2020

Multi Label Amharic Text Classification Using Convolutional Neural Network Approaches

Wubalem, Habtamu

BAHIR DAR UNIVERSITY

BAHIR DAR INSTITUTE OF TECHNOLOGY

SCHOOL OF RESEARCH AND POSTGRADUATE STUDIES

FACULTY OF COMPUTING

HUMAN SKIN DISEASE DETECTION AND CLASSIFICATION USING MACHINE LEARNING ALGORITHMS

HAYMANOT FILIE MOGES

BAHIR DAR, ETHIOPIA

August 18, 2020

HUMAN SKIN DISEASE DETECTION AND CLASSIFICATION USING MACHINE LEARNING ALGORITHMS

Haymanot Filie Moges

A Thesis submitted to the school of Research and Graduate Studies of Bahir Dar

Institute of Technology, BDU in Partial Fulfillment of the requirements for the Degree of Master of Science in Information Technology in the Faculty of Computing.

Advisor Name: Mekuanint Agegnehu (PhD)

Bahir Dar, Ethiopia

August 18, 2020

DECLARATION

I, the undersigned, declare that the thesis comprises my own work. In compliance with internationally accepted practices, I have acknowledged and refereed all materials used in this work. I understand that non-adherence to the principles of academic honesty and integrity, misrepresentation/ fabrication of any idea/data/fact/source will constitute sufficient ground for disciplinary action by the University and can also evoke penal action from the sources which have not been properly cited or acknowledged.

Name of the student: Haymanot Filie

Student's Signature________

Date of submission: 18/08/2020 G.C

Place: Bahir Dar

This thesis has been submitted for examination with my approval as a university advisor.

Advisor Name: Mekuanint Agegnehu (PhD)

Advisor's Signature: ________

Bahir Dar University

Bahir Dar Institute of Technology

School of Research and Graduate Studies

Faculty Of Computing

Approval of thesis for defense result

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

Name of Student Haymanot Fike Signature ____________ Date 14/08/2020

As members of the board of examiners, we examined this thesis entitled "Human Skin Disease Detection and Classification Using Machine Learning Algorithms"

By *Haymanot Fike*. 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
Mekuanint Agegnehu (PhD)		14/08/2020 G.C
Name of External examiner	Signature	Date
Fekade Getahun (PhD)		12/08/2020
Name of Internal Examiner	Signature	Date
Abrham Debasu (Ass.Professor)		14/08/2020 G.C
Name of Chairperson	Signature	Date
Gebeyehu Belay (PhD)		15/08/2020 G.C
Name of Chair Holder	Signature	Date
Derejaw Lake		14/08/2020 G.C
Name of Faculty Dean	Signature	Date
Belete Biazen		14/08/2020

iii

To my lovely family

ACKNOWLEDGEMENTS

First of all, I would like to thank to God for giving me the strength, ability and opportunity to undertake this thesis, persevere, and complete it satisfactorily. I praise and thank Him for His great help throughout this work. I would also thank the Ever-Virgin, St. Marry, Mother of our Lord.

I would like to express the greatest appreciation to my advisor Mekuanint Agegnehu (PhD). I am deeply thankful to him for giving his insights and constructive comments, continuous guidance, ideas, and suggestions from the initial proposal to the end of this research.

I am grateful to give my thanks for staff members of Bahir Dar Felege Hiywet General Specialized hospital and Tibebe Giyon Specialized hospital. Individually, I would like thanks Dr Mulunesh and Sister Tigist for their continuous support throughout data collection time.

My sincere thanks also go to my friends Getachewu Worku, Desalegn Ligaba, Selam Getinet, Firehiywot Mekuriawu, and Tigist Hidiru for their valuable assistance and support in this work. I would also thanks all my classmates who have spent many hours for discussion to make things go right and within time.

Finally, I am deeply grateful to my sisters Eseynat Filie, Yewubmirt Sharewu, Elishaday Amare, Tigist Yeshambel, Metadel Getu, my brother Alemu Filie and my husband Mitiku Melkamu for their endless love, support and encouragement. I would like to thank my father, Mr. Filie Moges, and my mother, Mrs. Aleminat Awoke and all the rest of my families and friends, who in one or another brought me up to success in my academic endeavor. None of this would have been possible without the love and patience of my family and friends. Their support and care helped me to complete this thesis satisfactorily.

ABSTRACT

The human skin plays a huge part in a person's physical appearance, and it is the biggest organ of the human body. It regulates body temperature, sensing from touching heat and cold. However, there are a number of risks that affect the skin, one of which is a disease. Fungus, bacteria, allergies, enzyme, and viruses cause most skin diseases. Identify the disease on the basis of manual feature extractions or based on the symptoms is time consuming and it requires extensive knowledge for perfect identification. Diagnosing, detection and classification of skin diseases is done by researchers previously. However, the recognition rate is still not enough and is dependent on feature selection, filtering and segmentation methods. In the previous work tinea pedies and tinea capitis are not identified. In this thesis, we developed a model using Convolutional Neural Network (CNN) for feature extraction and Support Vector Machine (SVM) for classification. CNN is state of the art for deep feature extraction, hence we used it for feature extraction. The model used to detect and classify human skin diseases such as tinea corpories, tinea pedies, and tinea capitis. The dataset is collected from patients at Bahir Dar Tibebe Giyon specialized Hospital, Bahir Dar Felege Hiywet General specialized Hospital and from skin disease image repository. The amounts of collected images are 65, 58, and 62 for tinea capitis, tinea corpories, and tinea pedies respectively. From the total 185 images 80 images are from medical image repository. After collecting datasets, Image augmentation, Image Preprocessing, and Image Segmentation techniques are applied to increase the performance of human skin disease identification. In image preprocessing, we normalize the image to 3 image sizes which are 200x200, 224x224, and 300x300. We also apply median filtering to remove noise and we used Histogram equalization to balance the intensity of image. In segmentation, we used integration of threshold and region based segmentation methods. From 1196 image dataset, we used 80% of image for training and 20% for testing. After the model is evaluated, we have achieved 95.6% test accuracy and 95.6% training accuracy which is 11.6% better than AlexNet model.

Keywords: Machine learning, skin disease, convolutional neural network, support vector machine

TABLE OF CONTENTS

DECLARATION .... ERROR! BOOKMARK NOT DEFINED.

ACKNOWLEDGEMENTS V

ABSTRACT VI

LIST OF ABBREVATIONS XI

LIST OF FIGURES XIII

LIST OF TABLES XV

LIST OF ALGORITHM XVI

CHAPTER ONE: INTRODUCTION 1

1.1. Background 1

1.2. Motivation 3

1.3. Statement of Problem 4

1.4. Objectives of the study 6

1.4.1. General objective 6

1.4.2. Specific objectives 6

1.5. Scope and limitation of the study 6

1.6. Methodology 6

1.6.1. Image acquisition and preprocess 7

1.6.2. Implementation Environment 7

1.6.3. Train HSDCN model 7

1.6.4. Model 8

1.7. Significance **8**
1.7.1. For dermatology 8
1.7.2. For customer 8
1.7.3. Economically 9
1.7.4. Researchers 9

1.8. Thesis organization **10**

CHAPTER TWO: LITRATURE REVIEW **11**

2.1. Introduction **11**

2.2. General Human skin structure **11**

2.3. Types of skin lesion in this study **15**
2.3.1 Fungal skin disease 15

2.4. Digital image processing **18**
2.4.1. Image Acquisition 20
2.4.2. Image preprocessing 20
2.4.3. Image segmentation 22
2.4.4. Feature extraction 25
2.4.5. Classification 27

2.5. Image Classification techniques **27**
2.4.6. Non deep learning Techniques 27
2.4.7. Deep learning Techniques 29

2.6. Component of Convolutional neural network **32**

2.7. Common Architectures in Convolutional Neural Networks **38**

2.8. Related works **40**
2.8.1. Detection and classification of skin disease using non deep learning approach 40

2.8.2. Detection and classification of skin disease using deep learning approach..... 41

2.9. Summery 43

CHAPTER THREE: METHODOLOGY 44

3.1. Introduction 44

3.2. The architecture of the proposed system 44

3.2.1. Image acquisition 45

3.2.2. Image Preprocess 46

3.2.3. Data augmentation 49

3.2.4. Image Segmentation.. 50

3.2.4. Feature learning 53

3.2.5. Classification... 54

3.3. Summery 56

CHAPTER FOUR: RESULTS AND DISCUSSION 57

4.1. Introduction 57

4.2. Dataset preparation.. 57

4.3. Implementation Tools.. 58

4.4. Hyper Parameters in the model 59

4.5. Evaluation techniques 60

4.6. Evaluation result.. 61

4.6.1. HSDCNet model using 200X200 preprocessed image and Relu at training phase 62

4.6.2. HSDCNet model using 200X200 preprocessed image and Relu at validation phase 63

4.6.3. HSDCNet model using 200X200 preprocessed image and Tanh 64

4.6.4. HSDCNet model using 224X224 preprocessed image and Relu 65
4.6.5. HSDCNet model using 224X224 preprocessed image and Tanh 66
4.6.6. HSDCNet model using 300x300 preprocessed image and Tanh 67
4.6.7. HSDCNet model using 200x200 image, segmentation and Tanh 68

4.7. Model evaluation 69

4.8. Comparison with deep neural network architecture 71
4.8.1. Alex net implementation result 72

4.9. Summarizing evaluation Results.. 75

4.10. Discussion 76

4.11. Summery 77

CHAPTER FIVE: CONCLUSIONS AND RECOMMENDATIONS......................... 78

5.1. Conclusions 78

5.2. Recommendation.. 79

REFERENCE 80

LIST OF ABBREVATIONS

AI	Artificial Intelligence
CBR	Case-based Reasoning
CCV	Color coherence vector
CM	color moments
CNN	Convolution Neural Network
DBM	Deep Boltzmann Machine
DBN	deep belief network
DNN	deep neural network
GLCM	Gray-Level Co-Occurrence Matrix
HSDCN	Human Skin Disease classification network
IDE	Integrated Development Environment
JPEG	Joint Photographer Expert Group
KNN	K-Nearest Neighbors
LDA	Linear Discriminate Analysis
MLP	Multilayer Perceptron Classifier
OpenCV	Open Computer Vision
RBF	Radial basis function

RBM	Restricted Boltzmann Machines
RGB	Red, Green, Blue
RNN	Recurrent neural networks
Spyder	Scientific Python
SVM	Support Vector Machine
SciPy	Scientific Python Development Environments

LIST OF FIGURES

Figure 2. 1 The general human skin structure (Hoffman, 2019) 15
Figure 2. 2 Tinea capitis 17
Figure 2. 3 Tinea Pedies 17
Figure 2. 4 Tinea Corporis 18
Figure 2. 5 Steps of digital image processing 20
Figure 2. 6 Basic convolutional neural network architecture (Antonio & Sujit, 2017) ... 32

Figure 3. 1 Proposed human skin disease identification architecture 45
Figure 3. 2 Original RGB image and the Grayscale image respectively 47
Figure 3. 3 the original and the segmented image respectively 51
Figure3.4 The original and the segmented image using region based techniques respectively 52
Figure 3. 5The original and the segmented image using integration techniques respectively 52

Figure 4. 1Sample data set 58
Figure 4 2 Training accuracy of the model using Training phase 62
Figure 4 3 Training Loss the model using Training phase 63
Figure 4 .4 Training and validation accuracy using preprocessed data and relu 63
Figure 4. 5Training and validation loss using preprocessed data and relu 63
Figure 4. 6Training and validation accuracy using Tanh, prepossessed data 64
Figure 4 .7Training and validation loss using Tanh, prepossessed data 64
Figure 4. 8Training and validation accuracy using 224x224 and Relu 65
Figure 4 .9Training and validation loss using 224x224 and Relu 66
Figure 4 .10Training and validation accuracy using 224x224 and Tanh 66

Figure 4 .11Training and validation loss using 224x224 and Tanh 67
Figure 4. 12 Training and validation accuracy using 300x300 and Tanh 67
Figure 4 .13Training and validation loss using 300x300and Tanh 68
Figure 4 .14Training and validation accuracy using integration segmentation method, 200x200 and Tanh 69
Figure 4 15Training and validation loss using integration segmentation method, 200x200 and Tanh 69
Figure 4 .16 Training Progress of HSDCNet 70
Figure 4 .17Confusion matrix result and classification report of HSDCNet 70
Figure 4 .18Alex net Training, validation accuracy and training and validation Loss 73
Figure 4 .19 Training progress of Alex Net Model 74
Figure 4 .20 Confusion matrix result and classification report of AlexNet Model............ 74

LIST OF TABLES

Table2. 1 Skin types based on the reaction to UV (Bartlett, 2020). 13

Table2. 2 Image contrast enhancement techniques 22

Table2. 3 image segmentation techniques.. 24

Table 4. 1 Number of collected and augmented dataset... 57

Table 4. 2 Hyper parameters and value in the model 60

Table 4. 3 Summarizing training model result 75

LIST OF ALGORITHM

Algorithm3- 1 Image processing of resized image 46
Algorithm3- 2 Image shape ordering 47
Algorithm3- 3 Balance the intensity of image 48
Algorithm3- 4 Noise removal using median filter 49
Algorithm3- 5 Algorithm of augment image 50

CHAPTER ONE: INTRODUCTION

1.1. Background

Skin is the body's outer covering, which protects us from microbes and the elements. The skin has ability to reduce harmful effect of ultraviolet radiation (UV) due to the pigment melanin that absolves UV radiation, which protects the cell’s nuclei from DNA damage. The skin regulates body temperature and also synthesizes vitamin D3. It permits the sensations of touch, cold, and heat. It is the largest organ of the body, with a total area of about 20 square feet. Skin has three layers. The first layer is called epidermis, which is the outermost layer of skin. It provides a waterproof barrier and creates our skin tone. The second layer beneath to the epidermis is called Dermis, which contains tough connective tissue, hair follicles, and sweat glands. The deeper subcutaneous tissue is called hypodermis, which is made of fat and connective tissue (Anatomy, 2019). Since skin is the outer covering of our body mostly it affected by bacteria, fungus and viruses. All the organs inside the human body are completely protected by skin. Therefore it is important to give attention to the complete care of the skin. Because any change in its normal functioning can cause to affect the other parts of the body. Skin disease here refers to disorders of exclusively the superficial layers of the skin. Skin diseases are the most common disease in humans among all age groups and a significant root of infection in Africa (Hay, et al., 2016).

Common diseases that attack human skin are eczema, melanoma, vitiligo, mycosis, papillomas, impetigo, scabies, herpes, dermatitis, wart, psoriasis, acne, tinea capitis, tinea corpories, tinea pedies, etc (Hay, et al., 2016). Skin diseases which may be of the bacterial, fungal, allergies, enzyme, etc. are very harmful to the skin and can spread throughout if not detected accurately as early as possible. Hence, identification of these diseases plays a vital role to detect the type of disease accurately in the early stage.

Diseases like tinea corpories, tinea pedies, 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 of four conditions: scabies, capitis, tinea capitis, or pyoderma (Hay, et al., 2016). Dermatology is the discipline dealing with diagnosis and management of human skin disease. In Ethiopia Dermatologist diagnosed using symptoms and sometimes they used a blood test for further analysis. The diagnosis process is seen to be intensely laborious, time-consuming and requires an extensive understanding of the domain.

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.

Human skin disease detection and diagnosis is done previously by researchers. Skin disease such as herpes, dermatitis, and psoriasis are recognized by the model proposed by researchers in (Li-sheng, Quan, & Tao, 2018). Eczema and Acne are also detected by the system developed in (Carl Louie Aruta, Calaguas, Gameng, & Prudentino, 2015). (Mrs. Jayashree, Aishwarya, Harsha, & Nisha, 2019) Also able to detect and classify melanoma, eczema, and impetigo. Melanoma, Bullae, Seborrheic keratosis, Shingles and Squamous cell are detected and classified (Sumithra, Mahamad, & Guruc, 2015). An expert system also developed using CNN to Diagnosis Atopic dermatitis, Acne vulgaris and Scabies (Sam, Philip, Nartey, & Nti, 2019). However, no research have been published which explores deep convolutional network and support vector machine for identification of three human skin disease such as tinea corpories, tinea pedies, and tinea capitis.

Nowadays, we are in the technology era, which makes possible for proper identification of human skin disease. In this thesis, we develop a model used for human skin disease detection and classification using deep learning techniques. We used a CNN for deep feature extraction and SVM for classification. The development of a proper methodology for skin disease identification is quite useful for enhance performance of identification. The main goal of this work is to construct deep convolutional neural networks to achieve effective identification for three skin diseases by using human skin disease images. The basic tasks performed during the development of model are identifying skin diseases, preparing datasets, and preprocessed datasets, image augmentation, image segmentation, extract features using feature extraction method, classifying the identified skin diseases, and finally evaluate the performance of the proposed method.

1.2. Motivation

For the last three decades many research and image classification model has been done, which shows that this area is so significant that it has gained continuous research attention and still a hot research issue. Even though there are many research papers which are done on skin disease detection and classification, still there is a gap to be filled. Most of the research work in this area is concentrated on white skin color image. But in reality there are many parameters which affect the detection (Kamulegeya, et al., 2019). In the previous work all common skin disease are not recognized. Medical image segmentation is a critical issue for identification of skin lesion area. After analyzing these problems we come up to contribute a little effort by develops our model including common skin disease those are not identifiable in the previous work.

1.3. Statement of Problem

Skin plays a significance role for our body. However, it affected by bacteria, viruses, allergies, enzymes, and Funguses. Common human skin diseases are eczema, melanoma, Vitiligo, mycosis, Papillomas, impetigo, scabies, herpes, dermatitis, wart, psoriasis, acne, tinea corpories, tinea pedies, tinea capitis, etc (Hay, et al., 2016). Those are very harmful to skin and can spread throughout if not detected accurately as early as possible. Skin diseases have psychological effects on humans. It affects us mentally because of visible effects of skin diseases in a human body.

Early detection and diagnosis of skin disease is critical to prevent it from further spread. In Ethiopia, Dermatologist diagnosis using symptoms and sometimes a blood test is used for further analysis. The diagnosis process is seen to be intensely laborious, time-consuming and requires an extensive understanding of the domain. The naked eye visualization of experts is the main old-style approach adopted for the recognition and identification of human skin disease.

