Detection and Prediction of Acral Lentiginous Melanoma in Dermoscopic Images using Deep Learning

Session: 2018 – 2020

Submitted by:

Qaiser Abbas 2018-MS-CS-23

Supervised by:

Dr. Muhammad Usman Ghani Khan

Department of Computer Science and Engineering

University of Engineering and Technology

Lahore Pakistan

Detection and Prediction of Acral Lentiginous Melanoma in Dermoscopic Images using Deep Learning

Submitted to the faculty of the Computer Science and Engineering Department of the University of Engineering and Technology Lahore in partial fulfillment of the requirements for the Degree of

Master of Science

in

Computer Science.

Internal Examiner	**External Examiner**
Signature:	Signature:
Name:	Name:
Designation:	Designation:
Chairman	**Dean**
Signature:	Signature:
Prof. Dr. Shazia Arshad	Name:

Department of Computer Science and Engineering

University of Engineering and Technology

Lahore Pakistan

Declaration

I declare that the work contained in this thesis is my own, except where explicitly stated otherwise. In addition this work has not been submitted to obtain another degree or professional qualification.

Signed: _____

Date: _____

Acknowledgments

Firstly, I'd like to express my thanks to my supportive supervisor, Dr. Muhammad Usman Ghani Khan, who has supported me throughout this research project. I am extremely grateful for our friendly conversations at the end of our meetings and his personal support throughout this research.

Secondly, my deepest gratitude to mam Farheen Ramzan for her constant support and valuable feedback and friendly behavior, who guided me at every step and provided feedback From the bottom of my heart I would like to say big thank you to the Bioinformatics Research Lab for providing technical assistance and help throughout my research work.

Lastly, A huge thanks to family and friends for their love and care. I'd like to thank my friends for their constant support and for believing in me especially I'd like to express my deepest gratitude to my best friends Anza Gul, Maheen Jafer, Zubair Farooq, Humayoun Faiz, Saqib Hassan, Imran Ali and Ali Hassan Khan Niazi.

To my parents for their unconditional love and for always being there for me.

To my family and friends for their support and encouragement

To all the people out there in world suffering from cancer. . .

Contents

List of Figures

List of Tables

Abbreviations

ALM	**A**cral **L**entiginous **M**elanoma
AM	**A**cral **M**elanoma
BN	**B**enign **N**evus
CNN	**C**onvolutional **N**eural **N**etwork
ResNet	**R**esidual **N**eural **N**etwork
CAD	**C**omputer **A**ided **D**iagnosis
DNA	**D**deoxyribo **N**ucleic **A**cid
BCC	**B**asal **C**ell **C**arcinoma
SCC	**S**quamous **C**ell **C**arcinoma
UV	**U**ltra **V**oilet
CT	**C**omputerized **T**omography
MRI	**M**agnetic **R**esonance **I**maging
SVM	**S**upport **V**ector **M**achines
AMNet	**A**cral **M**elanoma **Net**owrk
PET	**P**ositron **E**mission **T**omography
RCM	**R**eflectance **C**onfocal **M**icroscopy
ANN	**A**rtificial **N**eural **N**etwork
GLCM	**G**ray **L**evel **C**o-occuerence **M**atrix
ABCD	**A**rea **B**order **C**olor **D**iameter
kNN	**k** **N**earest **N**eighbor
VGG	**V**isual **G**eometry **G**roup
ISIC	**I**nternational **S**kin **I**maging **C**ollaboration
LOF	**L**ocal **O**utlier **F**actor
YOLO	**Y**ou **O**ly **L**ook **O**nce

R-CNN	**R**egion-based **C**onvolutional **N**eural **N**etwork
FCN	**F**ully **C**onvolutional **N**etwork
COCO	**C**ommon **O**bjects in **C**ontext
DSS	**D**ecision **S**upport **S**ystem
ROI	**R**egion **O**f **I**ntrest
ILSVRC	**I**mageNet **L**arge **S**cale **V**isual **R**ecognition Competition
TPU	**T**ensor **P**rocessing **U**nit
GANs	**G**enerative **A**dversarial **N**etworks
SGDM	**S**tochastic **G**radient **D**escent with **M**omentum
ROC	**R**eceiver **P**rocessing **U**nit
PPV	**P**ostive **P**redictive **V**alue
TPR	**T**rue **P**ositive **R**ate
ReLU	**R**ectified **L**inear **U**nit

