two. We Have to use the BRISK technique for the fact that CNN was not invariant for image transformation (rotation and illumination change) while BRISK is to make CNN invariant with scale and position and we have to applied Principal Component Analysis (PCA) to reduce the dimension of the combined features. After we extracted the features, the classification model is using SVM and SoftMax classifier. Experimental results show that the combination of deep features by CNN and local features by BRISK with the SVM classifiers can achieve better classification performance with an accuracy of 98.44%. And also, we compared the proposed Human skin diseases identification model with the current CNN architecture AlexNet, and the proposed model has better performance accuracy than AlexNet.

5.2. Recommendation

The proposed work can be further extended to improve the performance. Based on the investigation and findings of the study, we recommended for future and further research works:

1. This study only considered three common fungal Human diseased skin and healthy skin. As a result, future work should be considering and recognize other human skin diseases.
2. Compare the model with other predefined CNN architecture such as Google Net, VGG NET and Le Net with large dataset.
3. Use other segmentation techniques to increase the performance of the proposed model.

REFERENCES

Azadeh Noori Hoshyar , Adel Al-Jumailya, Afsaneh Noori Hoshyar (2014).The Beneficial Techniques in Preprocessing Step of Skin Cancer Detection System *Comparing a University of Technology, Sydney (UTS), Sydney, Australia andUniversity Putra Malaysia (UPM),Selangor,Malaysia. (http://creativecommons.org/ licenses/by-ncnd/3.0/).doi:10.1016/j.procs.2014.11.029*.

Awad, Ali Ismail & Hassaballah, Mahmoud. (2016*).* Image Feature Detectors and Descriptors; Foundations and Applications*. pp 1-5 doi: 10.1007/978-3-319-28854-3.*

Almaliki, Alaa & Alyousuf, Farah & Din, Roshidi. (2020). Review on techniques and file formats of image compression*. Bulletin of Electrical Engineering and Informatics. 9. 602-610. 10.11591/eei.v9i2.2085.*

Abrham Debasu Mengistu , Dagnachew Melesew Alemayehu and Seffi Gebeyehu Mengistu (2016).*Ethiopian Coffee Plant Diseases Recognition Based on Imaging and Machine Learning Techniques. international Journal of Database Theory and Application Vol.9,* No.4 (2016), pp.79-88 *http://dx.doi.org/10.14257/ijdta.2016.9.4.07*

A.Singh, S. Yadav and N. Singh, (2016)"Contrast enhancement and brightness preservation using global-local image enhancement techniques," *2016 Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC), ,* pp. 291-294, doi: 10.1109/PDGC.2016.7913162.

Awad, Ali Ismail & Hassaballah, Mahmoud. (2016). Image Feature Detectors and

Descriptors; *Foundations and Applications.* pp. 1-17 *10.1007/978-3-319-28854-3.*

Belal Ahmad , Mohd Usama , Chuen-Min Huang , Kai Hwang , M. Shamim Hossain , And Ghulam Muhammad (2020). Discriminative Feature Learning for Skin Disease Classification Using Deep Convolutional Neural Network. pp.39025-39033

Britannica, T. Editors of Encyclopaedia (2021, July 21*). image processing. Encyclopedia Britannica.* *https://www.britannica.com/technology/image-processing*

Desalegn Ashebr, 2020 Automatic Flower Disease Identification Using Deep Convolutional Neural Network. *Master's thesis. Department of Software Engineering, Bahir Dar University*

Dighe, P., & Guru, S.K. (2014). Survey on Image Resizing Techniques. Department of Computer Engineering, *D. Y. Patil college of Engineering, Akurdi, Pune 411 044, Savitribai Phule Pune University, India*. pp. 1444-1446

Enquhone Alehegn, *(*2017). Maize Leaf Diseases Recognition And Classifiaction Based On Imaging And Machine Learning Techniques. Master s thesis Bahir Dar, University pp. 46-48

Fan, L., Zhang, F., Fan, H. *et al.* (2019). *Brief review of image denoising techniques. Vis. Comput. Ind. Biomed. Art 2***,** *7.* *https://doi.org/10.1186/s42492-019-0016-7*

Gan W.S. (2020) Digital Image Enhancement. *In Signal Processing and Image Processing for Acoustical Imaging. Springer, Singapore.* PP.61-67 *https://doi.org/10.1007/978-981-10-5550-8_12.*