Detection, diagnosis and classification were done previously by different researchers. In the previous works, researchers used different methodologies and the performance rate of their models are also different. (Li-sheng, Quan, & Tao, 2018), (Carl Louie Aruta, Calaguas, Gameng, & Prudentino, 2015), (Mrs. Jayashree, Aishwarya, Harsha, & Nisha, 2019), (Petrellis, 2018), and (Sumithra, Mahamad, & Guruc, 2015) developed a model for identification of human skin disease however, the recognition rate is still not enough and is dependent on feature selection methods. A smart phone based application was developed by other researchers (Velasco, et al., 2019).

The human skin disease such as tinea pedies and tinea capitis are the most common disease that were not considered in the previous human skin disease identification model. In this thesis, we develop a model to detect and classify tinea corpories, tinea pedies, and tinea capitis.

Skin lesion segmentation must be done using an effective segmentation method. There are many researches being done for automatic skin lesion segmentation to solve the problem of accuracy in CAD of skin lesion. Skin lesion segmentation is an active research area because the performance of the classification is depends on the segmented image. Therefore, a lot must be done to improve performance of skin lesion segmentation because it is a fatal public health issue related to skin disease. Threshold and region based segmentation for human skin lesion segmentation solves the problem of human skin lesion detection.

Finally, this thesis answers the following research questions:

1. Which image segmentation method is better to improve skin disease classification performance?
2. Which image size is better to improve performance and computational time
3. Which activation function is better to improve performance and computational time
4. What is the performance of proposed Model?

1.4. Objectives of the study

1.4.1. General objective

The general objective of our thesis is to develop a model for human skin disease detection and classification using a machine learning algorithms.

1.4.2. Specific objectives

The specific tasks to achieve the general objective are listed below

- To compare segmentation algorithms that best suited for human skin disease identification.
- To compare three different image sizes
- To compare two different activation function
- To test and evaluate the performance of the proposed model

1.5. Scope and limitation of the study

The proposed model recognizes common skin diseases in Ethiopia such as tinea corpories, tinea pedies and tinea capitis. Disease types which cannot be identified by experts and not possible to capture by camera did not included in this study.

1.6. Methodology

In this research we have used a design science research methodology. It includes phases such as problem identification and motivation, objectives for a solution, design and development, evaluation, and communication. Problem identification is process of finding the gap or problem to be solved**.** Objectives of solution are followed after problem identification. Design and development phase includes data collection, preprocessed data, segmentation, feature learning, and finally classification. After build the model evaluation is taken to check the performance of the model.

1.6.1. Image acquisition and preprocess

We got the datasets from Bahir Dar Tibebe Giyon Specialized Hospital, Bahir Dar Felege Hiywet General Specialized Hospital, and from MedicineNet.com image repository. An image augmentation technique was applied to increase our datasets from the original dataset. Before train the network, we need to preprocess the skin image using resizing and rescale techniques to reduce computation time. We also need to minimize noise using filtering techniques as well as enhance contrast of image. Segmentation is a critical step in image processing for identification of human skin disease. It is a task done after preprocessed image. The performance of identification model is depending on the segmented image. If the lesion is segmented well the performance of the model improved. We select well suited segmentation techniques to improve the identification performance.

1.6.2. Implementation Environment

In order to implement the skin disease identification model, we made use several open source libraries. We used Anaconda environment and Python programming language for scientific computing. We used a free source Keras and TensorFlow library in this environment on windows 10 operating system. Keras is a free source CNN library written in Python. It is easily running on top of TensorFlow. TensorFlow is a representative math library and is also used for scientific computing tenders such as neural networks and these TensorFlow as a backend. Python code is written in Spyder IDE.

1.6.3. Train HSDCN model

In feature learning process we have two component image preprocessing (Resize image, Filter image, enhance contrast, image augmentation and segmentation), and deep learning (feature learning and classification). The first layer in CNN is a convolutional layer. The input to the first convolutional layer is a preprocessed image whereas the feature map is the input to the second convolutional layer. In order to get a feature map, the input image is convolved with or by the kernel filter to extract important features or find the most representative feature in a given region. After deep features extracted using the developed model, all features are input for SVM for final task of this work, which is classification.

The classification task is done using RBF kernel function .During convolution, we have considered important parameters: padding, stride, batch size, filter size, and dropout.

1.6.4. Model Evaluation

We have a total of 1196 image data set and we used 80% of dataset for train the model and the remaining 20% for testing. After building our model we evaluate the model using confusion matrix. It was evaluated by comparing its output against the observed data using precision, recall and f1-score values for evaluating classification accuracy. In addition, we conducted a comparison with state-of-the-art models Alex Net (Krizhevsky, E.Hinton, & Sutskever, 2012), which is used by previous work.

1.7. Significance

Skin disease is common in developing countries like Ethiopia. In Ethiopia dermatologist diagnoses skin infections based on symptoms. They spent a lot of time and labor during diagnosis and also sometimes health implications happened. This proposed method has advantage for dermatology, customer (patients) and also have economical advantage.

1.7.1. For dermatology

The proposed method helps for skin disease doctors (Dermatologist) to find out quickly of skin infection problems. Dermatologists used this proposed prototype as an assistant to detect and classify patients’ disease type. This helps for experts to make the accurate decision and minimize health implications.

1.7.2. For customer

As my observations’ in the hospitals most of patients spent much amount of time, labor and money. Most of the patients are come from other woredas, cities and kebeles. Customers (patients) can easily get the medication by the help of this proposed method and dermatologist. This reduces time, social consequences as well as health implications.

1.7.3. Economically

In Ethiopia skin disease experts are rare than other country. The proposed method solves this problem. Skin disease nurses can diagnosis by the help of this method. This reduces serious economic consequences.

1.7.4. Researchers

The models have great advantage for researchers. This thesis work have still gap as we described in chapter 5 section 5.2. Starting from these work researchers may start their work to fill this gap.

1.8. Thesis organization

The rest of the thesis is organized as follows. Chapter two discusses the literatures about general structure of the human skin, overview about common fungal infections and the related works. It discusses the basic concepts of Digital image processing and the general detection scheme of most CAD systems like preprocessing, feature extraction, feature selection, and classification of color images. Chapter three presents the proposed skin lesion detection and classification scheme and the steps followed. Chapter four discusses the data sets that have been used in the current study, the parameters used in the implementation part and the results obtained in all stages of the proposed method. The chapter also includes discussion on performance evaluation strategies employed in this thesis. Quantitative analysis is done in this chapter to evaluate the performance of the proposed method. The last chapter, Chapter 5, is about the conclusion and future works of the study

CHAPTER TWO: LITRATURE REVIEW

2.1. Introduction

Image processing is one of diagnostics techniques which grow dramatically. It forms core research area within engineering and computer science disciplines too. Currently, the use of digital image processing techniques have been applicable in many areas of interest such as medical visualization, law enforcement and inspection of the quality of agricultural products (V. Sandeep, K.Kanaka, & Keshavulu, 2013.)

In this chapter, the literatures related to the concepts that are basis for this thesis are reviewed. First, we present an overview of human skin structure and the fungal skin disease. Then, different techniques of digital image processing such as image acquisition, preprocessing, and different types of segmentation, feature extraction and classification are discussed in detail. Finally some related works reviewed and discussed.

2.2. General Human skin structure

Human skin is the major outer layer of human body. It surrounded all other parts of human body. 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. 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. Skin is composed of three primary layers: the epidermis, the dermis and the hypodermis (Hoffman, 2019).

Human skin has role to defense against germs that make you sick. Fungus, bacteria, and virus are causes for abnormal skin. In addition to providing a barrier that keeps germs out, the epidermis contains cells that help your immune system fight infection. Skin plays a big role in maintaining body temperature. On cold days, the brain reduces the amount of blood that travels to your skin. That helps retain heat within your body. On warm days, the brain increases blood flow to the skin, which has the opposite effect. The dermis contains sweat glands that help you to stay cool. Nerves are temperature sensitive. If you are in a warm environment, the sweat glands become active. The same thing can happen when you get nervous, which is why your hands, feet and armpits become sweaty in certain situations. Sweat travels up tiny tubes to the epidermis. When the sweat evaporates, it cools the surrounding skin (Bennett & Howard, 2020)

2.2.1. Epidermis

The epidermis, the outermost layer of skin, is made up of about 20 layers of tightly packed cells. It provides a waterproof barrier and creates our skin tone. The thickness of the epidermis varies depending on where on the body it is located. It is at its thinnest on the eyelids, measuring just half a millimeter, and at its thickest on the palms and soles at 1.5 millimeters (Brannon, 2020). The epidermis acts as a barrier that protects the body from ultraviolet radiation, harmful chemicals, and pathogens such as bacteria, viruses, and fungi (Hoffman, 2019). The epidermis is subdivided into five layers: stratum corneum, stratum lucidum, stratum granulosum, and stratum spinosum and stratum germinativum.

The epidermis can be impacted by more than just injury. This outermost layer is subject to both genetics and external forces that contribute to the aging of this skin. These factors include smoking, alcohol, and excessive UV exposure, all of which contribute to the development of wrinkles, sunspots, and the uneven thickening or thinning of the skin. The epidermis is also where rashes and blisters appear, caused by everything from infections and allergies. It is also the origin of both non-melanoma and melanoma skin cancers, and where certain diseases like diabetes and lupus can manifest with an array of dermatological symptoms (Brannon, 2020). Classification of the skin based on its reaction to UV radiation is shown in Table 1.1 below.

Table2. 1 Skin types based on the reaction to UV (Bartlett, 2020).

Type	Definition	Description
I	Always burns but never tans	Pale skin, red hair, freckles
II	Usually burns, sometimes tans	Fair skin
III	May burn, usually tans	Darker skin
IV	Rarely burns, always tans	Mediterranean
V	Moderate constitutional pigmentation	Latin American, Middle Eastern
VI	Marked constitutional pigmentation	Black

The skin's color is created by special cells called melanocytes. Melanocytes are located in the epidermis and it protects the skin from UV rays. These special cells produce a dark pigment called melanin. The amount of melanin in skin determines its color. If more melanin produced in epidermis, the skin becomes darker. Sometimes human being skin become yellow brown, which is because of ultraviolet light, simulates the release of melanin into the epidermis. However, tanned skin is damaged skin. So support the epidermis and put on lots of sunscreen is important for your skin (Bennett & Howard, 2020).

2.2.2. Dermis

The dermis is the second major layer of the skin beneath to the epidermis. The dermis contains a number of specialized cells and structures. These include: hair follicles, sweat glands, sebaceous glands that produce sebum which helps lubricate skin & hair (Hoffman, 2019). It is a thick layer made up of strong connective tissues. It is further divided into two levels. The first level is upper layer, which is made of loose connective tissue, called the papillary region. The second level is lower layer, which is made of tissue that is more closely packed, called the reticular layer. The dermis is made up of a matrix of collagen, elastin and network of capillaries and nerves. The collagen makes the skin strong, the elastin maintains its elasticity and the capillary network supplies nutrients to the different layers of the skin. Dermis also plays an important role in controlling our skin temperature and it mitigate against mechanical injury. When injured, the dermis makes well the formation of granulation tissue. Granulation tissue is rich in new blood vessels and many different cells. This tissue helps pull the edges of a cut or wound back together. Our body takes from 3 days up to 3 weeks to form this tissue (Virtual Medical, 2020)

2.2.3. Hypodermis

The hypodermis is the innermost (or deepest) and thickest layer of skin. It is also known as the subcutaneous layer or subcutaneous tissue. The deeper subcutaneous tissue or hypodermis is made of fat, connective tissue and elastin, which is an elastic protein that helps tissues return to their normal shape after stretching (Hoffman, 2019). The fatty layer differs in thickness with sex, race, and individual nutritional and hormonal status. This layer functions as an insulator, offering protection against the cold, and protects the body against heat as well through sweating (Bennett & Howard, 2020). The high levels of fat help insulate the body and prevent us from losing too much heat. The fat layer also acts as protection, padding our bones and muscles.

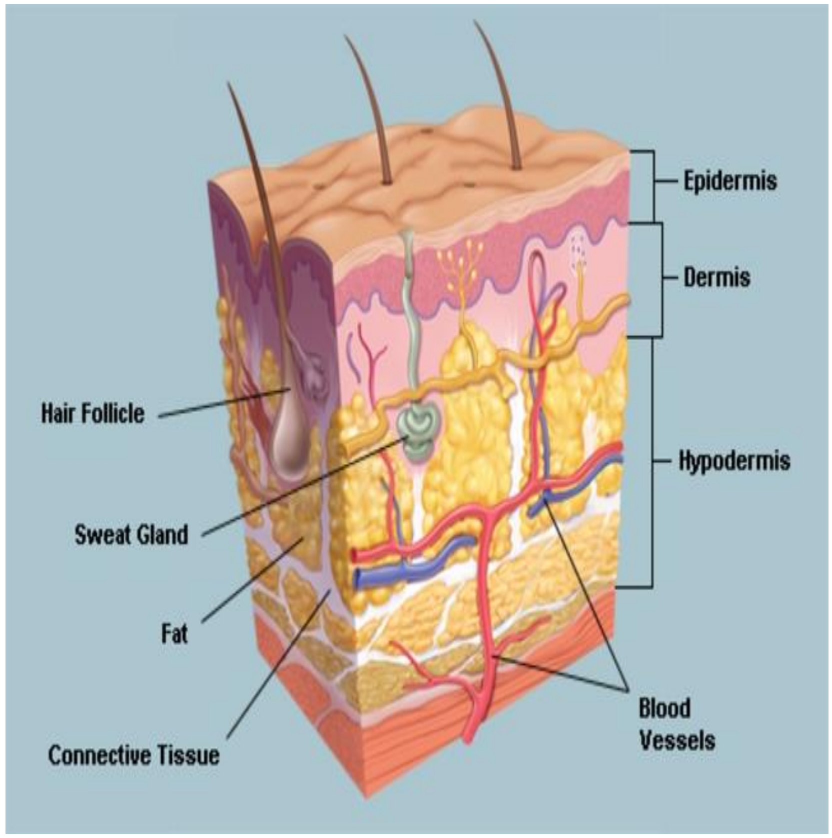

Figure 2. 1 The general human skin structure (Hoffman, 2019)

2.3. Types of skin lesion in this study

2.3.1 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 (Brendan P. Kelly, 2012). 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 corpories, Tinea capitis, tinea pedies, Tinea cruises, Pityriasis versicolor and etc (Brendan P. Kelly, 2012). Among them we included the most frequent occurred fungal infections those are Tinea capitis, tinea pedies and tinea corpories.

Tinea refers to superficial infection with one of three fungal genera called Microsporum, Epidermophyton, and Trichophyton collectively known as dermatophytes. These infections are among the most common diseases worldwide and cause serious chronic morbidity (Hay & Morris-Jones, 2012). 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 pedies, and tinea unguium. The clinical diagnosis can be unreliable because tinea infections have many mimics, which can manifest identical lesions. For example, tinea corporis can be confused with eczema, tinea capitis can be confused with alopecia areata, and onychomycosis can be confused with dystrophic toe nails from repeated low-level trauma (Mary Seaburyn, Rosenfeld, & W.Ely, 2014) .

Tinea Capitis

Tinea capitis is one of a skin disease caused by a fungus, its shape is usually round or oval. It is the commonest superficial fungal infection among primary school children (Oke, Onayemi, Olasode, Omisore, & AbimbolaOninla, 2014). It affects the skin around a head. Tinea capitis is almost exclusively a disease of childhood, and current evidence suggests that it occurs more often in children of Africa (Hay & Morris-Jones, 2012). Tinea capitis causes small red bumps and pustules on the scalp, as well as some scaling. Over time, these bumps may increase in number. The rash can also spread to cover a wider area. People who have tinea capitis may experience a localized area of scaling, itching, and pus filled bumps. They may also notice some hair loss.

Figure 2. 2 Tinea capitis

Tinea Pedies

Tinea pedies is the other disease caused by fungus and it mostly affects the leg and foot. It is a fungal infection that usually begins between the toes. It commonly occurs in people whose feet have become very sweaty while confined within tight fitting shoes. Signs and symptoms of athlete's foot include a scaly rash that usually causes itching, stinging and burning. Athlete's foot commonly starts with a red rash between the toes, typically between the fourth and fifth toe. The prevalence of tinea pedies is high particularly in people who wear occlusive footwear, such as athletes (Hay & Morris-Jones, 2012)

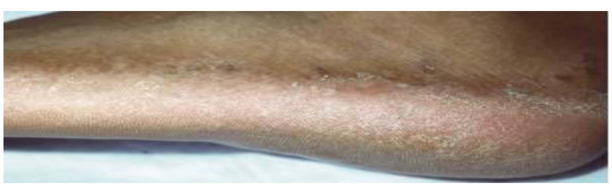

Figure 2. 3 Tinea Pedies

Tinea Corporis

This type of fungal infection is mostly affect the overall body like hand, leg, abdomen and neck. Present with apruritic erythematous rash with an active scaly palpable edge within. Tinea corporis is seen most commonly in children and young adults (Hay & Morris-Jones, 2012). It is characterized by ring shape. Signs of tinea corporis are the appearance of an itchy, red, circular rash in the shape of a ring. This rash may resemble a target or a bullseye, and it usually has raised edges. The most common infections in children are tinea corpories and tinea capitis (Mary Seaburyn, Rosenfeld, & W.Ely, 2014).

Figure 2. 4 Tinea Corporis

2.4. Digital image processing

A digital image is a representation of a two dimensional function, f(x, y), where x and y are spatial coordinates, and the amplitude at any pair of coordinates (x, y) is called the intensity or gray level of the image at that point. When x, y, and the amplitude values of f are all finite, discrete quantities, the image is called a digital image (Tigistu, 2017). Digital image is a composed of a finite number of elements, each of which has a particular location and value. These elements are referred to as picture elements (pixels). Pixel values typically represent gray levels, colors, heights, opacities etc. Common image formats include: 1 sample per point (B&W or Grayscale), 3 samples per point (Red, Green, and Blue), and 4 samples per point (Red, Green, Blue, and Alpha, a.k.a. Opacity) (Nikou, 2008).