Abstract

Melanoma is the rare and life threatening skin cancer and it is hard to diagnose because of the similarity between the melanoma and non-melanoma lesions. It is diagnosed by the expert dermatologists using various medical imaging techniques. The dermoscopy imaging technique is widely used by doctors which is non-invasive medical imaging technique. However, it is still difficult for a dermatologist for diagnosing melanoma due to the very minor differences between a melanoma and non-melanoma skin lesion. Melanoma has further four subtypes and Acral Lentiginous Melanoma is one of the rare melanomas. This type of melanoma occurs on palms and soles of human body which is more common in Asians and Africans. Most of the research in skin cancer diagnosis is related to classifying the lesions into melanoma and non-melanoma. Very limited work is done in further classification of subtypes of melanoma. In this research work we applied deep convolutional neural networks to classify the skin cancer into subtype of melanoma. We developed a deep learning based system to classify the dermoscopic images of skin cancer into acral melanoma and acral benign nevus. We have used dermoscopic images dataset of skin lesions from Yonsei University Health System South Korea for classification of skin lesions into acral lentiginous melanoma and acral benign nevus. We applied various image processing and data augmentation techniques to develop a robust automated system for acral melanoma detection. Due to acral melanoma's infrequent occurrences, limited data is available so its early diagnosis is hard. To overcome this problem, we applied data augmentation technique to develop an enhanced large dataset to train a deep learning model. Our convolutional neural network achieved an accuracy of 91% on unseen test dataset. Additionally, this enhanced dataset was used to train three different deep learning models which are AlexNet, Mobile Net V2 and ResNet-18. They achieved an accuracy of 95%, 93% and 97% respectively. From our analysis and results we found that our models performed well and these models were able to classify the skin cancer efficiently. Our findings prove that the automatic feature learning models can be used by dermatologists in clinical decision making process for early diagnosis of the acral melanoma using dermoscopic images.

Chapter 1

Introduction

The skin is body's largest organ and consists of two layers which are the dermis and epidermis. The epidermis is a meager layer of cells and over the inner dermis forms a defensive layer. It acts to prevent ailments on the inner layers of the skin. Melanocytes (skin cells) contain melanin that assimilates light vitality and protects against the harmful effects of the sun's bright rays. Melanin provides the skin with a form of shade [25]

Skin cancer is the unrestrained irregular production of skin cells. It was discovered by a French physician Rene Theophile Hyacinthe Laënnec (inventor of the stethoscope) in early 1800s. It happens when unrepaired DNA injury to skin cells, (most of the time produced by electromagnetic energy from sunshine or tannings beds) activity changes, or inherited faults [21]. As a result the top of the skin cells grow fast and procedure malignant tumors. Skin Cancer is categorized in major two types which are melanoma and non-melanoma. Melanoma is a type of cancer that begins in skin's pigment control cells which are called melanocytes. The name melanoma originates from Greek 'melanose' word. Which means black, reflective or the dark pigmentation. Non-melanoma skin cancer refers to a group of cancers in the upper layers of the skin that grow slowly. Most of the skin cancer occurs due to exposure to sun and ultraviolet rays however, it can also occur on skin which is not exposed to the sun or ultraviolet rays. The melanoma skin cancer is further classified in four subtypes which are the nodular melanoma, superficial spreading melanoma, acral melanoma and nodular lentigomaligna melanoma [38]. Whereas non-melanoma skin cancer is further classified into 2 major groups which are: basal cell carcinoma and squamous cell carcinoma.

The Acral lentiginous melanoma is most common in people with darker skin such as African, Hispanic and especially Asian ancestries [39]. It is not caused by

the exposure to the sun. It nearly makes up less than 5% of melanoma skin cancers [38]. Acral melanoma is a thin , flat spot of discoloured skin that is sometimes dark brown or black. It normally expands outward over the surface of the skin until it begins to develop down into the skin. It usually develops on the soles or palms or under nails sometimes. This happens on the body's non-hair-bearing surfaces, which can or can not be exposed to sunlight. Acral lentiginous melanoma is a result of malignant melanocytes at the membrane of the skin [31]. The pathogenesis of acral lentiginous melanoma remains unknown at this time. It has a poorer prognosis compared to that of cutaneous malignant melanoma [39].

According to recent reports [35] the incidence of skin cancer is growing, with the number of people dying from skin cancer every year almost doubling since 1990.. Among other cancer types, Melanoma is the deadliest which is responsible for death of around 55,000 people each year [35]. At least 1 in 5 individuals diagnosed with metastatic melanoma do not live more than five years, according to studies.

Detection and diagnosis of melanoma is a challenging due to visual similarity of the cancer infected parts and normal moles even the most experienced doctors cannot diagnose accurately [21]. Different medical imaging and biopsy techniques are applied to diagnose the melanoma such as dermatoscopy, CT scans and ultra sound. Out of these tests major tests are not required for early stage melanomas. However, dermatoscopy is usually used to analyze the skin lesions for early diagnosis of the skin cancer specially melanomas and benign nevus [22]. The dermatoscope contains magnifier, a transparent plate and a light source. These images are captured for early detection of cancer. The dermatologists who are expert in dermatoscopy can effectively diagnose the cancer.

So. there is also a major increase in the sensitivity (melanoma detection) as well as the precision relative to the inspection of the naked eye. The accuracy of the dermatoscopy was improved by up to 20% in the case of sensitivity and up to 10% in the case of precision compared to the naked eye test. [22]. This increases the precision of the dermatoscopy, reducing the incidence of needless surgical excisions of benign lesions. However, human error is still possible and an experienced dermatologist may misdiagnose the acral melanoma [39] which can be serious life threatening problem as the visual difference between melanoma and non-melanoma is zero sometimes.