G. Kumar and P. K. Bhatia, (2014)."A Detailed Review of Feature Extraction in Image Processing Systems," *2014 Fourth International Conference on Advanced Computing & Communication Technologies,* pp. 5-12, *doi: 10.1109/ACCT.2014.74.*

Ghosh, Anirudha & Sufian, A. & Sultana, Farhana & Chakrabarti, Amlan & De, Debashis. (2020). Fundamental Concepts of Convolutional Neural Network. Recent Trends and Advances in Artificial Intelligence and Internet of Things. (pp.519-567) *Doi: 10.1007/978-3-030-32644-9_36.*

Haymanot Filie, (2020). *Human Skin Disease Detectionand Classification Using Machine Learning Algorithms. Master s thesis Bahir Dar, University pp. 1-16*

Ju Xu, Mengzhang Li, and Zhanxing Zhu (2020). Automatic Data Augmentation for 3D Medical Image Segmentation. *School of Mathematical Sciences, Peking University, Beijing, China.*pp.1-3

Kamalahasan, Manjula & Vijayarekha, K & Vimaladevi, P. (2017). Review On Classification Algorithms In Image Processing. *Vol 2.* Issue 11 pp. 1-6.

K. Jeevitha , A. Iyswariya , V. RamKumar , S. Mahaboob Basha , V. Praveen Kumar.(2020). *A REVIEW ON VARIOUS SEGMENTATION TECHNIQUES IN IMAGE PROCESSSING, European Journal of Molecular & Clinical Medicine R.M.K. Engineering College, Kavaraipettai, Chennai, Tamilnadu, India. Vol 7,* Issue 4, PP. 1342-1346

K. Manjula, K. Vijaya Rekha, P.Vimaladevi(2017).Review On Classification Algorithms In Image Processing. *International Journal of Innovative Trends in Engineering & Research Volume 2,* Issue 11. pp.1-7

Lawton S (2019) Skin 1*: the structure and functions of the skin. Nursing Times [online]; 115,* 12, 30-33

Leelavathy S, Jaichandran R, Shobana R, Vasudevan, Sreejith S Prasad and Nihad (2020). Skin Disease Detection Using Computer Vision and Machine Learning

Technique. *European Journal of Molecular & Clinical Medicine ISSN 2515-8260 Volume 7*, Issue 4, pp.2999-3003.

Liu, Yu. (2018*)*. Feature Extraction and Image Recognition with Convolutional Neural Networks*. Journal of Physics: Conference Series.* pp.1-7 1087. 062032. *10.1088/1742-6596/1087/6/062032.*

Lorentzon, M. (2017). Feature Extraction for Image Selection Using Machine Learning (Dissertation). *Master of Science Thesis in Electrical Engineering Department of Electrical Engineering, Linköping University, 2017 Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-142095*

L. Deng and D. Yu. (2013) Deep Learning: Methods and Applications. Foundations and Trends in Signal Processing*, vol. 7, nos. 3–4,* pp. 197–387. DOI*:* 10.1561/2000000039

Manish Kumar and Rajiv Kumar (2016). An Intelligent System to Diagnosis the Skin Disease. *ARPN Journal of Engineering and Applied Sciences VOL. 11*, NO. 19, OCTOBER 2016 pp.11368-11370

Maneela Jain & Pushpendra Singh Tomar (2013). Review of Image Classification Methods and Techniques*. International Journal of Engineering Research & Technology (IJERT) Vol.* 2 Issue 8, pp. 852-859

M. Maheshan.H. Prasanna Kumar.(2019) *Performance of image pre-processing flters for noise removal in transformer oil images at diferent temperatures. Department of Electrical Engineering, University Visvesvaraya College of Engineering, Bengaluru, Karnataka, India. pp.* https://doi.org/10.1007/s42452-019-1800-x

Mrs. Jayashree Hajgude, Aishwarya Bhavsar, Harsha Achara, Nisha

Khubchandani.(2019) Skin Disease Detection Using Image Processing with Data Mining and Deep Learning . *International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 06* Issue: 04. pp. 4363-4366.