Digital image processing is manipulation of digital images using computers. Digital image processing focuses on developing a computer system that is able to perform processing on an image (Tigistu, 2017). Digital image processing has very wide applications in medical diagnosis, remote sensing, transmission and encoding, Machine/Robot vision, pattern recognition etc. (Daniel Hailemichael, 2015). DIP has been extensively used in various (human, animal, plant) disease diagnosis approaches and assisting experts to select the right treatment. It can either be used to recognize the symptoms of a disease on the skin or even in the molecular analysis using microscope images that display the anatomy of the tissues (Petrellis, 2018)

Digital image processing focuses on two major tasks. The first is for improvement of pictorial information for human interpretation. The second is for processing of image data for storage, transmission and representation for autonomous machine perception. Some argument about where image processing ends and fields such as image analysis and computer vision start. Image Processing is a technique to enhance raw images received from cameras/sensors placed on satellites, space probes and aircrafts or pictures taken in normal day-to-day life for various applications (Chitradevi & P.Srimathi, 2014) (Nikou, 2008).

Digital Image processing can be seen in three levels. These levels are termed as low, mid, and high level processes. Low level processes involve primitive operations such as image preprocessing to reduce noise, contrast enhancement, and image sharpening. A low level process is characterized by the fact that both its inputs and outputs are images. Mid-level processing involves tasks such as segmentation. Unlike low level processing, in mid-level its inputs are images, but its outputs are attributes extracted from those images. High level processing involves making sense of a group of recognized objects like classification, tracking etc (Daniel Hailemichael, 2015).

Basically different image processing applications may follow different steps. However, the fundamental steps that every image processing applications pass through are shown in Figure 2.5.

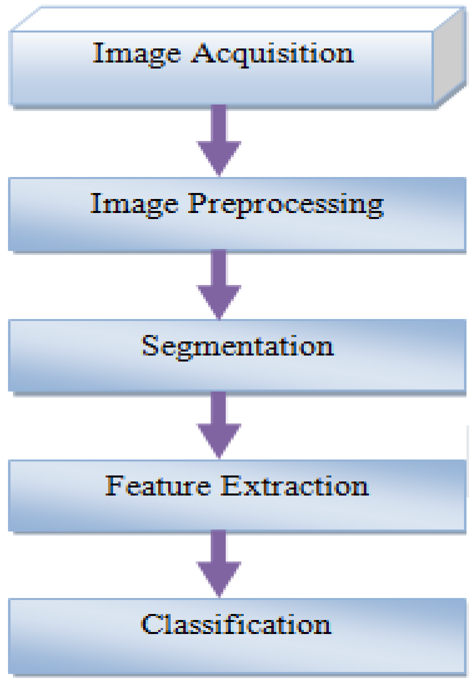

Figure 2. 5 Steps of digital image processing

2.4.1. Image Acquisition

Image acquisition is a process of gathering relevant data from different sources. It is the first step of digital image processing, before starting developing a system, first essential data must be available. There are different imaging methods of skin, some of them include: total cutaneous photography, dermoscopy, CSLM, ultra sound, MRI, OCT and multi spectral imaging (Ruela, 2015).

2.4.2. Image preprocessing

Image preprocessing is an essential step of detection in order to enhance the quality of original image by removing unrelated and surplus parts such as noise in the back ground of image. By any means the image is acquired there are artifacts to be removed and enhance the contrast for making the border detection easy to the system. This is an essential step in digital image processing because enhancing the quality of original image enhances the detection of lesion region. Thus selection of the best preprocessing techniques can greatly improve the accuracy of the CAD system. The objective of the preprocessing stage can be achieved through image enhancement and image resizes (Masood & Jumaily, 2013).

Image enhancement

Image enhancement is an important procedure to improve the visual appearance of the image and it can be categorized into image scaling, color space transformation and contrast enhancement. Image enhancement is the modification of image by changing the pixel brightness values to improve its visual impact. Images suffer from poor contrast and noise because of the limitations of imaging sub systems and illumination conditions while capturing image. Hence, it is necessary to enhance the contrast and remove the noise to increase image quality. In image processing, image enhancement is the most important stage that improves the quality (clarity) of images by removing blurring and noise, increasing contrast, and revealing image details for human viewing or machine interpretation. While image noise is a random effect that causes variation in image brightness or color information, image contrast is the difference in visual properties making an object in image distinguishable from other objects and the background. In general, image enhancement techniques can be divided into three categories (Maini & Aggarwal, 2010) (Bareilly & Vidyapith, 2013). The commonly used image enhancement techniques include point processing operation, logarithmic transform, histogram equalization and smoothing filters (Maini & Aggarwal, 2010).

Table2. 2 Image contrast enhancement techniques

Techniques	Description
Point Processing Operation	Each pixel value is replaced with a new value obtained from the old one
Logarithmic Transforms	Maps a narrow range of low gray levels into a wider range of gray levels
Histogram Equalization	If image which is predominantly dark. Then its histogram would be skewed towards the lower end of the grey scale and all the image detail is compressed into the dark end of the histogram. It become more clear image
Smoothing filters	Reduce noise or prepare the image for further processing such as segmentation
Gaussian filter	Screen noise with the high spatial frequencies and produce a smoothing effect
Median filter	It better in the sense of preserving useful details in the image. It is especially effective for removing impulse noise, which is characterized by bright and or dark high-frequency features appearing randomly over the image

2.4.3. Image segmentation

Image segmentation is defined as the process of partitioning an image into non-overlapping, constituent regions that are homogeneous with respect to some characteristic like gray level, color, texture, brightness, contrast and other (statistical) properties. First, the digital image is divided into two parts: background and foreground, where the foreground is the interesting objects and the background is the rest of the image. All the pixels in the foreground are similar with respect to a specific characteristic, such as intensity, color, or texture. The result of image segmentation is a set of segments that collectively cover the entire image.

Segmentation is the process used to locate objects and boundaries (e.g., lines or curves) in images. The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image. The segmented objects are often termed the foreground and the rest of the image is the background. It is applied to remove the unwanted part from the image and find the region of interest. For any given image, we cannot generally speak of a single, correct segmentation. Rather, the correct segmentation of the image depends strongly on the types of object or region we are interested in identifying. The central question in image segmentation is what relationship a given pixel have with respect to its neighbors and other pixels in the image in order that it be assigned to one region or another.

There are many segmentation techniques, but they can be categorized into detection of discontinuities and detection of similarities. The segmentation technique based on detection of discontinuities is the process of partitioning an image based on abrupt changes in intensity. Examples of such algorithms include all edge detection algorithms. On the contrary, detection of similarities is based on continuities. These techniques divide the entire image into sub regions depending on some similarity rules. Examples of such algorithms include thresholding, region growing etc. Whether the segmentation technique used is discontinuity or similarity, the end result of any segmentation process is a binary image (Daniel Hailemichael, 2015). Image segmentation methods are classified into threshold based, region based, and cluster based and edge based image segmentation (Heba Fathy, 2016). Image segmentation techniques are listed in table 2.2.

Table2. 3 image segmentation techniques

Methods	Description	Pros	Cons	Reference
Threshold based	Determining threshold and then the pixels are divided into groups based on that criterion. It include bi-level and multi-thresholding	✓ No need of previous information ✓ simplest method	✓ highly dependent on peaks ✓ spatial details are not considered	(Heba Fathy, 2016) (A.kaur & N.kaur, 2015) (Masood & Jumaily, 2013)
Region-based	Splitting the image into smaller then merging sub images which are adjacent and similar in some sense. It includes Statistical region merging, multi-scale region growing, and morphological flooding	✓ more immune to noise ✓ useful when it is easy to define similarity criteria	✓ expensive method in terms of time and memory	(Masood & Jumaily, 2013) (Kaganami & Beij, 2009)
cluster based	classifying pixels in an image into different clusters that exhibit similar features	✓ fuzzy uses partial membership ✓ useful for real problems	✓ determining membership function is not easy	(Masood & Jumaily, 2013) (Barghout & J.Sheynin, 2013)

MASK R-CNN	It can separate different objects in a image or a video. You give it a image, it gives you the object bounding boxes, classes and masks	✓ Simple, flexible and general approach	✓ High training time	(Masood & Jumaily, 2013)

2.4.4. Feature extraction

Features are the information extracted from images in terms of numerical values that are difficult to understand and correlate by human. If we consider the image as data the information extracted from the data is known as features. Features extracted from an image are much lower dimension than the original image. The reduction in dimensionality reduces the overheads of processing bunch of images. A good feature set contains discriminating information, which can distinguish one object from other objects. The selected set of features should be a small set whose values efficiently discriminate among patterns of different classes, but are similar for patterns within the same class. There are several image features which represent an image for classification/identification systems. Most popular among them are color, texture and shape of an image.

Color Features: are defined to a particular color space. Once the color space is specified, color feature can be extracted from the image. The most and common used color features include color histogram, color moments (CM), color coherence vector (CCV) and color correlogram, etc. The efficiency of the color feature resides in the fact that it is independent and insensitive to size, rotation and the zoom of the image (Seyyid, 2015).

Texture Features: texture is a repeating pattern of local variations in image intensity. It refers to surface characteristics and appearance of an object by the size, shape, density, arrangement, proportion of its elementary parts. It is generally believed that human visual systems use texture for recognition and interpretation. Color is usually a pixel property while texture can only be measured from a group of pixels (Dong, 2013). Based on the domain from which the texture feature is extracted, texture feature classified into two categories spatial texture feature extraction and spectral texture feature extraction (Seyyid, 2015). In spatial, texture features are extracted by computing the pixel statistics or finding the local pixel structures in original image domain, whereas in the spectral it transforms an image into frequency domain and then calculates feature from the transformed image. Among the current approaches used in image processing to describe texture, the so called statistical approach is the widely used because it produces good results with low computational costs. This method considers the distribution of gray levels and their interrelationship. The pixel values are used to construct numerical structures which are associated to the texture pattern of the image.

Shape Features: the shape of an object refers to its physical structure and profile. Shape features are mostly used for finding and matching shapes, recognizing objects or making measurement of shapes. Moment, perimeter, area and orientation are some of the characteristics used for shape feature extraction technique. The shape of an object is determined by its external boundary abstracting from other properties such as color, content and material composition, as well as from the object's other spatial properties. Shape features are commonly used in the object recognition and shape description. The shape features extraction techniques are classified as region based and contour based. The contour methods calculate the feature from the boundary and ignore its interior, while the region methods calculate the feature from the entire region.

The purpose of feature extraction is to reduce the original data set by measuring certain properties, or features, that differentiate one input pattern from another. When the pre-processing and the desired level of segmentation have been achieved, feature extraction technique is applied to the segments to obtain image features. Image features are those items which uniquely describe an image, such as size, shape, composition, location etc. Feature extraction refers to taking measurements, geometric or otherwise, of possibly segmented, meaningful regions in the image (Masood & Jumaily, 2013). The most common feature extraction methods used recently in medical image detection and classification systems is CNN.

2.4.5. Classification

Classification is the final phase in digital image processing responsible for making predictions or decisions about the extracted and selected feature information from the input. This phase is dependent on the previous phase's performance. The objective of image classification is to detect, identify and classify the features occurring in an image in terms of type of class these features represent on the field. (Masood & Jumaily, 2013).

Classification is the task of approximating a mapping function (f) from input variables (X) to discrete output variables (Y) (Catherine R. Alimboyong, 2018). In which the result is predicting the class of given data points. Classes are sometimes called as targets/ labels or categories. Classification is a two-step process, consisting training and classification step. In the training steps, a classification algorithm builds the classifier by analyzing and learning from a training dataset and their associated class labels. In the classification step the model is used to predict class labels for given data. The tasks depend on labeled datasets to learn the correlation between labels and data. It is a kind of supervised learning.

2.5. Image Classification techniques

Image classification techniques can be broadly divided into supervised and unsupervised (Kalra, Goswami, & Gupta, 2013) (Kumar, Kamble, Pawar, Patil, & Bonde, 2014). The most common classification methods used recently in medical image detection and classification systems are K-Nearest Neighbor Classifier, Support vector machine, Artificial Neural Network Based Methods, Deep Learning Methods. Classification algorithms designed by considering different assumptions (Brownlee, 2016).

2.4.6. Non deep learning Techniques

Naive Bayes Classifier is a classification technique based on Bayes Theorem. A Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature. Naive Bayes model is easy to build and particularly useful for very large data sets.

K-Nearest Neighbor is one of the classification algorithms. It takes a bunch of labeled points and uses them to learn how to label other points. To label a new point, it looks at the labeled points closest to that new point and has those neighbors vote, so the label which have most of the neighbors, is the label for the new point.

Support Vector Machines is one of the classification algorithms. SVM is a non-probabilistic binary linear classifier. The non-probabilistic aspect is its key strength. This aspect is in contrast with probabilistic classifiers such as the Naïve Bayes. That is, an SVM separates data across a decision boundary (plane) determined by only a small subset of the data (feature vectors) (Raj Bridgelall, 2017). It represents the training data as points in space separated into categories by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall. The major problem of SVM is, it used for binary class classification. However, it is possible to classify a given dataset in to more than two classes by using Kernel functions. Sigmoid, Polynomial, Radial basis function (RBF) are kernel types those enable SVM used for multi- class classification.

Artificial Neural Networks (ANN) is an information processing paradigm that is inspired by the way biological nervous systems process information. We are constantly analyzing the world around us. Without conscious effort, we make predictions about everything we see, and act upon them. When we see something, we label every object based on what we have learned in the past. Neural networks works with the same concept by adjusting the connection exist between neurons. It is composed of a large number of highly interconnected processing elements called neurons, which convert an input vector into some output.

Neural networks have the potential for solving problems in which some inputs and corresponding output values are known, but the relationship between the inputs and outputs either not well understood or difficult to translate into a mathematical function. These conditions are commonly found in tasks involving classification of agricultural and animal products, plant and animal disease identification etc. ANN is better from the above method because, it is nonlinear model that is easy to use and understand compared to statistical methods. Its non-parametric behavior allows better performance in complex problems because they don’t need background of statistic.

ANN can be classified into shallow and deep learning. A shallow learning has less number of neurons compared to Deep learning (DL) (Brownlee, 2016). Image classification has been done on both learning methods. But, there is basic difference between them. Shallow learning uses one neuron for each input which increases number of parameters. For example let us consider we have an image with 200 x 200 with 3 color channels it will cause difficulties while training and possible over fitting in the model. The problems become difficult when image size increase because the amount of weights becomes larger.

In shallow learning we lose spatial information when the image is flattened into the network. For example, if a picture of lesion appears in the top left of the image in one picture and the bottom right of another picture, the shallow learning will try to correct itself and assume that lesion will always appear in the bottom right of the picture. Hence shallow learning react differently to an input images and its shifted version and they are not translation invariant. But deep learning is translation invariant. Another difference is deep learning scale with data, whereas shallow learning converges. Shallow learning reaches a stable state at a certain level of performance when you add more examples and training data to the network. But deep learning often continues to improve as the size of the data increase (Raghavendra, Bhat, Gudigar, & Achary, 2018).

Another main difference is shallow learning don’t extract feature by themselves. They need feature extractor. Feature extractors are problem dependent and must be rewritten for each new dataset. The efficiency of the extractor depends on the assumption of a person who designs the extractor, because we simply cannot conceive every possible feature that will be useful (Catherine R. Alimboyong, 2018). We choose deep learning for our research work to extract deep features and classification SVM instead of Softmax.

2.4.7. Deep learning Techniques

Deep Learning is concerned with algorithms inspired by the structure and function of the brain called artificial neural networks. The performance of this type of model improves by training with more examples and by increasing its depth or representational capacity. In addition to scalability, another often cited benefit of deep learning models is their ability to perform automatic feature extraction from raw data (Brownlee, 2016).

There are many types of deep learning used for variety of tasks in Artificial intelligence (Razzak, Naz, & Zaib, 2018) like deep neural network, deep belief network, recurrent neural networks, Deep Boltzmann Machine, restricted Boltzmann machines and Convolutional neural network.

Deep neural network is neural network having more than two layers. It can model complex non-linear relationships. It is used for classification as well for regression. Deep neural networks use sophisticated mathematical modeling to process data in complex ways.

Deep belief network is unsupervised probabilistic deep learning algorithm. It can be defined as a stack of restricted Boltzmann machines (RBM), in which each RBM layer communicates with both the previous and subsequent layers. The nodes are connected to each other across layers, but no two nodes of the same layer are linked. That is, there is no intra-layer communication. The idea behind a DBN is that it is possible to stack more RBMs on top of each other by building a deep network, which presents the ability to extract a hierarchal representation of the input at multiple level of abstraction.

Deep Boltzmann Machine (DBM) is unsupervised, probabilistic, generative model with entirely undirected connections between different layers. It contains visible units and multiple layers of hidden units. Like RBM, no intra layer connection exists in DBM. Connections exists only between units of the neighboring layers network of symmetrically connected stochastic binary units DBM can be organized as bipartite graph with odd layers on one side and even layers on one side units within the layers are independent of each other but are dependent on neighboring layers. Deep Boltzmann Machine (DBM) have entirely undirected connections

Recurrent neural networks (RNNs) are a type of artificial neural network that are able to recognize and predict sequences of data such as text, genomes, handwriting, spoken word, or numerical time series data. They have loops that allow a consistent flow of information and can work on sequences of arbitrary lengths. The RNNs attempts to address the necessity of understanding data in sequences. Recurrent nets differ from feed forward nets because they include a feedback loop, whereby output from step n-1 is feedback to the net to affect the outcome of step n, and so forth for each subsequent step.

Convolutional neural network (CNN) is one of the main learning mechanisms to do images recognition, images classifications, objects detections etc. including automatic feature extraction. It is based on learning levels of representations. The higher lever concepts are defined from lower-lever ones and the same lower lever concepts can help to define many higher lever concepts. It learns multiple levels of representation and abstraction which helps to understand dataset such as images, audio and text. It is advantageous of simple structure, less training parameters because of shared weights and adaptability (Liu, Fang, Zhao, & Wan, 2017). CNN is now the go-to model on every image related problem. In terms of accuracy they blow the competition out of the water, it is also successfully applied to recommender systems, natural language processing and more, the main advantage of CNN compared to its predecessors is that it automatically detects the important features without any human regulation, CNN is also computationally efficient to learn feature.