To overcome this computer aided diagnosis methods can be utilized as they may make zero percent mistakes in diagnosing. Image processing and vision based machine learning techniques are widely used in computer aided diagnosis (CAD) research where the dermoscopic images are analyzed through computer programs

and results is obtained through these analyses which helps the dermatologists in their early decision making which results in accurate diagnosis [23] [28]. Traditional machine learning techniques requires extracted features which are required to classify and diagnose the skin cancer. However, this feature extraction process is very difficult and time consuming. Most of the work in literature is related to classification of melanoma and non-melanoma using image processing and machine learning techniques. Very limited work is done in further sub-type classification of skin cancer [15].

In recent years the trend has shifted to opt deep learning based models to perform end to end diagnosing tasks where a medical image is provided along with label and the model is trained on this data [7]. The main advantage of using deep learning is that feature engineering or feature extraction step can be skipped which is performed by the deep learning algorithms. This method results in robust and improved performance of automated diagnosis systems.

In order to assist the dermatologists in melanoma diagnosis many computer based diagnosis algorithms have been built which uses computer algorithms to diagnose and help dermatologists in early detection of melanomas [5], [10]. Deep learning algorithms such as CNNs and image processing techniques are the most important part of common computer based medical systems [39]. The recent studies [34], [7], [13] has proved the effectiveness of these algorithms in automatic detection tasks. Upon reviewing the literature, we found that very limited work is done to diagnose the acral melanoma because of its infrequent occurrences. The prior researches [20], [32], [34] only focused on categorizing the skin lesions images into some cancer subtype but they do not provide further information about the sub type of the cancer. For example [23] classifies the skin lesions into melanoma and non-melanoma. Only few [38], [39] of the researchers have worked on acral melanoma detection. The classification of melanomas in subtypes is very important for better diagnosis and it can increase the survival rates.

To overcome the gap in literature, in this research we have designed and implemented an automated end-to-end deep learning algorithm for diagnosis and classification of acral melanoma (AM) in dermoscopic images dataset. The methodology for the system development consists of (I) data pre-processing (II) deep learning model development (III) training and (IV) testing. The dermoscopic images dataset was preprocessed using different image processing techniques and the dermoscopic artifacts were removed for further analysis. To overcome the issue of limited data we applied different transformations on our dataset to artificially increase the images in dataset. After preprocessing and data augmentation the

deep learning models were used to extract features followed by training of the models on extracted features. Final step was to test the model on unseen test images dataset. Several convolutional neural networks were trained using enhanced dataset and hyper parameter tuning was performed to enhance the accuracy of the automated system. The objective was to achieve reasonable accuracy on unseen test data so that the computational model can be used effectively for acral melanoma (AM) detection and prediction.

placeins

Chapter 2

Literature Review

With the advancement in technology and medicine, the researchers from all over the world are developing and searching for effective automated systems for medical uses. A common disease like skin cancer is fatal if not treated properly. Hence the researchers from interdisciplinary fields like computer science, medicine, dermatology and healthcare are working together to develop effective automated systems for early diagnosis, treatment and prevention of fatal diseases. Cancer is a major cause of mortality in the world which motivates researchers from different fields to develop a thorough understanding of this fatal disease to prevent the deaths. Since cancer has many types which further include subtypes, the computer science researchers are developing and testing different medical imaging [36] and deep learning-based techniques for early detection so a suitable treatment can be planned in advance to save human lives.

2.1 Medical Imaging and Assessment of Melanoma

Melanoma skin cancer is diagnosed using numerous medical imaging techniques which include both invasive and non-invasive techniques [8]. An invasive procedure involves invading the body usually by cutting the body and inserting medical instruments into the body to assess different diseases. A non-invasive technique does not involve invading the body. Some of the medical imaging to diagnose and determine the treatment are Chest x rays, Ultrasound, Computed Tomography (CT) scan, Magnetic Resonance Imaging (MRI) scan, and Positron emission tomography (PET) scan. Most of these techniques are used to determine the treatment and may not require for early diagnosis.

2.1.1 Dermatoscopy

The early diagnosis is performed using dermoscopy which is also called surface microscopy where a skin lesion is analyzed under highly luminous and magnified lens powered device called dermatoscope. Once a doctor decides whether the skin lesion is melanoma than further skin biopsy techniques are applied where suspicious lesion area is removed and lab tests are performed under microscope.

2.1.2 Skin Biopsy

There are numerous methods of skin biopsy such as shave biopsy, punch biopsy, and optical biopsies. In shave biopsy top skin layers are shaved off using a tiny surgical blade. In a punch biopsy, the doctor uses a circular cookie cutter as a method to extract a deeper skin sample. While optical biopsy techniques do not require skin samples to be removed, for instance, reflectance confocal microscopy (RCM).