Srisha, Ravi & Khan, Am. (2017). Global Thresholding Techniques to Classify Dead Cells in Diffusion Weighted Magnetic Resonant Images.pp.463-465

Shorten, Connor & Khoshgoftaar, Taghi. (2019*).* A survey on Image Data Augmentation for Deep Learning*. Journal of Big Data. 6. 10.1186/s40537-019-0197-0.*

Shreyas F. (2014) Image Interpolation Techniques in Digital Image Processing. *Int. Journal of Engineering Research and Applications www.ijera.com ISSN: 2248-9622, Vol. 4,* Issue 10(Part I), October, pp.70-73

Sofia (2016). New Mixed Kernel Functions of SVM Used in Pattern Recognition. *Cybernetics and information technologies volume 16*, No 5 pp. 6-7. DOI: 10.1515/cait-2016-0047.

Sukhjinder Kaur. (2015). Noise Types and Various Removal Techniques. *International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE) Volume 4*, Issue 2, pp. 226-228.

Tyagi, Vipin. (2018). Understanding Digital Image Processing. *Department of Computer Science and Engineering Jaypee University of Engineering and Technology Raghogarh, Guna (MP), India .pp 1-3 .10.1201/9781315123905*

Vikas Kumar Mishra, Shobhit Kumar and Neeraj Shukla (2017). Image Acquisition and Techniques to Perform Image Acquisition. pp. 21-23 *DOI: 10.18090/samriddhi. v9i01.8333*

V B, Shereena & David, Julie. (2015). Significance of Dimensionality Reduction in

Image Processing. Signal & Image Processing: *An International Journal. 6.* 27-42. *10.5121/sipij.2015.6303.*

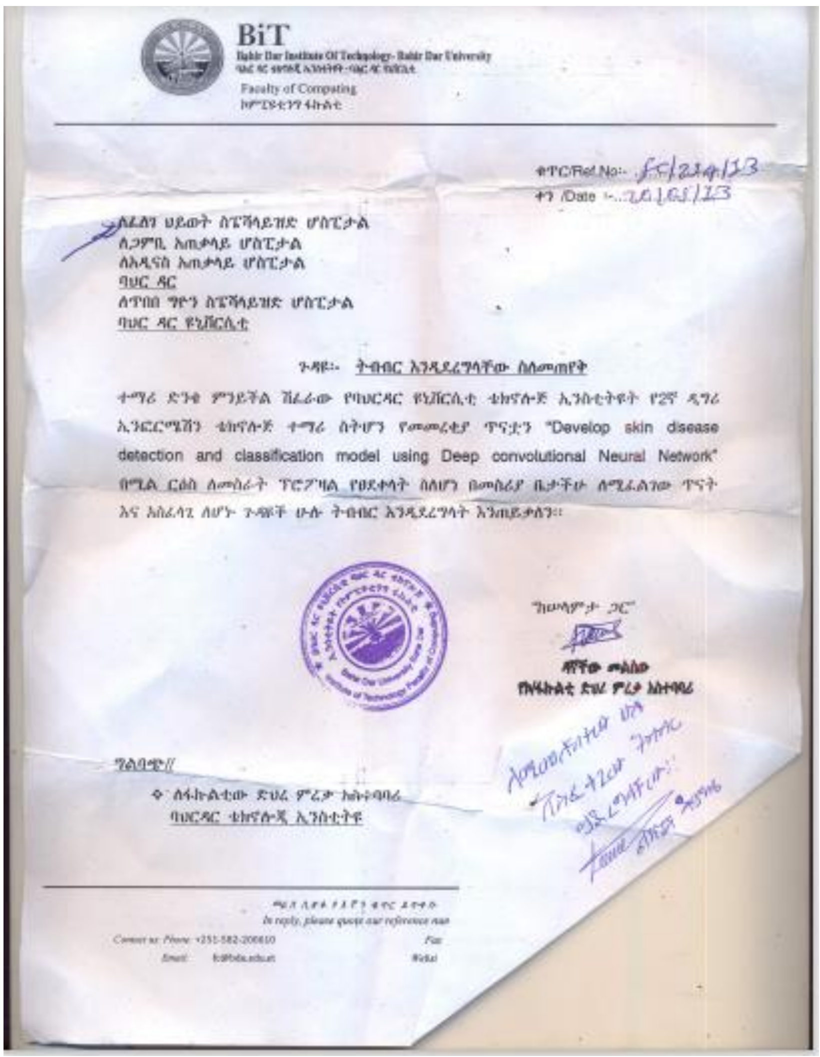
V. Pugazhenthi, Sagar Naik, Amruta Joshi, Shreya Manerkar, Vinita Nagvekar, et al. (2019). Skin Disease Detection and Classification. *International Journal of Advanced Engineering Research and Science (IJAERS6)* (5), pp.396-400. ff10.22161/ijaers.6.5.53ff. Ffhal-02141241f.