To choose which network fit for our problem, we must consider the type of problems and the data we use. Each deep network has its own characteristics as shown in Table 2.3. Our task is classifying the input image using good feature extractor, so it is classification and feature extractor problem and the data we have is labeled and its supervised learning. The data types are image which are represented in pixel. One pixel is dependent on the value from its 8 neighbors, the feature is spatial feature. We choose CNN for feature extraction with SVM for classification.

2.6. Component of Convolutional neural network

Even if there are different architectures the general overview of CNN contains convolution layer, pooling layer, and finally fully connected layer for classification of sub contents. The details of mathematical operation are as follows (Vedaldi & Lenc, 2015). CNN is a network that uses a convolution mathematical operation instead of general matrix multiplication.

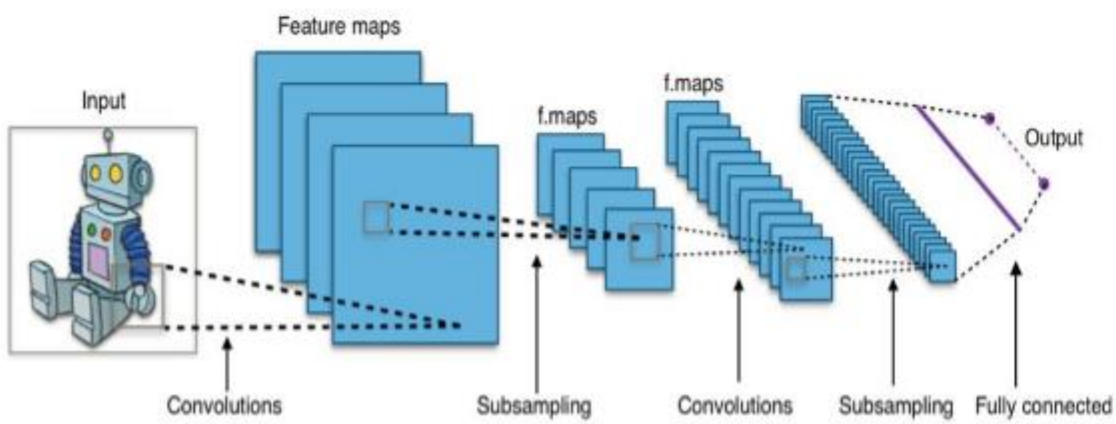

Figure 2. 6 Basic convolutional neural network architecture (Antonio & Sujit, 2017)

Convolution layer

The first layer of CNN is convolution layer. Convolution is an operation on two functions of real valued argument. For a given input image, “I” and convolving kernel “K” size of “mxn”, the convolution operation is written as Equation 2.1. The aim of CNN is to learn very deep features of the image via convolution operation, so that the model is capable of identifying the invariant features from raw image data.

$$S(i,j) = (I * K)(i,j) \sum_{m} \sum_{n} I(i-m, j-n) K(m,n) \qquad (2.1)$$

The input of CNN is a tensor, for the proposed problem the model takes an input of 3 order tensor (an image with 3 channels R, G, B) and size of width and length of each tensor “W” and “H”, then the model runs layer by layer until the output layer is reached. CNN is a sequence of layers suppose the first layer receives input image I ℰ R ^{HXWXDXN} where D is the number of channel or depth of image and N is the size of mini batch used for accelerating the training by dividing spatial input. For each element of "I" there is an index set {i,j,d,n} where 0≤i≤H, 0≤j≤W,0≤d≤D and 0<n≤N the output of lth layer Y which is an input for (l+1)th layer is computed using Equation 2.2 (Vedaldi & Lenc, 2015).

$$Y_{i+1,j+1,d} = \sum_{i=0}^{W}\sum_{j=0}^{H}\sum_{d}^{D} f_{i,j,d,n} * X_{i(1)} + i, j(l) + j, d(l) \qquad (2.2)$$

where "f" is the kernel function used for generating weights to be adjusted on the training and i, j, d of "l" are values of index in the "lth" layer or input layer. When applying convolution in actual training there are other hyper parameters selected manually additional to kernel. The hyper parameters are: output depth (depth of feature maps), stride (how far a filtering curve "f" is moving) and padding (for controlling the spatial size of the input volume) (Vedaldi & Lenc, 2015).

The result of convolution operation "Y" Equation 2.2 is called a feature map which is convolution of kernel and input. Convolution is needed to amplify different edges of the input image depending on the type of bias used. If V feature maps are gained by convolving a single channel of input X detection or object recognition is by combining all the feature maps from different channel.

Pooling layer

The second layer is pooling layer or sub sampling layer that comes after convolution layer for reducing the spatial size. The data representation is progressively reduced over the network and it helps to control over fitting. There is no parameter learning in this layer, it results a new size spatial input for the next convolution layer without changing the depth or feature maps and preserving the location of edges. Max pooling and average pooling are the most common down sampling methods (Patterson & Gibson, 2015). In this work we used the Max pooling .The output of pooling layer for a new filter size of HXD is in Equation 2.3, the output returns the maximum in the selected neighborhood for max pooling.

$$Y_{i,j,d} = Max\{X_{i,j,d}\} \; where, 0 \leq i \leq H, 0 \leq j \leq W, d \qquad (2.3)$$

Adding pooling layers to the CNN architecture helps to minimize cost of computation by reducing the number of parameters to be learned for the next convolution layer. Pooling is also helpful for handling inputs of varying size.

Activation Layer

The third layer is activation layer applied after the convolution layer. This layer is used to increase the non-linearity of CNN. Image contains semantic information which is nonlinear pixel inputs the output also must be highly nonlinear to map the input and output of CNN (Vedaldi & Lenc, 2015).

Fully Connected layer

The forth layer of CNN general architecture is fully connected layer used to compute the class of prediction. This layer is applied at the end of deep CNN model. The feature output matrix is converted into a vector form and elements fully connected to each other. For each elements of the input to the fully connected layer the category of its class is computed using Softmax activation function of Equation 2.4. Image pixel values in the form of vector are input for the function whereas output is the probability distribution of an item over P probabilities.

$$\alpha\left(Y_i = \frac{e^{Y_i}}{\sum_{j=1}^{p} e^{Y_j}}\right) \tag{2.4}$$

Where p is the number of prediction class for Y input with index of "I"

The above explanation is the forward run of CNN architecture the backward run computes the prediction loss, which helps to adjust the weights for the right prediction e.g. cross entropy loss used for multi class classification problems (Patterson & Gibson, 2015). Cross entropy Equation 2.5 compares the actual distribution and predicted distribution i.e. ground truth comparison to the predicted value.

$$H(Y, Y') = \sum_i Y_i \log\left(\frac{1}{Y_i'}\right) \tag{2.5}$$

Activation Function

Activation functions 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. It serves to introduce non linearity in the modeling capabilities of the network. A neural network without Activation function would simply be a linear regression Model, which has limited power and does not perform well. There are three types of Activation Functions binary, linear and nonlinear function.

Binary Function is a threshold based activation function. If the input value is above or below a certain threshold, the neuron is activated and sends exactly the same signal to the next layer. It does not allow multi value outputs, for example, it cannot support classifying the inputs into one of several categories.

Linear Activation Function is activation function takes the form A = cx. It takes the inputs, multiplied by the weights for each neuron, and creates an output signal proportional to the input. Linear function is better than binary function because it allows multiple outputs, not just yes and no. All layers of the neural network collapse into one with linear activation functions, no matter how many layers in the neural network. The last layer will be a linear function of the first layer because a linear combination of linear functions is still a linear function. So a linear activation function turns the neural network into just one layer. Also using back propagation to train the model because the derivative of the function is a constant, and has no relation to the input. So it's not possible to go back and understand which weights in the input neurons can provide a better prediction.

Non-Linear Activation Functions allow the model to create complex mappings between the network inputs and outputs. It is essential for learning and modeling complex data, such as images, video, audio, and data sets which are non-linear or have high dimensionality. They allow back propagation because they have a derivative function which is related to the inputs. They allow stacking of multiple layers of neurons to create a deep neural network. Multiple hidden layers of neurons are needed to learn complex data sets with high levels of accuracy. We need a neural network model to learn and represent a complex function which maps inputs to outputs that is why we need non linearity.

Some of the most common nonlinear Activation function is listed below: Sigmoid is activation function of form

$$f(x) = \frac{1}{1} + \exp(-x) \qquad (2.6)$$

It’s recommended to be used only on the output layer so that we can easily interpret the output as probabilities since it has restricted output between 0 and 1. It is computationally expensive.

Tanh is activation function and it is computationally expensive. The mathematical form is

$$f(x) = \frac{1 - \exp(-2x)}{1 + \exp(-2x)} \qquad (2.7)$$

RELU is another activation function. Compared to other activation function, it is computationally expensive. Softmax is activation function which used for classification with network; RBF is a kernel which used to make classifier able to classify the input in to multi-classes. There is no definitive guide for which activation function works best on specific problems. It’s a trial and error process by trying different set of functions and sees which one works best on the problem at hand. In this thesis, we show the comparison result of Relu and Tanh. The mathematical formula expressed in equation 2.8

$$f(x) = \max(0, x) \qquad (2.8)$$

Padding

In order to preserve the same dimensionality, possible to appling padding to border the input with zeros. (Arden, 2020). Padding is a border zero values which are placed outside the matrix of the image. In a convent having many layers, this effect of decrease the output can be countered by zero padding the input. We can have the result of canceling dimensional discounts and preserving the input dimension at the output. So, padding may be necessary in order to prevent outputs from becoming too reduced. In our research we applied Same Padding: In this case, we add ‘p’ padding layers such that the output image has the same dimensions as the input image.

Dropout Layer

Dropout is a technique used to prevent a model from overfitting. The in 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. This process applied during the training phase and the validation phase (Arivazhagan & Vineth, 2018).

2.7. Common Architectures in Convolutional Neural Networks

Different CNN architectures have proposed from 2012 like VGG16, VGG19, AlexNet, Google net, ResNet, Densenet and so on. The technologies make researchers able to improvement classification across natural images by taking advantage of the quite luxurious classification method (Preetom & Chowdhury, 2018).

AlexNet

AlexNet is one of CNN architecture, which has 60 million parameters and 650,000 neurons. AlexNet consists 8layers including five convolutional layers and 3 fully connected layers with the final 1000-way Softmax which produces a distribution over the 1000 class labels. Relu was applied at every convolutional and fully connected layer. The network's input is 150,528 (224 x 224 x 3) dimensional. The 1st convolutional layer filters the input image with 96 kernels of size 11 x 11 x 3 with a stride of 4 pixels. The 2nd convolutional layer takes as input the pooled output of the first convolutional layer and filters it with 256 kernels of size 5 x 5 x 48. The 3rd, 4th, and 5th convolutional layers are connected to one another without any pooling or normalization layers. The fully connected layers have 4096 neurons each. 11*11 filter size leads higher memory consumption and they do not used batch normalization. AlexNet won the competition in the ILSVRC-2012 to classify 1.2 million images into 1000 different classes by achieving a winning top-5 test error rate of 15.3%, compared to 26.2% achieved by the second-best entry (Krizhevsky, E.Hinton, & Sutskever, 2012).

VGG

VGG is another CNN architecture ,which consists of 16 learned layers including thirteen convolutional layers with filter size of 3 x 3, five pooling layers which follow some of the convolutional layers and three fully connected layers with the final 1000-way Softmax which produces a distribution over the 1000 class labels. All learned layers were applied with ReLU nonlinearity. Their main contribution was a detailed evaluation of networks of increasing depth using architecture with very small (3 x 3) convolution filters, which are convolved with the input at every pixel with stride size of 1. VGG takes the second place in the contest of the ILSVRC-2014 to classify 1.2 million images into 1000 different classes by achieving a top-5 test error rate of 7.3 % (Karen & Andrew, 2014). They did not apply batch normalization to accelerating the training time and stable the model in order to learn features by updating the parameter.

GoogleNet

GoogleNet is a CNN model, which contains 22 layers with no fully connected layers and only 5 million trainable parameters and decreases 12 times than AlexNet. For the convolutional operation the model used 1*1, 3*3, 5*5 filter size for pooling 3*3 with stride 2 and 224*224-pixel value used to rescale or normalized images, it achieved a top-5 error rate of 6.67% in the ILSVRC-2014 to classify 1.2 million images into 1000 different classes(Sakshi, Kumar, Mishra, & Asopa, 2018). But they applied 5*5 filter size, this need higher computation time so in our case the architecture not fit with our dataset related to characteristic features this also lead overfitting and underfitting problems.

ResNet

ResNet is the deepest CNN model than VGG, AlexNet, and GoogleNet, which consists of 152 layers. ResNet won the 1st place in the ILSVRC-2015 classification competition to classify 1.2 million images into 1000 different classes by achieving a winning top-5 test error rate of 3.57%. It also won the 1st places on the tasks of ImageNet detection, ImageNet localization, COCO detection, and COCO segmentation (Kaiming, Xiangyu, Shaoqing, & Jian, 2015).

2.8. Related works

So far, several researchers developed there model for diagnosis, detection, and classification of skin disease. Several methods have been applied by many researchers in order to identify and recognize human skin diseases from images. In section 2.8.1 and 2.8.2 review of some approaches used by numerous researchers are explained and we try shows the gap we want to fill.

2.8.1. Detection and classification of skin disease using non deep learning approach

In (Li-sheng, Quan, & Tao, 2018), the authors proposed a method for skin disease detection and classification. They used GLCM for feature extraction and SVM for classification. However, their method recognized three types of diseases, namely, herpes, dermatitis, and psoriasis. Their evaluation result is 85%, 90%, 95% of the recognition rate respectively. In this thesis we used CNN for feature extraction and SVM for classification, since CNN is the state of the art today to extract deep features.

Diagnosing different types of skin diseases were done by researchers 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%, 34% for Eczema, Psoriasis, Acne, Skin Cancer, Scabies and Seborrheic Dermatitis respectively (Carl Louie Aruta, Calaguas, Gameng, & Prudentino, 2015). From their result, only Eczema and Acne perform high accuracy. However, the rest 4 types of diseases less perform than the two skin diseases.

Segmentation and classification of skin lesions was proposed (Sumithra, Mahamad, & Guruc, 2015). Authors used SVM and KNN for classification purposes. Melanoma, Bullae, Seborrheic keratosis, Shingles, Squamous cell were detected and classified using this method. They compare the two algorithms by using individually and as fusion. Their performance result shows SVM-KNN fusion outperforms the individual classifiers. The overall result of the proposed method is 46.71% and 34% of F-measure using SVM and k-NN classifier respectively and with 61% of F-measure for a fusion of SVM and k-NN.

From this, the performance result is low, because they used a collection of datasets from internet resources and they didn't use a good feature extraction method.

2.8.2. Detection and classification of skin disease using deep learning approach

Skin-disease detection system was implemented with the Tensor flow library and Python. Researchers used CNN for feature extraction. This web-based system used to classifying and Diagnosis Atopic dermatitis, Acne vulgaris and Scabies (Sam, Philip, Nartey, & Nti, 2019). Their system was able to detect widespread skin diseases (atopic dermatitis, acne vulgaris, and scabies) in Ghana. Disease identification accuracy of 88% for atopic dermatitis, 85% for acne vulgaris, and 84.7% for scabies. The overall classification accuracy is 88%.

Skin diseases are hazardous and often contagious, especially melanoma, eczema, and impetigo. The fundamental problem is only an expert dermatologist is able to detect and classify such disease. Sometimes, the doctors also fail to correctly classify the disease and hence provide inappropriate medications to the patient. Authors propose a skin disease detection method. Their system is mobile-based so it can be used even in remote areas. They also compared Support Vector Machine (SVM) and Convolutional Neural Networks (CNN). Based on their result CNN performs better classier performance than SVM (Mrs. Jayashree, Aishwarya, Harsha, & Nisha, 2019). However, their system classifies 3 types of disease melanoma, eczema, and impetigo. Skin diseases like tinea corpories, tinea pedies, and tinea capitis are not recognized in this paper.

A Mobile net CNN model also used by authors to propose a smart phone based skin disease classification system. This system recognized 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 (Velasco, et al., 2019). Even though they include tinea corpories, they didn't consider other fungal skin disease like tinea pedies, and tinea capitis.

(Kamulegeya, et al., 2019), were tried to test one artificial intelligent application, which is developed for skin disease identification for dermatology. They used diversity black skin color images as a data set they have got 17% accuracy. Finally they conclude as the designed AI is poor for fungal infection like tinea (capitis and corpories). To fill this gap we develop a method by including those fungal skin disease images.

2.9. Summery

Generally, this chapter explained the general science of Human skin structure, digital image processing and different techniques of image preprocess, segmentation, feature extraction, and classification. Some of the very recent research on medical image processing also reviewed. The reviewed literature is from different medical image processing studies but most of them are on human skin lesion identification. Addressing all the related works is not as such easy but the notable literatures are presented in this chapter. As it is the main aim of this work most of presented literatures are from support of automatic skin lesion detection and diagnosis. From the reviewed literatures, we tried to find the gap to fill in this thesis work.

CHAPTER THREE: METHODOLOGY

3.1. Introduction

In this work, we explain the design and system architecture of the proposed work. The steps in image processes such as skin lesion preprocessing, segmentation, feature extraction, model training and classification that are used in the system architecture are also explained in detail. The model has two components such as image processing and deep learning. It has three phases training, validation, and testing. In this work, our objective is to address the identification problem of skin disease using convolutional neural network.

3.2. The architecture of the proposed system

The architecture of the proposed work is depicted in the figure 3.1. Its major components include image preprocessing, image segmentation, feature extraction, skin lesion classification. In image processing part we used image size rescale, resize and filtering techniques for noise removal. Secondly, image segmentation applied to find region of interest and to separate the foreground and background of image. Thirdly feature extraction applied by the developed CNN model. Deep learning techniques used for a feature extraction by automatically applying convolutional layers, pooling layer and other hyper parameter like batch normalization, activation function, padding, stride, learning rate and optimization function.. The proposed model extract feature and used for classification of human skin disease. Finally we are classifying the image into predefined class using SVM, specifically RBF kernel function.