2.1.3 Chest X rays

This test may be performed to determine the metastatic melanoma in lungs. Acral melanoma is malignant melanoma which possess the ability to metastasize hence it can affect the nearby body organs hence chest x rays can be used to determine whether it has spread to lungs or not.

2.1.4 Ultra Sound

Ultrasound uses sound waves to create images from the inside of the body on a screen. This test could be performed to examine the lymph nodes adjacent to the tumour, particularly if the tumours are enlarged. Ultrasound is a fast and simple procedure and does not require exposure to radiation. Ultrasound is used to help direct a biopsy needle to a suspected lymph node.

2.1.5 Computed Tomography (CT) Scan

The CT scan used x-rays to produce accurate, cross-section images of the body. CT scans may reveal information in soft tissues such as the internal organs of the body. This test can show whether any lymph nodes are swollen or whether organs such as the lungs or the liver have unusual spots that may be caused by the spread of melanoma. CT scans can also be useful to help direct a biopsy needle to a suspicious region of the human body.

Imaging Technique	**PROs**	**CONs**
Dermatoscopy	Non-invasive and helpful for Early diagnosis	Requires expert doctor
Ultrasound	quick and easy technique, doesn't require exposure to radiation	Requires improved anatomical knowledge and a formal educational program.
Chest X Rays	Can be used to determine whether melanoma has spread to the lungs or not	Exposure to high radiation levels can have a range of effects
CT Scan	can display the detail in soft tissues	Needs exposure to radiation and the use of a contrast material
MRI Scan	helpful for looking at the brain and spinal cord	Must remain still in an enclosed machine, problem for claustrophobic patients
PET Scan	Useful for detection of advanced melanoma	radioactive material is combined with glucose, dangerous for Diabetes patients

Table 2.1: Medical Imaging Tests for Melanoma Dignosis

2.1.6 Magnetic Resonance Imaging (MRI)

MRI scans use radio waves and heavy magnets instead of x-rays to produce accurate images of the body parts. MRI scans can be very useful in looking at the brain and spinal cord and in testing whether or not melanoma has spread to these organs.

2.1.7 Positron Emission Tomography (PET) Scan

A PET scan has the potential to indicate whether the cancer has spread to lymph nodes or other areas of the body. It is most useful for the detection of advanced melanoma in humans. For this examination, radioactive sugar is injected into the body, which collects primarily in cancer cells. A special camera is then used to take an image of the radioactivity areas in the body. Following table summaries the techniques.

2.2 Machine Learning based methods for melanoma diagnosis

Researchers have used machine learning techniques for melanoma diagnosis. They have designed and developed various feature extraction methods and classification techniques for dermoscopic data. Various classification algorithms have been applied on dermoscopic datasets such as ANNs and SVMs. Additionally, various image processing approaches have been applied for pre-processing and artifacts removal for better analysis of dermoscopic images.

2.2.1 Support Vector Machines (SVMs)

In a research by Hiam et al. [2] dermoscopy image database was preprocessed, and segmented by utilizing threshold techniques. The feature were extracted using Gray Level Co-occurrence Matrix and ABCDE rule. Then classification was performed using SVM. The results demonstrated 92.1% classification accuracy.

In another research conducted by Poornima et al. [28] authors presented a method for the detection of Melanoma Skin Cancer using Image processing. The skin lesion image was used as input to the algorithm after which image enhancement was applied. The features were extracted using image processing and segmentation of skin lesion images. These extracted features were used to classify the image as melanoma or non-melanoma based on model prediction.

2.2.2 Artificial Neural Networks (ANNs)

Antony et al. [3] developed a machine learning technique to classify the skin lesions. First the image was processed, segmented and features were extracted. Finally, these features were used for classification. They have utilized a neural network model for classification task.

In a separate research which was conducted by Praveenkumar et al. [29], they presented an automated system for melanoma skin cancer using ANN. They extracted features of the segmented images using two dimensional 2D wavelet transformation. These images were used in ANN for classification. The system was able to achieve an accuracy of 97% on dataset.

2.2.3 KNN and Random Forests

Murugan et al. [23] presented a machine learning based system for classifying lesions using images dataset. The watershed algorithm was applied for segmentation which increased the precision of the automated system. They compared

the performance of famous machine learning algorithms such as k-Nearest Neighbor, Random Forests and Support Vector Machines. The results show that SVM performed well on the dataset.

Authors	**Feature Extraction**	**Classification**	**Accuracy%**
Murugan et al	ABCD	SVM	89.43
Murugan et al	ABCD	Random Foreststst	76.87
Murugan et al	ABCD	kNN	69.54
Praveenkumar et al	2D-Wavelet	ANN	97
Antony et al	ABCD	ANN	86.66
Hiam et al	GLCM, ABCD	SVM	92.1

Table 2.2: Machine Learning Methods for Melanoma Dignosis

2.3 Deep Learning based methods for melanoma diagnosis

The traditional ML methods need hand crafted features which are required for diagnosis. However, for automatic diagnosis of dermoscopic images data, manual feature extraction is difficult. User defined and crafted features based approached have limitations and performance is not optimal if these techniques are utilized. by learning problem specific features robust automated systems can be developed. Recently, deep learning is being utilized in the dermoscopy for automated analysis of images which learns useful features without feature engineering. In deep learning , features are extracted automatically, so models have robust and improved performance.