Yamashita, R., Nishio, M., Do, R.K.G. *et al.* (2018*). Convolutional neural networks: an overview and application in radiology*. *Insights Imaging 9*, pp.611–629 *https://doi.org/10.1007/s13244-018-0639-9*

Yuan, W., Hu, F. & Lu, L. *(2021*. A new non-adaptive optimization method: Stochastic gradient descent with momentum and difference. *Appl Intell).* *https://doi.org/10.1007/s10489-021-02224-6*

(2021, April 1). Retrieved from Math Works: https:// www.mathworks .com/help/vision/ug/ local-feature-detection-and-extraction.html#bulrsev-1 last time accessed oct 25/2021.

Appendix



BiT

Bahir Dar Institute Of Technology- Bahir Dar University

Faculty of Computing

ኮምፒዩቲንግ ፋኩልቲ

ቁጥር/Ref.No:- ፍና/234/13

ቀን /Date :- 20/05/13

ለፈለገ ህይወት ስፔሻላይዝድ ሆስፒታል
ለጋምቢ አጠቃላይ ሆስፒታል
ለአዲሳስ አጠቃላይ ሆስፒታል
ባህር ዳር
ለጥበበ ግዮን ስፔሻላይዝድ ሆስፒታል
ባህር ዳር ዩኒቨርሲቲ

ጉዳዩ፡- ትብብር እንዲደረግላቸው ስለመጠየቅ

ተማሪ ድንቁ ምንይችል ሽፈራው የባህርዳር ዩኒቨርሲቲ ቴክኖሎጅ ኢንስቲትዩት የ2ኛ ዲግሪ ኢንፎርሜሽን ቴክኖሎጅ ተማሪ ስትሆን የመመረቂያ ጥናቷን "Develop skin disease detection and classification model using Deep convolutional Neural Network" በሚል ርዕስ ለመስራት ፕሮፖዛል የፀደቀላት ስለሆነ በመስሪያ ቤታችሁ ለሚፈለገው ጥናት እና አስፈላጊ ለሆኑ ጉዳዮች ሁሉ ትብብር እንዲደረግላት እንጠይቃለን።

"ከሠላምታ ጋር"