As shown in figure 3.1, the first phase of the model is the training phase. In the training phase, we first acquire human skin disease images. Each image has three channels Red, Green, and Blue (RGB). Then, pre-process the acquired image to enhance its quality, rescales, and reduces nose. After preprocessing image data augmentation applied to increase the train dataset and to overcome overfitting problem. Secondly image segmentation technique applied to locate the region of interest that may contain the characteristic features of the disease. Thirdly train the model using training datasets. The model has a convolutional layer, pooling layer, and fully connected layer, to extract relevant features. Finally, we apply the SVM for classification function to classify into a predefined set of classes. The second phase is validation, its similar with the training phase by Appling preprocessing to feature learning and additional batch normalization used to stable the model to learn feature by update parameters, dropout and activation function the final phase is testing the model by test image dataset the convolution applied similar the validation phase.

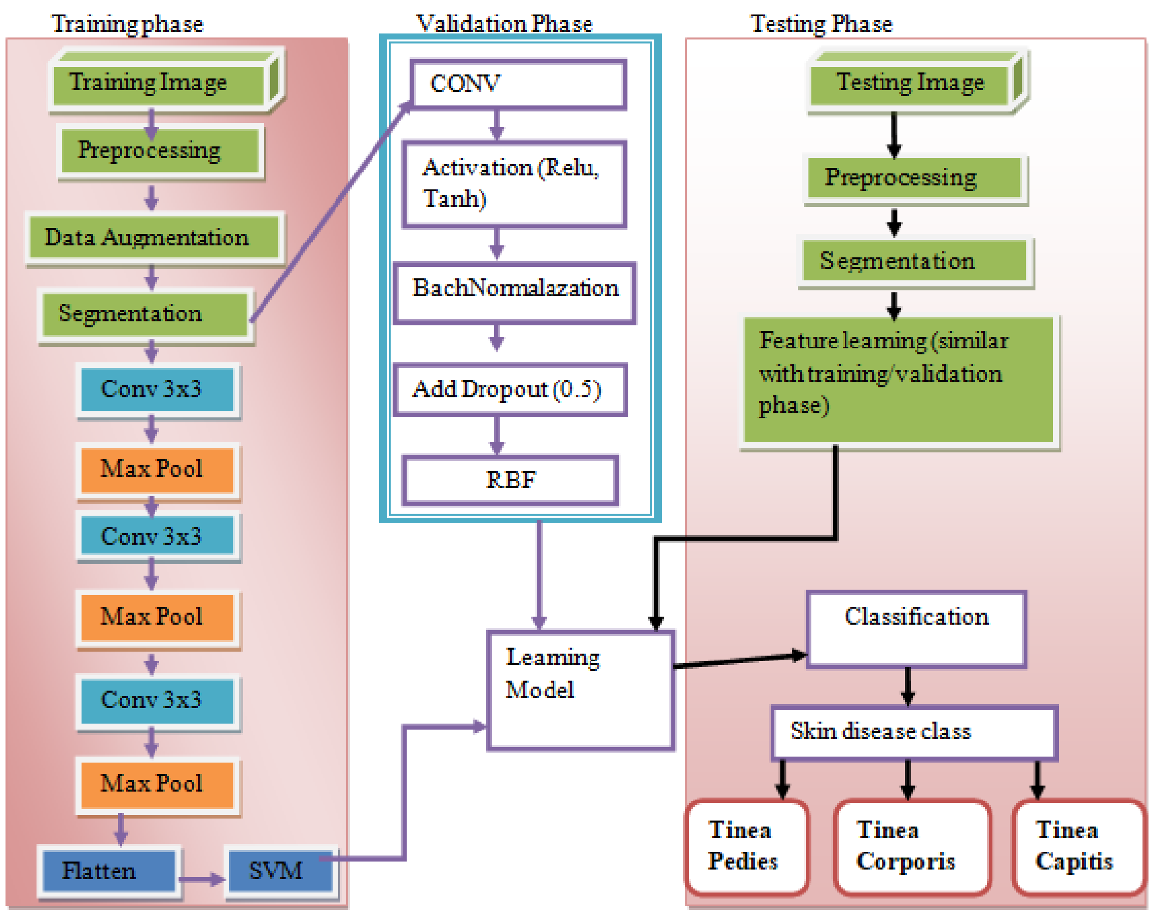

Figure 3. 1 Proposed human skin disease identification architecture

3.2.1. Image acquisition

We got the datasets from patients at Bahr Dar Tibebe Giyon Hospital and Bahr Dar Felege Hiywet General Specialized Hospital. We select Bahir Dar hospitals, since the selected fungal disease are common cases in these two hospitals. To maximize the number of collected datasets we used additional datasets from public repositories.

3.2.2. Image Preprocess

Image is acquired there are artifacts to be removed and enhance the contrast for making the border detection easy to the system. This is an essential step because enhancing the quality of the original image enhances the detection of lesion region. For example, in our data set hair is a noise to be removed. We used the following image preprocess techniques.

Image resizes: This step is necessary when you need to increase or decrease the total pixel for the standardization of image. Moreover, resizing the image reduces processing time and computational cost. The collected images have different size, width in ranging between 100 and 2000 pixels and length 100 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 method and the large pixel's size consumes too much computational cost and time. We compare three types of image size by resized the dataset in to 200X200, 224X224 and 300X300 pixels.

As part of skin lesion preprocessing, we resized each skin lesion image, converted RGB image to Grayscale and balance the image intensity using histogram equalization. Noise minimized is also performed using median filter.

Input: sample image
Output: resized Image(200,200)
Begin:
image = sample image image=resize(image,200x200)
Return resized Image
END

Algorithm3- 1 Image processing of resized image

Convert RGB in to Grayscale: since RGB is 3 channel image color format, it must change in to 1 channel to apply threshold segmentation. In this work, after resize image the color format changes from RGB to Grayscale.

Input: resized Image
Output: Grayscale
Begin:

If image data format == channels first:
Channel dimension = 1; // the position of the depth is at the first coordinate
InputShape = (depth, height, width);
Else:
Channel dimension = -1; // the depth will be located next to height and width
Grayscale = (height, width, depth);

Return Grayscale
END

Algorithm3- 2 Image shape ordering

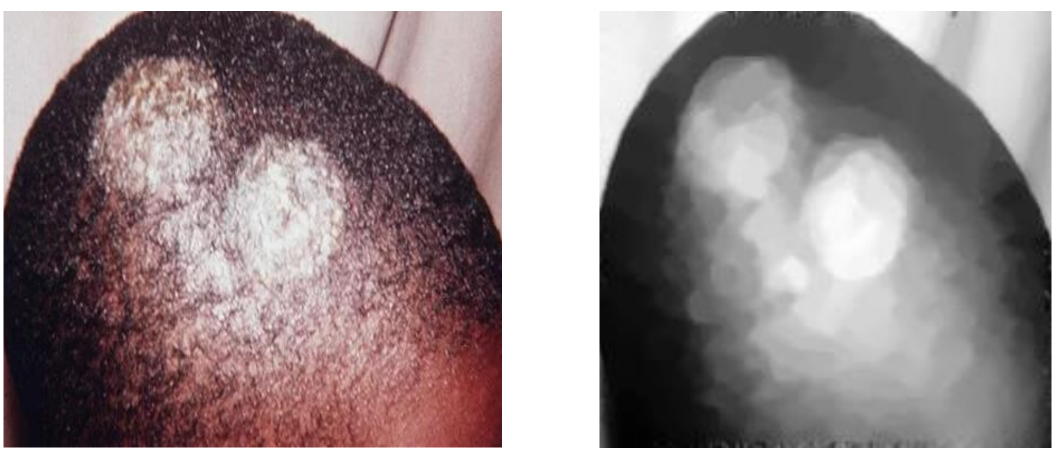

Figure 3. 2 Original RGB image and the Grayscale image respectively

Histogram Equalization: it is a technique used to balance the intensity of image. It is a common technique for enhancing the appearance of images. Suppose we have an image which is predominantly dark. Then its histogram would be skewed towards the lower end of the grey scale and all the image detail is compressed into the dark end of the histogram. If we could `stretch out' the grey levels at the dark end to produce amore uniformly distributed histogram then the image would become much clearer.

Input: Grayscale
Output: balanced intensity image
Begin:
Histogram=Grayscale Histogram = equalizeHist(Histogram)
Return Histogram
END

Algorithm3- 3 Balance the intensity of image

Median filtering: In this thesis minimizing the effect of hair is addressed through the use median filter. It is another smoothing filter that is used to reduce noise in an image, somehow like a mean filter. However, it performs better than a mean filter in the sense of preserving useful details in the image. It is especially effective for removing impulse noise, which is characterized by bright and or dark high-frequency features appearing randomly over the image. Statistically, impulse noise falls well outside the peak of the distribution of any given pixel neighborhood, so the median is well suited to learn where impulse noise is not present, and hence to remove it by exclusion. The median filter is demonstrably better than another algorithms at removing noise because it preserves edges for a given, fixed window size. So, median filtering is very widely used in digital image processing (Chande & Gupta, 2013)

Input: histogram_image
Output: Filter
Begin:

Image=histogram_image
Set height=histogram_image.shape[0]
Set wedith=histogram_image.shape[1]
Initialize x, y
for i in np.arange(x,height-x):
for j in np.arange(x,wedith-x):

Set new_data=[]// create memory
for k in np.arange(-x,y):
for l in np.arange(-x,y):
Set a=histogram_image.item(i+k,j+l)
new_data.append(a)
new_data.sort()
Calculate midpoint z
Set m the midpoint value from array of new_data[z]
histogram_image.itemset((i,j),m)

Return histogram_image
END

Algorithm3- 4 Noise removal using median filter

3.2.3. Data augmentation

Data augmentation is one of regularization method to prevent machine learning models from over fitting (Adrian, 2017). In this thesis data augmentation is applied after preprocessed the collected images. We have applied random transformation (such as zooming, rotation, shearing, and resizing) to our dataset to balance the insufficiency of our data.

Input: preprocessed image
Output: Augmented images
Begin:
Initialize rotation range (RR), width shift range (WSR), height shift range (HSR), shear range (SR), zoom range (ZR)
Set horizontal flip (HF) to true Aug_image = ImageDataGenerator (RR, WSR, HSR, SR, ZR, HF) Return Aug_image
End

Algorithm3- 5 Algorithm of augment image

3.2.4. Image Segmentation

Segmentation is dividing an image into different parts, each part representing homogeneous regions that share the same features such as intensity, color, texture. Image segmentation is a very important task in imaging science because the segmentation result can affect all the subsequent steps of automatic image analysis. We will use region based and threshold segmentation techniques. Now threshold segmentation is apply on the processed dataset. This method partition an input image into pixels of two or more values through comparison of pixel values with the predefined threshold value T individually (Kamdi & R.K.Krishna, 2012) .It transforms an input image to segmented binary image. In this work, we select Global threshold based and region based segmentation techniques.

Global thresholding: This threshold algorithm is applicable when the intensity distribution of objects and background pixels are sufficiently distinct. In the global threshold, a single threshold value is used in the whole image. If G(x, y) is a threshold version of f(x, y) at some global threshold T, then the G(x,y) will be given as in equation

$$G(x,y) = \{1, if\ f(x,y) \geq T\ otherwise\ 0 \qquad (3.1)$$

There are a number of global thresholding techniques. Otsu thresholding, optimal thresholding, histogram analysis thresholding, iterative thresholding, maximum correlation thresholding, clustering thresholding, Multispectral and Multi-thresholding techniques are among the global tresholding techniques (K.Bhargavi & Jyothi, 2014)

Figure 3. 3 the original and the segmented image respectively

The region based segmentation: Region based technique is also known as “similarity based segmentation” used to find out region directly. (Yogamangalam & Karthikeyan, 2013) It partition an image into uniform sub-regions based on some properties such as texture, color, intensity etc. Pixels belong to same intensity characteristics and closed to each other can be group together and assumed to be in same object. Region contains more information because it covers more pixels than edges. Watershed algorithm, region split and merge algorithm and region growing algorithm are the some commonly used methods of region based technique. It is a simple and robust method used to partition image into uniform regions correctly and give original image with clear edges using spatial information but it use high computation power. (Elayaraja & Suganth, 2014). Finding region of interest is the major task to segment the relevant feature only. Steps of to find region of interest are listed below.

In the first step, preprocessed images converted into array that Keras or TensorFlow works. Secondly, get the dimensions (hight=row and width=column) of the image. Thirdly select the initial kernel points; kernel point selection depends on the properties regions we want to segment like intensity value. Next assign x and y integers value such that their sum is equal=255 like x=100, and y=155. (x , y). Then assign i, j to cover the point x and y using this iteration, loop continues until we cover the segmented region that we want to segment and the loop continues until the length becomes zero, while the length is reduced by 8 units, each time length and width are reduced by 8 since the size of array x and y is eight and print value of pixel then update the x and y coordinate of the point p, where we want to segment.

Finally check whether the point p is inside the region points (an area where we want to segment). if p is inside the region points assign the image region (we want to segment) a value of 1 Otherwise, assign the image region (that cannot be segmented) a value of 0


Figure 3. 4 The original and the segmented image using region based techniques respectively


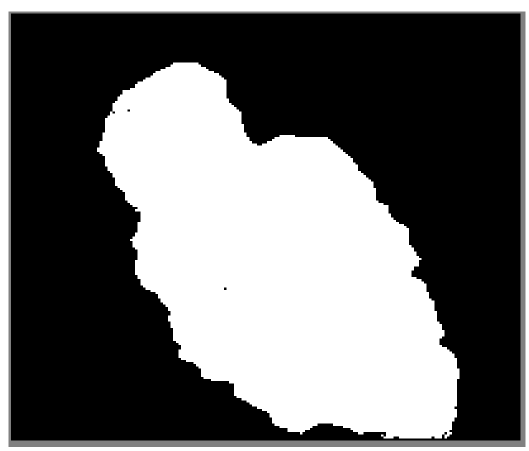

Figure 3. 5The original and the segmented image using integration techniques respectively

From the experiment result as shown in the figure 3.3, figure 3.4 and figure 3.5 the more clearly segmentation techniques is the integration one. So in this thesis, we used hybrid segmentation (region based & threshold)

3.2.4. Feature learning

The feature extraction is the process to represent a raw image in a reduced form to facilitate decision making such as pattern detection, classification or recognition. Finding and extracting reliable and discriminative features is always a crucial step to complete the task of image recognition and computer vision. Extracting as much information as possible from the available data sets is crucial to creating an effective solution. To extract deep features CNN is better (Ji & Yuyu, 2019).

In Convolutional network the neurons are arranged in 3 dimensions width, height and depth. The Convolutional layers consist of a set of learnable filters. Every filter is arranged with width and height, and extends through the full depth of the input volume.

As shown in the figure 3.1 the first layer in the proposed model is convolution layer. The first Convolutional layer accepts 200x200x3 image and have filter with 3x3x3 where the last 3 represent the depth of the filter which is similar to the input image. As we slide the filter over the width and height of the input volume we will produce a 2 dimensional activation map that gives the responses of that filter at every spatial position. Then we will stack these activation maps along the depth dimension and produce the output volume. In the first Convolutional layer 32 filters are applied which result in 200x200x32 activation map. The max polling layer will perform down sampling operation along the spatial dimensions (width, height), resulting in volume 100x100x32.

The second Convolution layer accept the result of the first convolutional layer 100x100x32 and apply its 64 filters to the input which result 100x100x64 activation map. Max poling follows to perform down sampling and result 50x50x64 volume. Its function is to progressively reduce the spatial size of the representation to reduce the amount of parameters and computation in the network, and also control over fitting.

The third convolutional layer applies 128 filters. It accepts the result of the second max pooling layer 50x50x64 and apply its 128 filters to the input which result 50x50x128 activation map. Max poling follows to perform down sampling and result 25x25x128 volume.

The final fully connected layer compute the class scores, resulting in volume of size 1x1x3, where the 3 numbers correspond to a class score, among the 3 categories of our dataset. We observe adding more layers do not improve the performance on our dataset. It leads to over fitting, increased memory consumption and computation time. The removal of one layer from the model result poor performance because the model will not generalize enough with less number of layer in the model.

Activation function is also added on each Convolutional layer. There are different activation functions. There is no definitive guide for which activation function works best on specific problems. It's a trial and error process by trying different set of functions and sees which one works best on the problem at hand. From a number of activation function comparison was done between Relu and Tanh.

Bach normalization is added on each convolutional layer. It used to increase the stability of a neural network. Batch normalization normalizes the output of a previous activation layer by subtracting the batch mean and dividing by the batch standard deviation.

3.2.5. Classification

Classification is the final task of this work. As we discussed in the literatures, there are different algorithms used for classification purpose. In this thesis work we select SVM for this purpose.

Support vector Machine and Kernel

Support Vector Machine (SVM) is a linear model for classification and regression problems. The algorithm creates a line or a hyper plane which separates the data into classes (Pupale, 2018). SVM is a popular and powerful supervised learning algorithm is for binary class (Zoltan, 2018). To make suitable for non linear or more than two class non linear activation function must be applied. RBF is a kernel function which, used for classification instead of Softmax. (Macedo, 2016) Shows a better performance using CNN with SVM than the individual CNN. As described in section 3.2 Figure 3.1 in the fully connected layer we used SVM for classification and RBF non linear activation function.

Classification is done by using the knowledge from the learning model, which is constructed by using training and validation phase. Training dataset is used in both training and validation phase. Hence, the features extracted by CNN used as input for SVM. By using the knowledge from the learning model, we can classify each image (in the testing dataset) into a specific or predefined class (Tinea Capitis, Tinea pedies, and Tinea Corpories).

3.3. Summery

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 preprocessing, segmentation, feature learning and classification are described clearly. The model has three phases: training phase, validation phase, and testing phase. The training and validation phase use training dataset to form the learning model. The testing phase, uses testing dataset and pass via the same procedures as in the training and validation phases.

CHAPTER FOUR: RESULTS AND DISCUSSION

4.1. Introduction

In this section the dataset used in this thesis work, tools used in the implementation part and parameters used in the model are described in details. We also presented the evaluation result and discussion about the extracted result.

4.2. Dataset preparation

We take image data sets from Bahr Dar Tibebe Giyon specialty Hospital and Bahr Dar Felege Hiywet General Specialized Hospital. In addition to this, we also used images in repository MedicineNet.com. Since the pandemic of corona virus, it was difficult for me to collect as much data set up to this date. After collecting some data from hospitals and repository data augmentation apply on the collected data to minimize over fitting problem. From 1196 image data 80% of the dataset used for training and 20% used for testing.