2.3.1 2D Convolutional Neural Networks and Transfer Learning

Many authors have used convolutional neural networks and different pretrained neural networks for feature extraction and dermoscopic image classification tasks. Due to the automatic feature extraction methods these algorithms are very popular among different computer vision researchers

The most related and recent work in Acral melanoma diagnosis was done in their work [39]. Where the researchers from Korea applied a CNN on a dataset which was obtained from their university hospitals. They got 750 images from different people out of which nearly half of the images were of Acral Melanoma and rest

of the images were of acral benign nevi. They used augmentation and transfer learning approach to train the CNN model. VGG16 was used as pretrained model. Their model got 80% accuracy, 92% sensitivity and 75% specificity.

A CNN was developed and implemented to investigate whether a CNN can be used to detect Acral Melanoma and help physicians in the early detection [19]. They trained a CNN on dermoscopic images of benign nevi and acral melanoma along with some intermediate tumors. The ResNet 50 was used as pretrained convolutional neural network for feature extraction and images were preprocessed to increase the accuracy. Their results showed that a CNN can help the physician in the early detection of the acral melanoma so it can be further avoided. A CNN helped the physician to decide accurately and increased the accuracy of the physician up to 8%.

The researchers [13] used transfer learning approach to detect the melanoma on PH2 which is a benchmark dataset for dermoscopic images. They showed that classification rates can be improved by removing different artifacts in the dermoscopic images. They preprocessed images by removing hairs in the images and used AlexNet as pre-trained CNN for feature extraction. They got 86% sensitivity and 94% specificity for melanoma detection.

Their [16] study was aimed at comparing the different deep learning architectures for melanoma detection. They evaluated latest deep learning architectures to detect the melanoma in dermoscopic images. They preprocess images for enhancing images quality and removing noise from images. The data augmentation technique was used to reduce overfitting. Their experiments showed that data augmentation and image preprocessing increase the classification rates significantly. They got 92% accuracy, 93% precision and 92% recall.

This [27] study was aimed to survey the data augmentation techniques for deep learning. As per their research deep learning needs a huge amount of data to learn the features. If the data is not enough then overfitting occurs and system does not work as per the requirement. They categorized the image augmentation into two groups, image manipulation and deep learning approach. The image manipulation includes kernel filters, geometric transformations, random erasing and color space transform. While deep learning included adversarial training, neural networks and Generative Adversarial Networks. These augmentation techniques avoid overfitting by enlarging small datasets.

Hosny et al. [13] presented automatic skin lesion classification system using transfer learning and pre-trained convolutional neural networks. They used three

datasets MEDNODE, DERM and International Skin Imaging Collaboration (ISIC) for testing and training in their experiments. They showed that augmentation and transfer learning can increase the classification rates. They got 87.31% accuracy. 62.02% sensitivity, 79.07% specificity and 73.07% precision on ISIC dataset with using image augmentation techniques. When they applied data augmentation techniques the results obtained were 95.91% accuracy, 88.47% sensitivity, 93.00% specificity and 92.34% precision.

Han et al. [10] used clinical images of malignant cutaneous tumors using a deep learning method. Their system was able to classify 12 different diseases of skin cancer. They used Edinburgh, ASAN, MED-NODE dataset for training which consisted of 19,398 images. ASAN, Hallym and Edinburg datasets were used for validation purposed while ASAN, MED NODE were used for testing system. They 96% accuracy on melanoma detection using ASAM dataset and 88% accuracy on Edinburgh dataset. Their system was also used to detect other cancer types such as basal cell carcinoma etc and their accuracies were between 83% and 90%.

In their experiments [5] the deep learning system outperformed the 136 out of the 157 experienced dermatologists of the hospitals in a German university. The results obtained from the system were compared with the board-certified dermatologists where the system outperformed 136 out of 157 in melanoma detection task. They used 12,378 images from the International Skin Imaging Collaboration 2019 dataset for training the deep learning system. Out of these images 100 images were used to compare the performance of the system with the human experts. For the purpose of outlier detection, they utilized Local Outlier Factor (LOF) method. The specificity of the deep learning system was 86.5% as compared with the human experts who got only 60%. The sensitivity was also 74.1% for both, doctors and deep learning system.

In this work [7] they demonstrated classification of skin lesions using single Convolutional Neural Network with end-to-end training from images. They trained a convolutional neural network on 1,29,450 clinical images out of which 3,374 were dermoscopic images containing 2032 various skin abnormalities. The images were double the amount of the images as used by the previous researchers in their research work. They used GoogleNet Inception V3 which is pre trained on ImageNet database. They utilized the transfer learning approach to increase accuracy of the system. Three divided the whole problem in three classes where CNN achieved 72.1% accuracy as compared to the two expert dermatologists who got 65.56% and 66% accuracy respectively.