ፋኩልቲ ድህረ ምረቃ አስተባባሪ

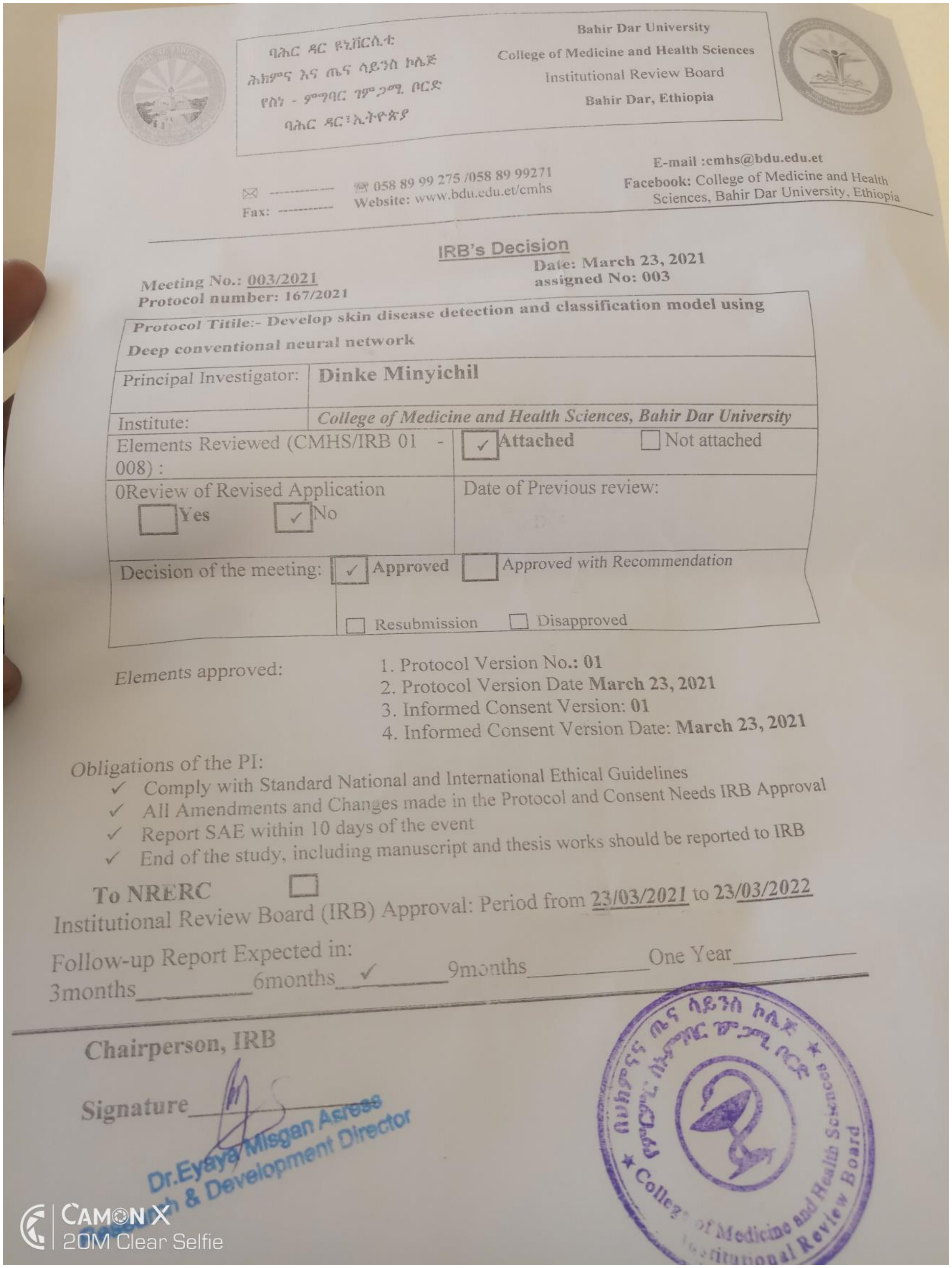
ግልባጭ//

ፍ ለፋኩልቲው ድህረ ምረቃ አስተባባሪ
ባህርዳር ቴክኖሎጂ ኢንስቲትዩ

In reply, please quote our reference number

Contact us: Phone: +251-582-206610 Fax:

Email: Website:



ባሕር ዳር ዩኒቨርሲቲ
ሕክምና እና ጤና ሳይንስ ኮሌጅ
የስነ - ምግባር ግምገማ ቦርድ
ባሕር ዳር፣ኢትዮጵያ

Bahir Dar University
College of Medicine and Health Sciences
Institutional Review Board
Bahir Dar, Ethiopia

✉ ------------ ☏ 058 89 99 275 /058 89 99271
Fax: ----------- Website: www.bdu.edu.et/cmhs

E-mail :cmhs@bdu.edu.et
Facebook: College of Medicine and Health Sciences, Bahir Dar University, Ethiopia

IRB's Decision

Meeting No.: 003/2021
Protocol number: 167/2021

Date: March 23, 2021
assigned No: 003

Protocol Titile:- Develop skin disease detection and classification model using Deep conventional neural network	
Principal Investigator:	**Dinke Minyichil**
Institute:	*College of Medicine and Health Sciences, Bahir Dar University*
Elements Reviewed (CMHS/IRB 01 - 008) :	☑ **Attached** ☐ Not attached
0Review of Revised Application ☐ **Yes** ☑ No	Date of Previous review:
Decision of the meeting:	☑ **Approved** ☐ *Approved with Recommendation* ☐ Resubmission ☐ Disapproved

Elements approved:

1. Protocol Version No.: **01**
2. Protocol Version Date **March 23, 2021**
3. Informed Consent Version: **01**
4. Informed Consent Version Date: **March 23, 2021**

Obligations of the PI:

- ✓ Comply with Standard National and International Ethical Guidelines
- ✓ All Amendments and Changes made in the Protocol and Consent Needs IRB Approval
- ✓ Report SAE within 10 days of the event
- ✓ End of the study, including manuscript and thesis works should be reported to IRB

To NRERC ☐

Institutional Review Board (IRB) Approval: Period from 23/03/2021 to 23/03/2022

Follow-up Report Expected in:
3months________6months__✓____9months__________One Year__________

Chairperson, IRB

Signature________

Dr.Eyaya Misgan Asress
Research & Development Director

College of Medicine and Health Sciences Institutional Review Board

CAMON X
20M Clear Selfie