Table 4. 1 Number of collected and augmented dataset

Class	Original collected image	After Augment image
Tinea capitis	65	413
Tinea Corpories	58	415
Tinea Pedies	62	368
Total dataset	185	1196

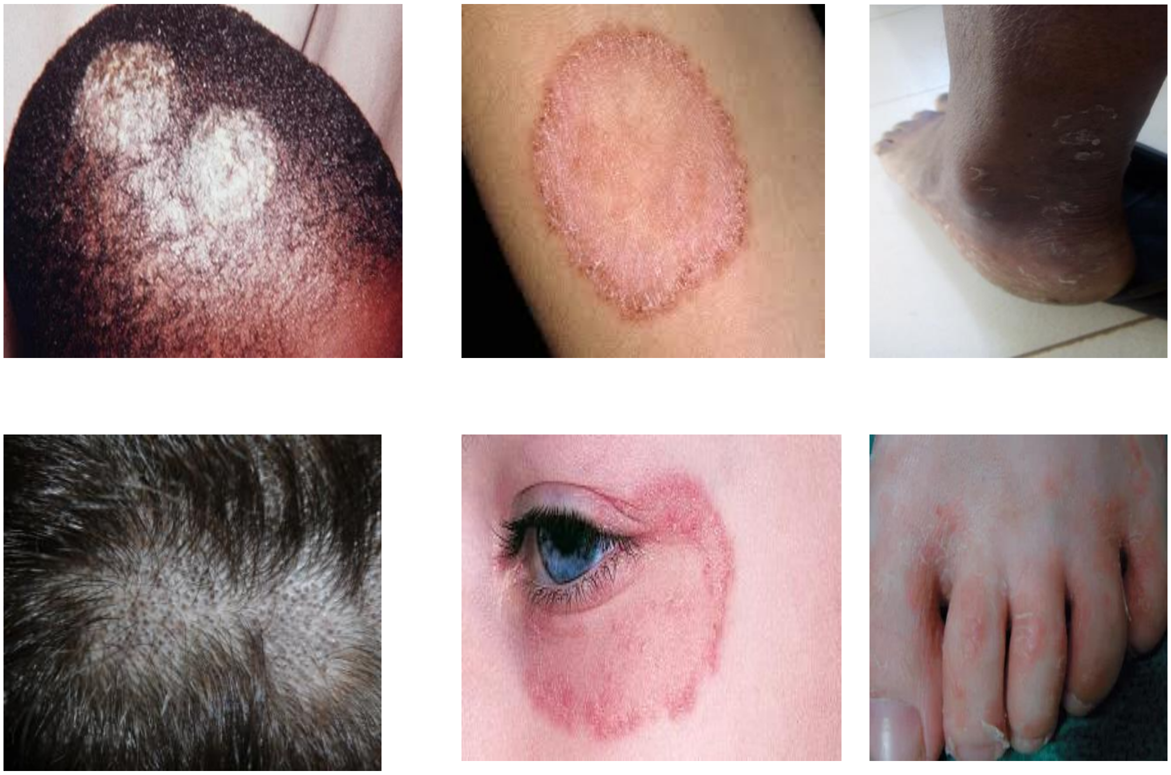

Figure 4. 1Sample data set

4.3. Implementation Tools

In order to implement the skin disease identification model, we made use several open source libraries. The following libraries are used on Windows 10 Pro with Processor Intel(R) Core(TM) i3, CPU 2370MHz and RAM 4.0 GB.

- Python: We choose Python as the programming language to develop the algorithms. The main reason for its selection is mainly because of its simplicity and code readability.
- The IDE used is Anaconda 3, Spyder 4.0.1 and Jupyter notebook 6.0.3
- OpenCV-Python which is a library of Python bindings designed to solve computer vision problems. CV2 module is inside this library which used to read image, resize image, change RGB to grayscale etc…
- NumPy: is the fundamental package for scientific computing with Python that contains a powerful N-dimensional array object. This library is needed in order to treat images

as matrices. NumPy is the key library for image manipulation. It is very fast, which allows our algorithms to run with high computational efficiency, which is part of the desired features of the proposed work. OpenCV Python makes use of Numpy, which is a highly optimized library for numerical operations.

- Scikit-learn is a free machine learning library for the Python. It features various classifications, regression and clustering algorithms including support vector machines, random forests, etc. and is designed to interoperate with NumPy and SciPy.
- Tensor flow is also free machine learning library used for mathematical operation.
- Keras library, used to create layers in cnn model
- SVC : is a module from SVM , which used to for classification purpose

4.4. Hyper Parameters in the model

While training the model, there are some hyper parameters i.e. variables that determine the network structure for optimized result of training. The first groups of hyper parameters are determined for the network based on number of training data set by choosing the batch size. Those are:

1. Number of Epoch: How many times the model reads all the data set

2. Batch Size: Parts from all dataset to be visited at a time

3. Number of iterations: The number of batches to compete the dataset in one epoch.

The second group of hyper parameters is discussed in visualization Section 3.2, which may different for different convolution layer. This are listed below and the value of the parameters described in the Table 4.2.

1. Kernel Size: Filter size to generate feature maps.

2. Padding: For keeping the size of output equal to the input of the next layer.

3. Stride: The distance of moving the kernel over convolving the image.

4. Pooling Size: Down sampling size of convolved result.

5. Learning Rate: a positive fraction determining the step of learning of the network neurons.

6. Dropout: a mechanism of regularization of the training to overcome over fitting

Based on our dataset we set values for the above parameters. We have set the value by making a comparison according to training time and the performance. The hyper parameters and the value assigned for each parameter is described below in the table 4.2.

Table 4. 2 Hyper parameters and value in the model

Hyper Parameter	Value
Epoch	50
Batch size	32
Iteration	30
Kernel size	3x3 for all layers
Padding	same
Pooling	2
Learning rate	0.001
Dropout	0.5

4.5. Evaluation techniques

To test the accuracy of the model, 20% of the lesions in the database are used as test images. The widely used measurement metrics such as precision, recall, f1 score and support are used to compute the accuracy of the system which is defined as follow.

Precision is defined as the ratio of

$$\frac{Tp}{Tp + FP} \qquad (4.1)$$

Where Tp is the number of true positives and Fp is number of false positives. The precision is intuitively the ability of the classifier not to label as positive a sample that is negative.

The recall is the ratio of

$$\frac{Tp}{Tp + FN} \tag{4.2}$$

Where Tp is the number of true positives and FN the number of false negatives. The recall is intuitively the ability of the classifier to find all the positive samples.

The f1score can be interpreted as a weighted harmonic mean of the precision and recall and is defined as

$$F1\ Score = \frac{2 * (\text{Recall } * \text{ Precision})}{\text{Recall} + \text{ Precision}} \tag{4.3}$$

Support is defined as the number samples of the true response that lie in that class. Accuracy is defined as a ratio of correctly predicted observation to the total observations.

$$Accuracy = \frac{Tp + TN}{Tp + Fp + FN + TN} \tag{4.4}$$

Where TN is true negative which is defined as the ration of tn/(tn+fp). Classification report () function displays the precision, recall, f1-score and support for each class

4.6. Evaluation result

The proposed model trained many times using different activation function (tanh, relu), dataset (original data and preprocessed data) using 200x200 , 224x224 and 300x300 image size. Finally we compare our model HSDCNet with Alex net CNN architecture.

In the first experiment we train the model using 200X200 original images using Relu and using Tanh differently. From this experiment we got 53% Training accuracy, 54% testing accuracy. However, during relu replaced by Tanh, the model achieves 60% and 58% training and testing accuracy respectively. This result indicates that image preprocessing is critical task to be done on our image datasets to increase the performance of identification.

4.6.1. HSDCNet model using 200X200 preprocessed image and Relu at training phase

After visualizing the first experiment result, we apply image preprocessing techniques on the collected image. As described in chapter three figure 3.1 the model consists training, validation and testing phase. To show the performance difference, we trained the model using training phase and validation phase differently. In this experiment we achieved 80% training accuracy and 75% validation accuracy. Training phase result is described below in the figure 4.2.and figure 4.3. The rest experiment done using validation phase.

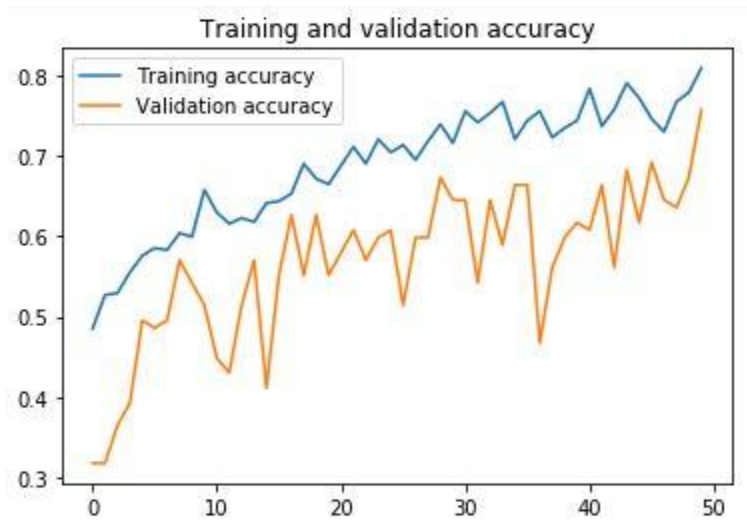

Figure 4 2 Training accuracy of the model using Training phase

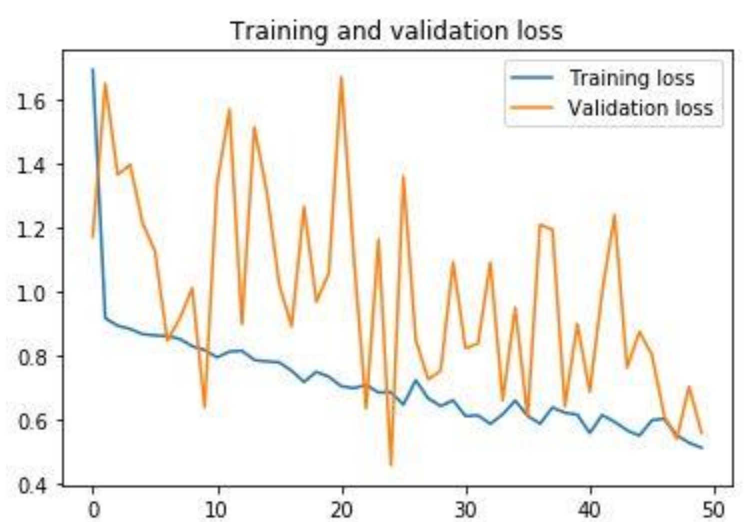

Figure 4 3 Training Loss the model using Training phase

4.6.2. HSDCNet model using 200X200 preprocessed image and Relu at validation phase

In the this experiment the model trained at validation Phase using 200X200 image size and the preprocessed image with Relu as activation function. From this we have achieved 88%, 87% for training and testing accuracy respectively. Hence the preprocessed image enhances the performance of model than used original image. The plot diagram shows this accuracy and loss result below in figure 4.4 and figure 4.5

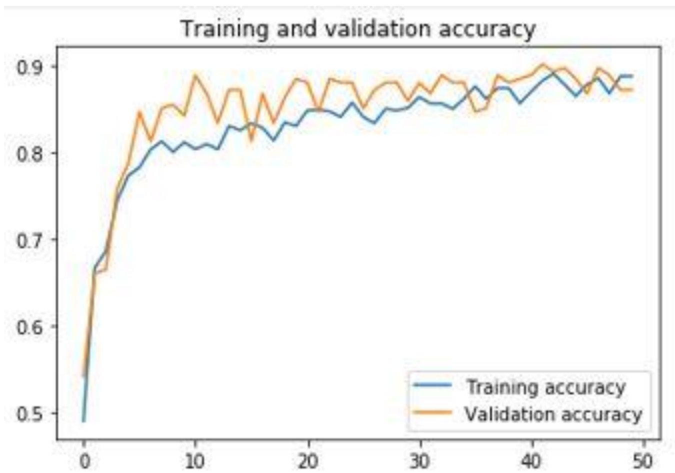

Figure 4 .4 Training and validation accuracy using preprocessed data and relu

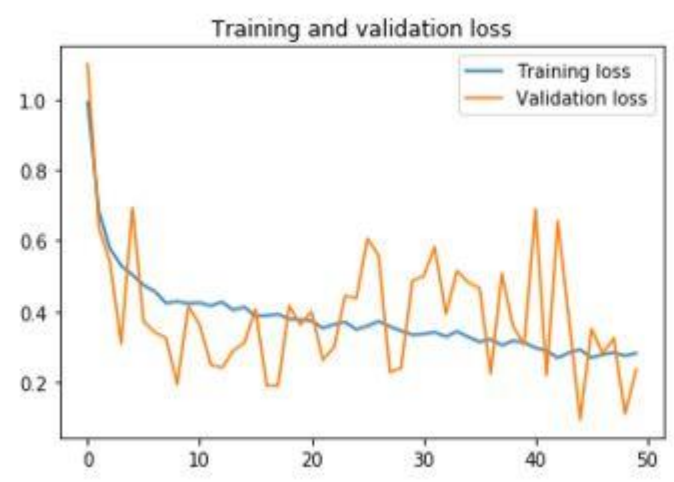

Figure 4. 5Training and validation loss using preprocessed data and relu

4.6.3. HSDCNet model using 200X200 preprocessed image and Tanh

In this experiment the model trained at validation Phase using 200X200 image size and the preprocessed image with Tanh as activation function. From this we have achieved 89%, 89.4% for training and testing accuracy respectively. This result shows that Tanh activation function enhance the performance than Relu. The plot diagram shows this accuracy and loss result in figure 4.6 and figure 4.7 below.

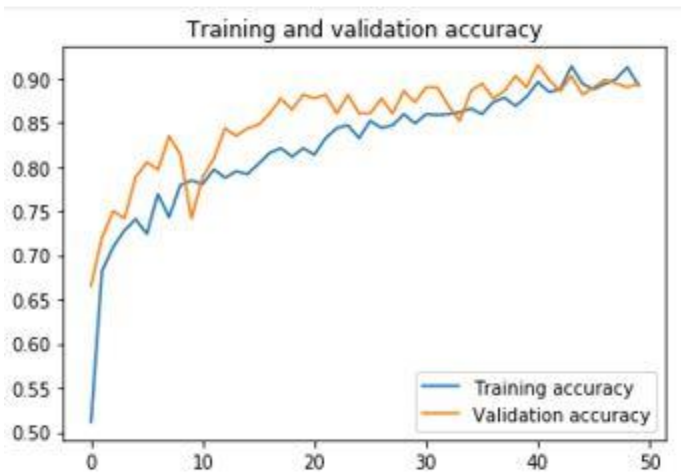

Figure 4. 6Training and validation accuracy using Tanh, prepossessed data

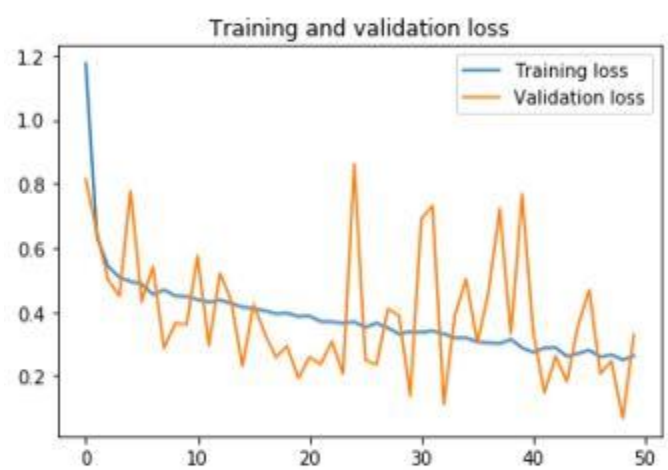

Figure 4 .7Training and validation loss using Tanh, prepossessed data

4.6.4. HSDCNet model using 224X224 preprocessed image and Relu

In this experiment the model trained at validation Phase using 224X224 image size and the preprocessed image with Relu as activation function. From this we have achieved 86%, 87% for training and testing accuracy respectively. This result shows as image size increases the performance model do not have more change. However, it takes more computational time and memory than the smaller one. The plot diagram shows the accuracy and loss result in figure 4.8 and figure 4.9 below.

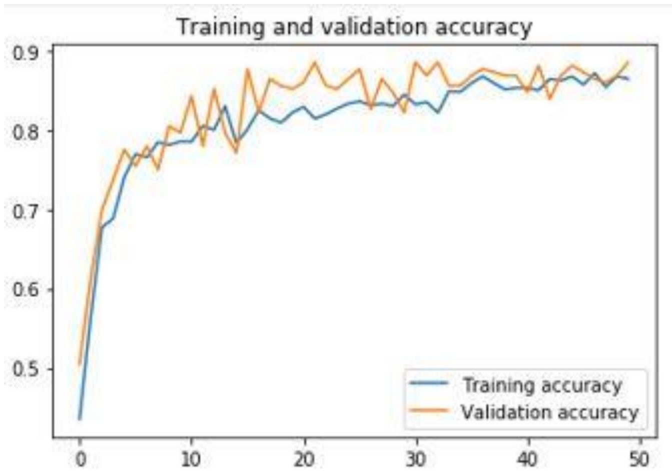

Figure 4. 8Training and validation accuracy using 224x224 and Relu

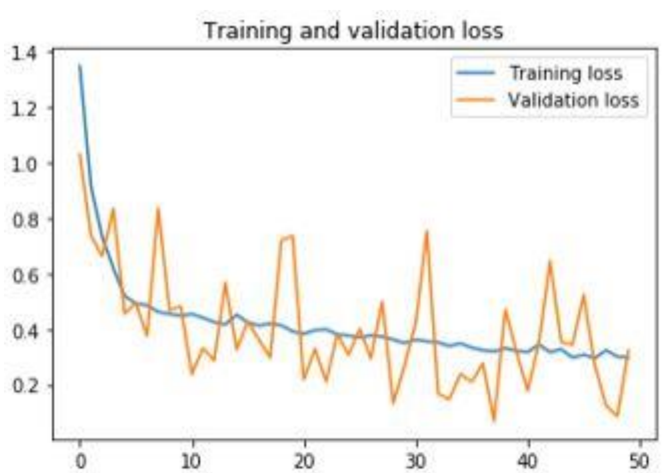

Figure 4 .9Training and validation loss using 224x224 and Relu

4.6.5. HSDCNet model using 224X224 preprocessed image and Tanh

In this experiment the model trained at validation phase using 224X224 image size and the preprocessed image with Tanh as activation function. From this we have achieved 92.2%, 89% for training and testing accuracy respectively. This result shows as image size increases the performance model do not have more change. However, it takes more computational time and memory than the smaller one. Hence, the training accuracy is more than testing accuracy overfitting problem also happened. The plot diagram shows the accuracy and loss result in figure 4.10 and figure 4.11 below.