In another [11] research utilized the region based CNNs to localize the keratinolytic

skin cancer on face. This system was trained on 1106886 images to detect and localize the cancer on faces. The dataset used in the research was obtained from Korean national hospitals. They found that deep learning might be utilized to diagnose the skin cancer and it can be localized using a region-based classifier. Besides this 1570 dermoscopic images were collected from 386 patients and other clinical images were used from MED-NODE dataset. They got an AUC (Area Under Curve) of 0.910 along with 89.2% sensitivity and 77.9% specificity.

Li et al. [20] proposed three methods based on deep learning for skin lesion analysis. These were: lesion segmentation, lesion classification and feature extraction. The deep learning method which compromised of two fully convolutional residual network was utilized to extract the lesion segmentation and classify the lesion. ISIC-2017 dataset of dermoscopic images was used for evaluation of the system. Data augmentation was used to increase the accuracy.

Salido et al. [32] proposed the system which preprocesses the images by removing unwanted objects such as hair and then segmenting the skin lesion automatically. After removing artifacts and noise from the images they developed a deep convolutional neural network which was applied on both images, processed and un-processed. Their experiments showed that the processed images demonstrated a great classification accuracy. They achieved 93% accuracy and sensitivity in 84-94% range

Fatma Sherif et al. [34] also used DCNNs for detection and categorizing of the melanoma. They used the ISIC challenge 2018 dataset to train the system. The system was tested and validated on the same dermoscopic images dataset. They got 96.67% accuracy for melanoma detection tasks.

Halil et al. [37] used latest deep learning algorithm for melanoma detection. YOLO (You Only Look Once) and GrabCut algorithm was used to detect and segment the melanoma affected body parts. The YOLO is used for detection purposes which has great detection results. It's very fast and computationally inexpensive. After this GrabCut algorithm was applied to segment the detected area on image. They used ISIC 2017 dataset and got an accuracy of 93.39%

To automate the melanoma detection Li et al. presented a deep learning method using an enhanced R-CNN. Mask R-CNN used Faster-RCNN for detection purposes and then Fully Convolutional Network (FCN) is used to produce the segmentation mask of the detected object. They have used ResNet 101 as baseline model with ImageNet pretrained weights on HAM1000 dataset. They achieved 78.3% accuracy on testing images from the HAM1000 dataset.

In this work the researchers diagnosed the malignant cancer of lip. The dataset was obtained from Seoul National University. They trained a deep convolutional neural network using transfer learning where they applied Inception ResNet V2 to fine tune their model. Their model showed an accuracy of 82% and sensitivity of 75% whereas the 80% specificity was achieved.

In this work [38] researchers presented an algorithm which was able to differentiate the ridge and furrow patterns of lesions on acral areas for acral melanoma detection. They used image processing techniques to classify the pattern in dermoscopic images. They utilized the width ratio of bright and dark patterns in dermoscopic images and used this information to classify the image. Nearly 300 images were used and they got 99.7% accuracy with a sensitivity and specificity of 100% and 99.1% respectively.

Following table presents an overview of deep learning based methods used for melanoma classification using dermoscopic images dataset.

Authors	**Method**	**Classification**	**Accuracy%**
Yuan et al(2017)	Conv-Deconv Network	Melanoma	93.40
Yang et al(2017)	Ridge and Furrown Pattern	Acral Melanoma	99
Codella et al(2017)	Ensemble Learning	Melanoma	85.50
Sarjar et al(2018)	Residual ConvNet	Melanoma	99
Yu et al(2018)	VGG16	Acral Melanoma	83
Julie et al(2018))	AlexNet	Melanoma	86
Li et al(2018)	CNNs	Melanoma	82
Salido et al(2018)	Deep CNN	Melanoma	93
Kassani et al(2019)	ResNet-50	Melanoma	92
Hosny et al(2019)	AlexNet	Melanoma	87
Sherif et al(2019)	Deep CNN	Melanoma	96.57
Kassani et al(2019)	Deep CNNs	Melanoma	92

Table 2.3: Deep Learning Methods for Melanoma Dignosis

array

Chapter 3

Research Gap

The prediction of melanoma is not a new research area as we have discussed in literature review section. However, the classification problem of predicting skin cancer is now shifting from two class problem of distinguishing between melanoma and non-melanoma to a far more challenging problem of predicting multiple subtypes of melanoma and non-melanoma. This research has the potential to be applied in clinical settings where a physician can not only effectively diagnose melanoma but also classify the subtypes of of and whether the patient's melanoma will progress to metastatic melanoma or not. Early and accurate melanoma diagnosis will help physicians to make proper and timely decisions for diagnosis and treatment. Based on the reviewed literature the research gap is identified as followed:

3.1 Melanoma Subtype Classification

The problem of melanoma classification by researchers is limited to two class classification problem [13]. Various researches worked on classifying melanoma and non-melanoma. However, a very limited dataset is available for further classification of melanoma subtypes and a limited work is done in this field. In this research, the objective is to classify the melanoma in its further subtype which is acral melanoma (AM).