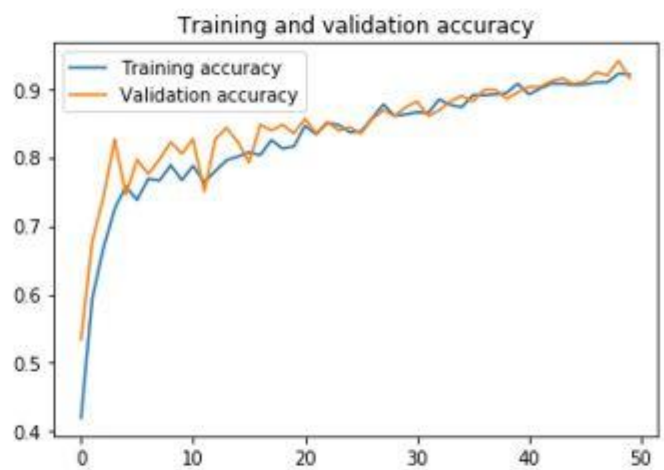

Figure 4 .10Training and validation accuracy using 224x224 and Tanh

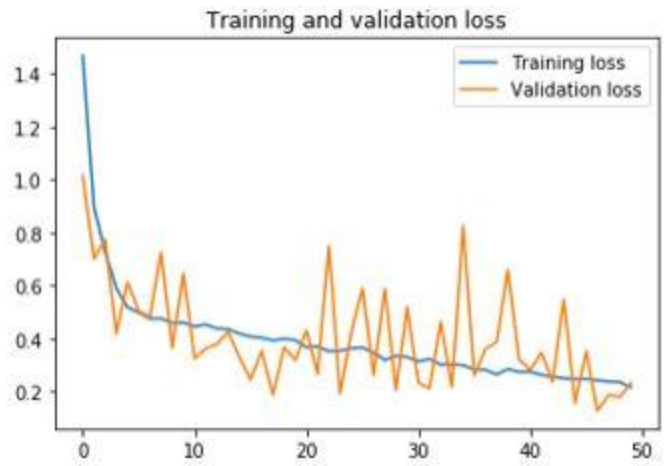

Figure 4 .11Training and validation loss using 224x224 and Tanh

4.6.6. HSDCNet model using 300x300 preprocessed image and Tanh

In this experiment the model trained at validation phase using 300x300 image size and the preprocessed image with Tanh as activation function. From this we have achieved 93.3%, 89.4% for training and testing accuracy respectively. This result shows as image size increases the performance model do not have more change. However, it takes more computational time and memory than the smaller one. Hence, the training accuracy is more than testing accuracy overfitting problem also happened. The plot diagram shows the accuracy and loss result in figure 4.12 and figure 4.13 below.

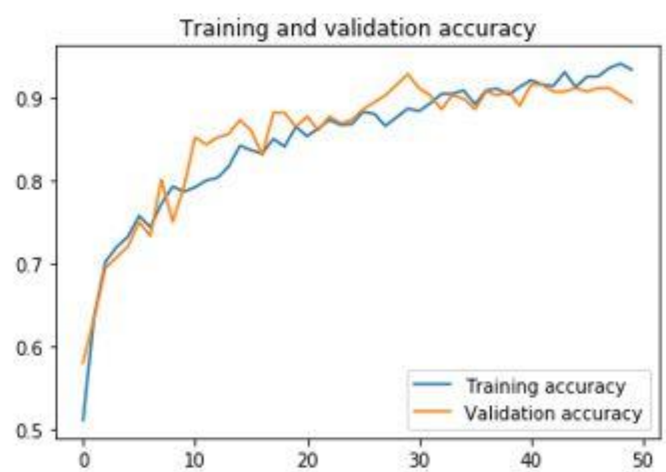

Figure 4. 12 Training and validation accuracy using 300x300 and Tanh

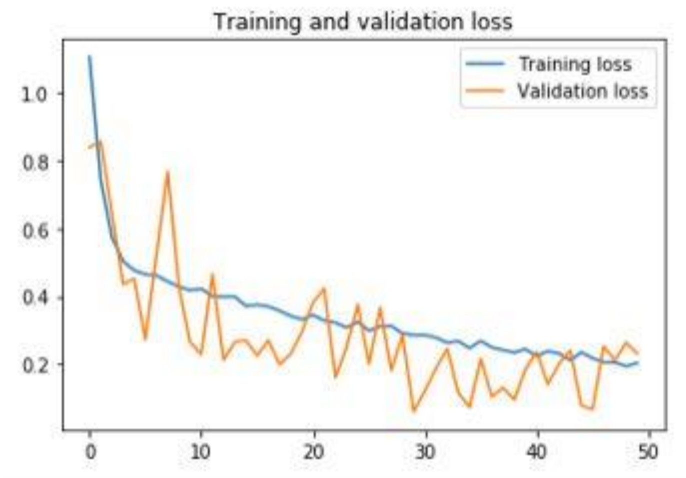

Figure 4 .13Training and validation loss using 300x300and Tanh

4.6.7. HSDCNet model using 200x200 image, segmentation and Tanh

In this experiment the model trained at validation phase using 200x200 image size and the preprocessed image with Tanh as activation function. From this we have achieved 95.6%, 95.6% for training and testing accuracy respectively. It takes averagely 3 hours. This result shows as image segmentation plays critical role to increases the performance of model .It takes medium computational time and memory than the larger image size and AlexNet model. Hence, the training accuracy and testing accuracy are equal there is no overfitting and underfitting problem. The plot diagram shows the accuracy and loss result in figure 4.14 and figure 4.15 below.

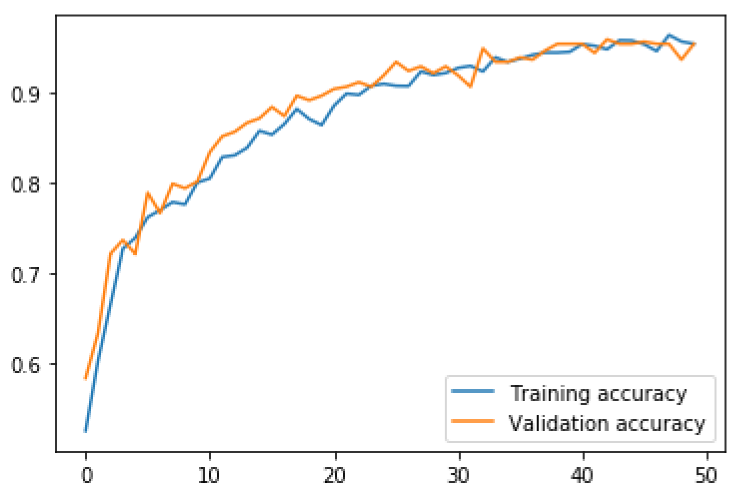

Figure 4 .14Training and validation accuracy using integration segmentation method, 200x200 and Tanh

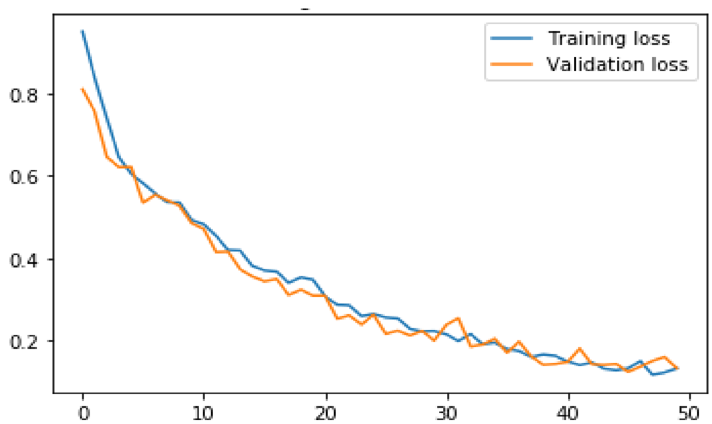

Figure 4 15Training and validation loss using integration segmentation method, 200x200 and Tanh

4.7. Model evaluation

Based on the above training result we evaluate our model using confusion matrix and deep learning evaluation techniques. As shown below in Figure 4.16, classification report is described based on the confusion matrix result. The evaluation measurements' are Precision, recall, F1-score and accuracy. The overall prediction of the model is measured by accuracy.

```
Epoch 49/50
30/30 [==============================] - 175s 6s/step - loss: 0.2505 - accuracy:
0.9125 - val_loss: 0.0285 - val_accuracy: 0.9491
Epoch 50/50
30/30 [==============================] - 167s 6s/step - loss: 0.1328 - accuracy:
0.9569 - val_loss: 0.1327 - val_accuracy: 0.9561
Model: "sequential_1"
```

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 200, 200, 32)	896
batch_normalization_1 (Batch	(None, 200, 200, 32)	128
max_pooling2d_1 (MaxPooling2	(None, 100, 100, 32)	0
conv2d_2 (Conv2D)	(None, 100, 100, 64)	18496
batch_normalization_2 (Batch	(None, 100, 100, 64)	256
max_pooling2d_2 (MaxPooling2	(None, 50, 50, 64)	0
conv2d_3 (Conv2D)	(None, 50, 50, 128)	73856
batch_normalization_3 (Batch	(None, 50, 50, 128)	512
max_pooling2d_3 (MaxPooling2	(None, 25, 25, 128)	0
flatten_1 (Flatten)	(None, 80000)	0
dense_1 (Dense)	(None, 128)	10240128
dropout_1 (Dropout)	(None, 128)	0
dense_2 (Dense)	(None, 3)	387
dropout_2 (Dropout)	(None, 3)	0

```
Total params: 10,333,763
Trainable params: 10,333,763
Non-trainable params: 0
```

```
> 95.611
```

Figure 4 .16 Training Progress of HSDCNet

```
Confusion Matrix
[[77 1  1]
 [2  75 2]
 [5  3  70]]
Classification Report
```

	Precision	recall	f1-score	support
T-capitits	0.91	0.97	0.93	79
T-corpories	0.94	0.94	0.94	79
T-pedies	0.95	0.89	0.91	78
accuracy			0.93	236
macro avg	0.93	0.93	0.92	236
weighted avg	0.93	0.93	0.92	236

Figure 4 .17Confusion matrix result and classification report of HSDCNet

The performance of the model will be poor either by overfitting or underfitting the data. The training of the model is plotted to see possibility of overfitting and underfitting in the model. Overfitting happens when a model learns the detail and noise in the training data to the extent that it negatively impacts the performance of the model on new data. When the training accuracy is above the test accuracy it means the model is overfitting. Our model is not overfitting as shown in the Figure 4.14 there is no significant difference between the value of training and test accuracy.

Underfitting refers to a model that can neither model the training data nor generalize to new data. When validation loss is below the training loss the model is underfitting. As shown in the figure 4.15 our model is not underfitting.

The model achieves 95.6% accuracy in 50 epochs. Continuing the training above 50 epochs, the model try to learn the data and the noise and the performance is not changing but overfitting happen.

4.8. Comparison with deep neural network architecture

Comparison is done with deep neural networks using the same dataset and parameter with architectural difference in the model. As we described in the literature section 2.7, there are difference types of deep neural network models such as AlexNet, ResNet, VGG16, VGG19, and GoogleNet. From these models we select AlexNet to evaluate with our model. The reason select this model is related to hardware dependency, since, this architecture run on the CPU and it has less number of layers than the rest CNN models.

The AlexNet architecture consists of five convolutional layers, some of which are followed by maximum pooling layers and then three fully-connected layers and finally a 1000-way softmax classifier.

First Layer: The input for AlexNet is a 224x224x3 RGB image which passes through the first convolutional layer with 96 feature maps or filters having size 11×11 and a stride of 4.The image dimensions changes to 55x55x96. Then the AlexNet applies maximum pooling layer or sub-sampling layer with a filter size 3×3 and a stride of two. The resulting image dimensions will be reduced to 27x27x96.

Second Layer: Next, there is a second convolutional layer with 256 feature maps having size 5×5 and a stride of 1. Then there is again a maximum pooling layer with filter size 3×3 and a stride of 2. This layer is same as the second layer except it has 256 feature maps so the output will be reduced to 13x13x256.

Third, Fourth and Fifth Layers

The third, fourth and fifth layers are convolutional layers with filter size 3×3 and a stride of one. The first two used 384 feature maps where the third used 256 filters.

The three convolutional layers are followed by a maximum pooling layer with filter size 3×3, a stride of 2 and have 256 feature maps.

Sixth Layer

The convolutional layer output is flatten through a fully connected layer with 9216 feature maps each of size 1×1.

Seventh and Eighth Layers

Next is again two fully connected layers with 4096 units.

Output Layer:

Finally, there is a Softmax output layer ŷ with 1000 possible value

4.8.1. Alex net implementation result

The performance of Alex net model is trained using tanh activation function the model achieves 84% training accuracy at 50 epochs and 86 testing accuracy. As shown in the figure 4.18 the training and validation accuracy, training and validation loss is plotted. It takes a lot computational time than HSDCNet averagely more than 5 hours .From the result since the validation loss is lower than the training loss, underfitting problem is happen. Hence the proposed HSDCNet model performs better than AlexNet model.

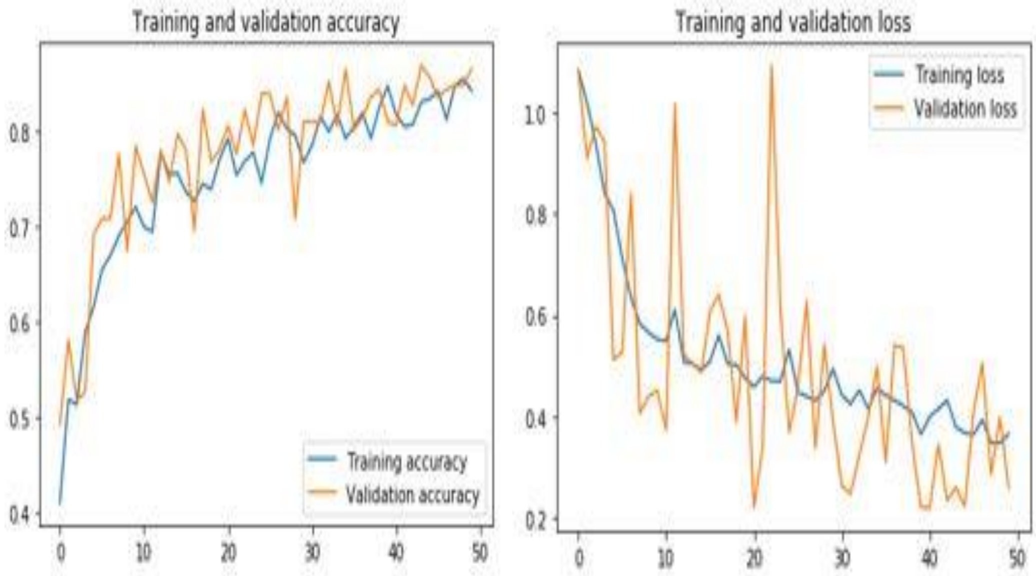

Figure 4 .18Alex net Training, validation accuracy and training and validation Loss

We evaluate AlexNet model using similar parameters as evaluation of our model HSDCNet. These are Recall, F1 Score, Accuracy, and Precision based on the confusion matrix result. As described below in figure 4.20 the accuracy in the confusion matrix is lower than HSDCNet model. The training progress of AlexNet is described in figure 4.19.

```
8475
Epoch 49/50
30/30 [==============================] - 309s 10s/step - loss: 0.3479 - accuracy:
0.8542 - val_loss: 0.3985 - val_accuracy: 0.8475
Epoch 50/50
30/30 [==============================] - 305s 10s/step - loss: 0.3658 - accuracy:
0.8417 - val_loss: 0.2595 - val_accuracy: 0.8644
Model: "sequential_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_1 (Conv2D)            (None, 54, 54, 96)        34944
_________________________________________________________________
activation_1 (Activation)    (None, 54, 54, 96)        0
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 27, 27, 96)        0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 17, 17, 256)       2973952
_________________________________________________________________
activation_2 (Activation)    (None, 17, 17, 256)       0
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 8, 8, 256)         0
_________________________________________________________________
conv2d_3 (Conv2D)            (None, 6, 6, 384)         885120
_________________________________________________________________
activation_3 (Activation)    (None, 6, 6, 384)         0
_________________________________________________________________
conv2d_4 (Conv2D)            (None, 4, 4, 384)         1327488
_________________________________________________________________
activation_4 (Activation)    (None, 4, 4, 384)         0
_________________________________________________________________
conv2d_5 (Conv2D)            (None, 2, 2, 256)         884992
_________________________________________________________________
activation_5 (Activation)    (None, 2, 2, 256)         0
_________________________________________________________________
max_pooling2d_3 (MaxPooling2 (None, 1, 1, 256)         0
_________________________________________________________________
flatten_1 (Flatten)          (None, 256)               0
_________________________________________________________________
dense_1 (Dense)              (None, 4096)              1052672
_________________________________________________________________
activation_6 (Activation)    (None, 4096)              0
_________________________________________________________________
dense_2 (Dense)              (None, 4096)              16781312
_________________________________________________________________
activation_7 (Activation)    (None, 4096)              0
_________________________________________________________________
dense_3 (Dense)              (None, 1000)              4097000
_________________________________________________________________
activation_8 (Activation)    (None, 1000)              0
_________________________________________________________________
dense_4 (Dense)              (None, 3)                 3003
_________________________________________________________________
activation_9 (Activation)    (None, 3)                 0
=================================================================
Total params: 28,040,483
Trainable params: 28,040,483
Non-trainable params: 0
_________________________________________________________________
> 86.441
```

Figure 4 .19 Training progress of Alex Net Model

```
Confusion Matrix
[[66 4 9]
 [10 59 10]
 [4  5  69]]
Classification Report
              precision    recall  f1-score   support

  T-capitits       0.73      0.73      0.73        79
 T-corpories       0.69      0.74      0.76        79
    T-pedies       0.72      0.77      0.74        78

    accuracy                           0.71       236
   macro avg       0.71      0.71      0.71       236
weighted avg       0.71      0.71      0.71       236
```

Figure 4 .20 Confusion matrix result and classification report of AlexNet Model

4.9. Summarizing evaluation Results

The overall evaluation result of Human Skin Disease Classification Network (HSDCNet) and the comparison result with AlexNet using Different image size(200x200,224x224,300x300), activation function(Relu,Tanh), original image, preprocessed image, and the integration segmentation techniques are summarized in the table 4.8 below.