3.2 Dermatoscopy technique

Various invasive and non-invasive techniques are being used to diagnose the melanoma [8]. For example researches have used digital images [10] in classification of melanoma vs non melanoma cancer. However, with the advancement in optical imaging technology the new dermatoscopy technique is very useful in early diagnosis of skin cancer and its subtypes and it increases the accuracy of the physician

in diagnosing the disease before time which can help in timely decision. In this research we have used images taken using dermatoscope which can increase the classification accuracy.

3.3 Automatic Feature Extraction

The traditional machine learning algorithms and techniques required manual feature extraction which is very difficult and time-consuming process [23]. The classification accuracy of machine learning algorithms depends on these manually extracted features. If these features are not accurate and cannot distinguish various classes then the classification algorithm cannot effectively classify skin lesions. This manual feature extraction results in limited accuracy. In this research we have used deep learning-based models which do not require manual feature engineering; hence the classification accuracy increases and a robust automated system is developed.

3.4 Data Preprocessing Artifacts removal

In the previous acral melanoma detection tasks, raw dermoscopic images are used in direct classification [39].However, the raw images results in a low classification accuracy. In this research we have preprocess data and different dermoscopic artifacts are removed using image processing techniques then these images are used to train the deep learning models.

3.5 Data augmentation technique

Due to infrequent occurrences of the acral melanoma, a large dermoscopic images dataset is not available unlike other skin cancer subtypes. ISIC provides dataset which contains nearly 25 thousand images of different skin cancer but it doesn't contain any images related to the further subtypes of melanoma. Recently a research was done in where they utilized 750 dermoscopic images of AM and BN. As the deep learning required a large dataset so this small dataset cannot be used to develop a robust automated system. In this research we have applied various data augmentation techniques to artificially increase dermoscopic images and additionally we used transfer learning for development of a robust automated system.

Chapter 4

Problem Statement

The research objective of research is to study and investigate techniques and approaches to detect and classify subtype of melanoma skin cancer. The objective is to make distinction between the normal benign nevus and acral melanoma skin lesions using dermoscopic images which is mostly neglected by deep learning community due to non-availability of large publicly dermoscopic images dataset. This research will assist the dermatologists in early diagnosis of AM so appropriate decisions can be made by doctors and proper treatment can be planned before time. This problem will be addressed by using image processing methods for the processing and analysis of dermoscopic images followed by deep learning techniques for melanoma subtype classification. The problem statement for this research is as follows:

There is a plethora of techniques in medical imaging for detection and diagnosing Melanoma Skin Cancer however, there is no fully automated computational framework for classifying and diagnosing subtypes of melanoma. This research presents an automatic end-to-end deep learning based computational model for automatic feature extraction, AM detection and classification using dermoscopic images dataset.

Chapter 5

Research Questions

There are some research objectives that have been targeted in this research to justify the problem statement. The first research question is to investigate melanoma and to identify various subtypes of this disease. Second research question is to study multiple skin related medical imaging techniques used in literature to analyze and study the AM. Third research question is to study and utilize preprocessing techniques used to process and analyze dermoscopic images. Fourth research question is to explore dermoscopic artifacts removal process to enhance the automated system accuracy. Fifth research question is to determine a deep learning based classifier that best performs in AM diagnosis and evaluating its performance on unseen test data. The sixth research question is to determine the how transfer learning and data augmentation might be helpful when working with small dataset. The seventh and last research question is to analyze how this automated system can be used practically in clinical settings to help the physicians in early diagnosis.

5.1 Research Question 1

What are the multiple sub types of melanoma and what are stages of this lethal cancer?

Melanoma is a lethal skin cancer caused by melanin producing cells and abnormal growth of these cells which effects a large world population. Researchers from interdisciplinary fields are interested in finding solutions to deal with this lethal disease. In computer science domain, this problem has been targeted by researchers to automate the process of diagnosis, prediction and classification of the cancer. In current research our primary target is to develop a computational framework that will help physicians to diagnose, classify and predict disease in

early stages. Melanoma is a skin related lethal cancer, which can be metastasized and affect the nearby healthy tissues of human body and other organs. For development of a better computational model, deep understanding of the melanoma is required which can be helpful in the later automated diagnosis. The studies on Melanoma revealed that it is a metastatic disease and computer science researchers have proposed different models to diagnose and treat the melanoma in early stages.

5.2 Research Question 2

What are different medical imaging techniques used to diagnose melanoma and its subtypes? What are the advantages and disadvantages these techniques? Which performs better for early diagnosis of disease?