Table 4. 3 Summarizing training model result

Dataset	**Image size**	**Activation function**	**Training Accuracy**	**Testing Accuracy**	**Training Time**
Original dataset	200x200	Relu	53%	54%	4hr &25minutes
Original dataset	200x200	Tanh	60%	58%	4hr &32minutes
Processed @Train_P	200x200	Relu	80%	75%	4hr&5minutes
Processed@Val_P	200x200	Relu	88%	87%	3hr&40 minutes
Processed@Val_P	200x200	Tanh	89%	89.4%	3hr&46 minutes
Processed@Val_P	224x224	Relu	86%	87%	4hr&40minutes
Processed@Val_P	224x224	Tanh	92.2%	89%	4hrs&45minutes
Processed@Val_P	300x300	Tanh	93.3%	89.4%	5hr&5 minutes
Threshold & region-based@Val_P	200x200	Tanh	95.6%	95.6%	3hr&10 minutes
AlexNet	224x224	Tanh	84%	86%	5hr&7 minutes

4.10. Discussion

As discussed in the section 4.5 the model trained more than 5 times by taking different parameters. In the section 4.9, the evaluation results are summarized. Firstly the model trained using the original data and Relu as activation function, result achieved 53% training accuracy and 54% testing accuracy. Secondly Relu replaced by the Tanh, from this 60%, 58% is training and testing accuracy respectively. From the third experiment, we used preprocessed 200x200 image and we achieved 88%, 89% for training accuracy using Relu and Tanh activation function respectively, which is better accuracy than the original data set result. The accuracy becomes 86% and 87% training and validation accuracy using relu activation function and preprocessed 224x224 image size. Using the preprocessed 224x224 image size and Tanh activation function, we have got 92.1% training and 89% validation accuracy. Using preprocessed image, 300x300 image size and Tanh activation function we have got 93.3% of training accuracy. From this result Tanh activation function increase the training and validation accuracy than Relu activation function. However, as we compare different image size, the training time increase since as image size become larger the no of pixels becomes high and the trained parameters also more than the smallest one, which is the factor to increase the training time. Finally based on the above result we choose the smallest image size 200x200 and Tanh activation function. By applying integration of threshold and region based segmentation techniques the accuracy becomes 95.6%, which is the better result than the other experiment result. The comparison result of AlexNet model is 84% training and 86% testing accuracy. Hence, our model HSDCNet achieves better result than the state of the art AlexNet model.

4.11. Summery

In this chapter, the total dataset and number of image in each class is described. Implementation tools used to create the system, the evaluation result of our model and comparison with CNN architecture AlexNet is described in detail. The challenge in the experiments is overfitting and underfitting in the training and validation loss or accuracy. Ours model, HSDCNet achieves higher classification accuracy than AlexNet. It was also trained faster and weighs less as compared to the AlexNet model. There is no underfitting and overfitting problem the validation and training accuracy is almost the same.

CHAPTER FIVE: CONCLUSIONS AND RECOMMENDATIONS

5.1. Conclusions

Many works are done so far in the area of Diagnosing, detection and classification of skin diseases. In this thesis, we developed a method to detect and classify human skin diseases common in developing countries like Ethiopia using machine learning algorithms. These are tinea corpories, tinea pedies, and tinea capitis. The datasets collected from patients at Bahir Dar Tibebe Giyon specialized Hospital and Bahir Dar Felege Hiywet General specialized Hospital using camera. In this work we described an algorithm for skin disease detection and classification. We used a deep learning approach Convolutional Neural Network for feature extraction and Support Vector machine for classification.

The classifier is separately trained using original, preprocessed data with Relu and Tanh activation function and using different image size. In the first experiment, the model trained using original 200x200 images with relu and tanh activation function separately . Secondly using the preprocessed data, 200x200 image size and relu activation tanh activation function. Thirdly using 224x224 image size, preprocessed data and tanh and Relu activation function separately. Using 300x300 image size, preprocessed data and tanh activation function, we have got 93.3% training accuracy , which is better result however, it takes much amount of training time than the smallest one(200x200). Using the integration segmentation techniques we have got 95.6%. From the result achieved, we can conclude that use Tanh perform better than relu in our dataset. Integration segmentation techniques perform better result than using single segmentation techniques.

Finally, we compare AlexNet model with our proposed model. From the comparison we have got 84% training accuracy. Based on this result our model, HSDCNet is better than this AlexNet model. This is because of the number of dataset we used is small. However AlexNet uses more number of parameters so over fitting or underfitting problem happened. Since the validation loss below training loss underfitting problem happened.

5.2. Recommendation

There is still gap from this proposed work. The proposed work can be further extended to improve the performance. Therefore, the following are some notable future work recommendations observed while implementing this research work:

- In this work the amount of dataset used are not much. It is better to increase dataset since Deep learning needs a huge amount of data
- Use other segmentation techniques to increase the performance of proposed method. Since image segmentation has a major role for the performance of classification.
- In this work only 3 classes are considered, it is better increase number of class to identify normal skin from all common skin disease type.
- Apply Gabor filter for texture feature extraction
- Compare the model with other CNN architecture such as GoogleNet, VGG, and ResNet etc.

REFERENCE

A.kaur, & N.kaur. (2015). Image Segmentation Techniques. *International Research Journal of Engineering and Technology, vol.2, pp.944-947*.

Adrian, D. R. (2017). *Deep Learning for Computer Vision with Python, PyImageSearch, 2017*.

Anatomy, P. o. (2019, NOV 7). *skin-problems-and-treatments/picture-of-the-skin*. Retrieved from https://www.webmd.com

Antonio, G., & Sujit, p. (2017). deep learning with keras.

Arden. (2020, May 8). Retrieved from Towards Data Science: https://towardsdatascience.com/applied-deep-learning-part-4-convolutional-neural-networks-584bc134c1e2

Arivazhagan, S., & Vineth, S. (2018). Mango Leaf Diseases Identification Using Convolutional Neural Network. *International Journal of Pure and Applied Mathematics, vol. 120, no. 6, pp. 11067-11079,*.

Bareilly, & Vidyapith, B. (2013). Various Image Enhancement Techniques- A Critical Review. *International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 3,*, 1-5.

Barghout, L., & J.Sheynin. (2013). "Real-world scene perception and perceptual organization: Lessons from Computer Vision",. *Journal of Vision, vol. 13*.

Bartlett, J. a. (2020, April 9). *Basic Biology of the Skin*. Retrieved from https://www.samples.ipbup.com/9780763761578/03ch pg029-032 Wiles.indd.pdf/

Bennett, & Howard. (2020, Feb 27). *Ever wonderd about your skin*. Retrieved from https://www.washingtonpost.com/lifestyle/kidspost/ever-wondered-about-your-skin/2014/05/23/02cc93ca-ba79-11e3-96ae-f2c36d2b1245_story.html

Brannon, H. (2020, 27 feb). *"Understanding the Epidermis*. Retrieved from https://www.verywell.com/anatomy-of-epidermis-1068881

Brendan P. Kelly, M. (2012). Superficial Fungal Infections. *skin disorders*, 1-19.

Brownlee, J. (2016). *Master Machine Learning Algorithms.*

Carl Louie Aruta, C. L., Calaguas, C. R., Gameng, J. K., & Prudentino, M. V. (2015). Mobile-based Medical Assistance for Diagnosing Different Types of Skin Diseases Using Case-based Reasoning with Image Processing . *International Journal of Conceptions on Computing and Information Technology* , 1-4.

Catherine R. Alimboyong, A. A. (2018). Classification of Plant Seedling Images Using Deep Learning. *IEEE Region 10 Conference*.

Chande, R., & Gupta, l. (2013). Image Filtering Algorithms and Techniques: A Review. *International Journal of Advanced Research in Computer Science and Software Engineering*, 1-5.

Chitradevi, B., & P.Srimathi. (2014). An Overview on Image Processing Techniques. *International Journal of Innovative Research in Computer and Communication Engineering Vol. 2, Issue 11*.

Daniel Hailemichael, L. (2015). Development of Automatic Maize Quality Assessment System Using Image Processing Techniques.

Dong, p. T. (2013). A Review on Image Feature Extraction and Representation Techniques. *International Journal of Multimedia and Ubiquitous Engineering*.

Elayaraja, P., & Suganth, M. (2014). Survey on Medical Image Segmentation Algorithms. *International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE)*.

Enbiale, W., & Ayalew, A. (2018). Investigation of a Scabies Outbreak in Drought-Affected Areas in Ethiopia. *Tropical medicine and infectious Diseases*, 1-9.

Hay, R., & Morris-Jones, R. (2012). *The diagnosis and managemen to tinea*, 1-11.

Hay, R., Bendeck, S. E., Chen, S., Estrada, R., Anne, H., McLeod, T., & Mahé, A. (2016). *Skin Diseases.*

Heba Fathy, E. (2016). Performance Improvement of Plant Identification Model based on PSO Segmentation. *I.J. Intelligent Systems and Applications*, 1-6.

Hoffman, M. (2019, NOV 7). *skin-problems-and-treatments*. Retrieved from Picture of the Skin (Human Anatomy): https://www.webmd.com/

Ji, M., & Yuyu, Y. b. (2019). “Dimension reduction of image deep feature using PCA. *J. Vis. Commun. Image R. 63 (2019) 102578*, 1-8.

K.Bhargavi, & Jyothi, S. (2014). "A Survey on Threshold Based Segmentation Technique in Image Processing. *International Journal of Innovative Research and Development, Vol. 12(12)*, .

Kaganami, H. G., & Beij, Z. (2009). Region based detection versus edge detection. *IEEE Transactions on Intelligent Information Hiding and Multimedia Signal Processing, pp.*, 17-21.

Kaiming, H., Xiangyu, Z., Shaoqing, R., & Jian, S. (2015). *Deep Residual Learning for Image Recognition.* arXiv preprint arXiv: 1512.03385, pp. 1-12,.

Kalra, K., Goswami, A. K., & Gupta, R. (2013). A comparative Study of Supervised Image Classification Algorithms for Satellite Images . *International Journal of Electrical, Electronics and Data Communication, ISSN: 2320-2084 Volume-1*.

Kamdi, S., & R.K.Krishna. (2012). Image Segmentation and Region Growing Algorithm. *in International Journal of Computer Technology and Electronics Engineering (IJCTEE) Volume 2, Issue* .

Kamulegeya, L. H., Okello, M., Bwanika, J. M., Musinguzi, D., Lubega, W., Rusoke, D., . . . Börve, A. (2019). Using artificial intelligence on dermatology conditions in Uganda: A case for diversity in training data sets for machine learning. 1-30.

Karen, S., & Andrew, Z. (2014). Very deep convolutional networks for large-scale image recognition. 1-14.

Krizhevsky, A., E.Hinton, G., & Sutskever, I. (2012). ImageNetClassificationwithDeepConvolutional NeuralNetworks. *Advances in neural information processing systems*, 1-9.

Kumar, M., Kamble, M., Pawar, S., Patil, P., & Bonde, N. (2014). Survey on Techniques for Plant Leaf Classification. *International Journal of Modern Engineering Research (IJMER),Vol.1, Issue.2, pp-538-544 ISSN*.

Li-sheng, W., Quan, G., & Tao, J. (2018). Skin Disease Recognition Method Based on Image Color and Texture Features. 1-11.

Liu, T., Fang, S., Zhao, Y., & Wan, P. (2017). Implementation of Training Convolutional Neural Networks.

Macedo, D. (2016). Improving Image Classification Accuracy using Hybrid Systems of Support Vector Machines and Convolutional Neural Networks. 1-3.

Maini, R., & Aggarwal, H. (2010). A Comprehensive Review of Image Enhancement Techniques. *JOURNAL OF COMPUTING, VOLUME 2, ISSUE 3,*, 1-6.

Mary Seaburyn, S., Rosenfeld, S., & W.Ely, J. (2014). *Diagnosis and Management of Tinea Infections*, 1-10.

Maseleno, A., & Md. Mahmud, H. (n.d.). Skin Infection Detection using Dempster-Shafer Theory . *[EEE/OSA/[APR International Conference on [nfonnatics, Electronics & Vision* , 1-5.

Masood, A., & Jumaily, A. (2013). Computer Aided Diagnostic Support System for Skin Cancer: A Review of Techniques and Algorithms. *InternationalJournalofBiomedicalImaging Volume2013*, 1-23.

Mrs. Jayashree, H., Aishwarya, B., Harsha, A., & Nisha, K. (2019). Skin Disease Detection Using Image Processing with Data Mining and Deep Learning. *International Research Journal of Engineering and Technology (IRJET)* , 1-4.

Nikou, C. (2008). Digital Image Processing. In P. H. R. Gonzalez and R. Woods. Digital Image Processing, *Digital Image Processing* (pp. 1-47). University of Ioannina : C. Nikou – Digital Image Processing (E12).

Oke, O. O., Onayemi, O., Olasode, O. A., Omisore, A. G., & AbimbolaOninla, O. (2014). *The Prevalence and Pattern of Superficial Fungal Infections among School Children in Ile-Ife, South-Western Nigeria*, 1-8.

Patterson, J., & Gibson, A. (2015). Deep Learning.

Petrellis, N. (2018). A Review of Image Processing Techniques Common in Human and Plant Disease Diagnosis. *Symmetry2018, 10, 270, doi:10.3390/sym10070270*.

Petrellis, N. (2018). Using Color Signatures for the Classification of Skin Disorders. *7th International Conference on Modern Circuits and Systems Technologies*, 1-4.

Połap, D., Winnicka, A., Serwata, K., sik, K. K., & ´niak, M. W. (2018). An Intelligent System for Monitoring Skin Diseases. 1-19.

Preetom, S. A., & Chowdhury, R. R. (2018). Identification and Recognition of Rice Diseases Using Deep Convolutional Neural Networks. *Bangladesh University of Engineering and Technology, pp*, 1-56.

Pupale, R. (2018, Jun 16). *Support Vector Machines(SVM) — An Overview*. Retrieved from medium-com-pupalerushikesh-svm-f4b42800e989: https://towardsdatascience.com

Raghavendra, U., Bhat, N., Gudigar, A., & Achary, U. (2018). Automated system for the detection of thoracolumbar fracture using a CNN architecture",. *Future Generation Computer Systems*.

Raj Bridgelall, P. (2017). *Introduction to Support Vector Machines* .

Razzak, M. I., Naz, S., & Zaib, A. (2018). *Deep Learning for Medical Image Processing.* Retrieved from Challenges and the Future, Classification in BioApps: https://doi.org/10.1007/978-3-319-65

Ruela, M. (2015). Imaging & Visualization A system for the detection of melanomas in dermoscopy images. *Computer Methods in Biomechanics and Biomedical Engineering* .

S.Lakshmi, & V.Sankaranarayanan, D. (2010). A study of edge detection techniques for segmentation computing approaches. *, IJCA Special Issue on Computer Aided Soft Computing Techniques for Imaging and Biomedical Applications (CASCT),*.

Sakshi, I., Kumar, G. A., Mishra, S. P., & Asopa, P. (2018). Conceptual Understanding of Convolutional Neural Network- A. *in International Conference on Computational Intelligence and Data Science (ICCIDS 2018)*.

Sam, S. A., Philip, A. A., Nartey, C. N., & Nti, K. I. (2019). A Web-Based Skin Disease Diagnosis Using Convolutional Neural Networks. *I.J. Information Technology and Computer Science, 2019*, 1-7.

Seyyid, A. M. (2015). A Comparative Study of Feature Extraction Methods in Images Classification. *I.J. Image, Graphics and Signal Processing*.

Shereena, V. B., & Julie, M. D. (2015). SIGNIFICANCE OF DIMENSIONALITY REDUCTION IN IMAGE PROCESSING. *Signal & Image Processing : An International Journal (SIPIJ) Vol.6, No.3*, 1-16.

Sumithra, R., Mahamad, S., & Guruc, D. S. (2015). Segmentation and Classification of Skin Lesions for Disease Diagnosis. *International Conference on Advanced Computing Technologies and Applications* (pp. 1-10). India: Elsevier B.

Tigistu, G. (2017). Automatic Flower Disease Identification using image processing . *Master's Thesis. Department of computer Science, Addis Ababa University, Addis Ababa, Ethiopia*.

Tribhuvan Singha, R. S. (2018). Software Cost Estimation using Environmental Adaption method. *ELSEVIER*.

V. Sandeep, V., K.Kanaka, D., & Keshavulu. (2013.). Seed image analysis: its applications in seed science research. *International Research Journal of Agricultural Sciences, vol. 1*, 30-36.

Vedaldi, A., & Lenc, K. (2015). MatConvNet Convolutional Neural Networks for MATLAB. *ACM MULTIMEDIA 2015 OPEN SOURCE SOFTWARE COMPETITION*, 1-4.

Velasco, J., Pascion, C., Alberio, J. W., Apuang, J., Cruz, J. S., Gomez, M. A., . . . Jr. Jorda, R. (2019). A Smartphone-Based Skin Disease Classification Using MobilNet CNN. *International Journal of Advanced Trends in Computer Science and Engineering*, 1-7.

Virtual Medical, C. (2020, feb 27). *Human Skin*. Retrieved from https://www.myvmc.com/anatomy/human-skin

Yogamangalam, R., & Karthikeyan, B. (2013). Segmentation Techniques Comparison in Image Processing. *International Journal of Engineering and Technology (IJET)*.

Zoltan, C. (2018, Nov 12). *SVM and Kernel SVM*. Retrieved from /svm-and-kernel-svm-fed02bef1200: https://towardsdatascience.com