Early stages of melanoma are diagnosed using various imaging methods such as Dermatoscopy, MRI, PET, CT, Chest X rays, Ultrasound and various others. In this research we have investigated the dermoscopic imaging techniques used for melanoma diagnosis. We have studied how physicians make diagnostic decisions by manually analyzing dermoscopic images and which features they look for such as ABCD rule. And how can physicians distinguish between various subtypes of melanoma disease by manual inspection of medical images? In order to conduct this research, dermoscopic data related to melanoma subtype was required. We needed dermoscopic imaging data from multiple cancer patients related to subtypes of melanoma. For the development of the classification model, small dataset was acquired. Therefore, we have utilized reliable histopathological confirmed dermoscopic images of acral melanoma and acral nevus.

5.3 Research Question 3

Which pre-processing techniques are used for dermoscopic images analysis and images enhancement? How to prepare these dermoscopic images for classification?

The collected raw dermoscopic images data requires pre-processing in order to remove noise and be able to perform further analysis followed by the final classification step. Pre-processing is an important step that involves various methods to transform images. This reduces noise and improve images quality. This results in increase of accuracy of computational model. Some of the pre-processing techniques for dermoscopic imaging data include remove non-lesion areas, brightness correction, resizing to a standard size for respective deep model and cropping high resolution images for computational friendly algorithm, etc. In this research, we

have investigated different techniques based on OpenCV and python tools available for image preprocessing, cropping and preparation.

5.4 Research Question 4

What are dermoscopic images artifacts removal techniques used for disease analysis? What are recent trends in deep learning for artifacts removal using image processing?

The raw dermoscopic images data which contains artifacts and these should be removed before classification. Some of the common dermoscopic images artifacts include hair, black frames, ruler markings, and other markings performed by physician to detect and localize the skin lesion manually inspect the skin lesions. These artifacts are not recognized by the deep learning algorithm hence these must be removed before final classification step. These artifacts affect the performance of the classifier. To remove these artifacts, in this research we have used various image processing techniques such as segmentation, cropping, hair removal and various thresholding techniques to remove unwanted artifacts in order to enhance the performance of automated classifier.

5.5 Research Question 5

What classification algorithms are being used to for Melanoma? Which classifier can give optimal performance? What will be the output of the classifier? How to evaluate the classifier's performance?

Our problem in this research is a supervised binary classification problem. Statistical feature extraction along with machine learning algorithms have been used by researchers for Melanoma classification. Much of the work in this field is for two class problems where a learning algorithm learns to differentiate a healthy and cancerous skin lesions. Our aim in this research is to implement classification techniques to distinguish subtypes of melanoma i.e. AM. In literature and industry, various deep learning methods are being used to classification tasks. We have implemented different types of deep learning algorithms and network architectures for classification. The input to the deep learning network was a pre-processed skin lesion and true labels. The labels represent the stage of disease such as AM and BN. Several experiments were performed by tuning hyper parameters to identify best classification model. The final classification model was then used to make predictions on unseen data. The output of the prediction is the label representing the subtype of melanoma cancer. The model performance was evaluated on unseen dataset data using various statistical classification metrics.

5.6 Research Question 6

How transfer learning and data augmentation might be helpful when working with small dataset of dermoscopic images?

Many computer vision tasks include a large dataset such as COCO and ImageNet which are good for development of a classifier as deep learning requires a huge amount of data to learn useful features. However, in medical domain a large publicly available datasets are not available easily due to privacy concerns. To overcome this issue many researchers have used transfer learning which is a learning paradigm in which features learned by a classifier are transferred to another task. To tackle small dataset issue in medical imaging domain researchers, use this transfer learning technique to overcome the data problems. In this research we have utilized various latest pre trained deep neural network architectures such as AlexNet, MobileNet V2 and ResNet-18 for classification of skin cancer lesions. Additionally, we have applied different data augmentation techniques to artificially increase the dataset for development of a robust classifier.

5.7 Research Question 7

How this automated system can be used practically in clinical settings to assist the physicians in diagnosis?

Cancer is the lethal disease worldwide. According to the recent research and survey an estimated 1.8 million new cancer cases will be diagnosed in US. A total of 606,520 cancer deaths are expected in the US in 2020 [1]. Skin cancer is a kind of cancer which affects the skin. In US more than 5 million skin cancer patients are diagnosed [1]. One of most lethal kind of skin cancer is melanoma. It contributes to the 4% of deaths worldwide [2]. Melanoma is responsible for nearly 10,000 deaths a year in US alone [1]. Out of these 10,000 deaths nearly 5% deaths are due to acral melanoma. It has largest impact on quality of life. The diagnosis and prognosis of acral melanoma is hard due to various reasons and researchers are using automated AI powered systems to help in early diagnosis This research can contribute to lower the disease burden and costs due to melanoma and its subtypes. Melanoma disease can be effectively diagnosed in early stages, so that suitable medical treatments can be made by physicians. These timely and early treatment can increase patient's survival results.

Chapter 6

Research Methodology

Our method consists of various steps and procedures including pre-processing, artifacts removal, data augmentation classification and prediction. Dermoscopic images data has been acquired from a well-known South Korean university hospital. Image processing methods have been applied to remove noise from images. Deep learning approaches have been used for classification of AM. A flow chart of methodology is given in figure

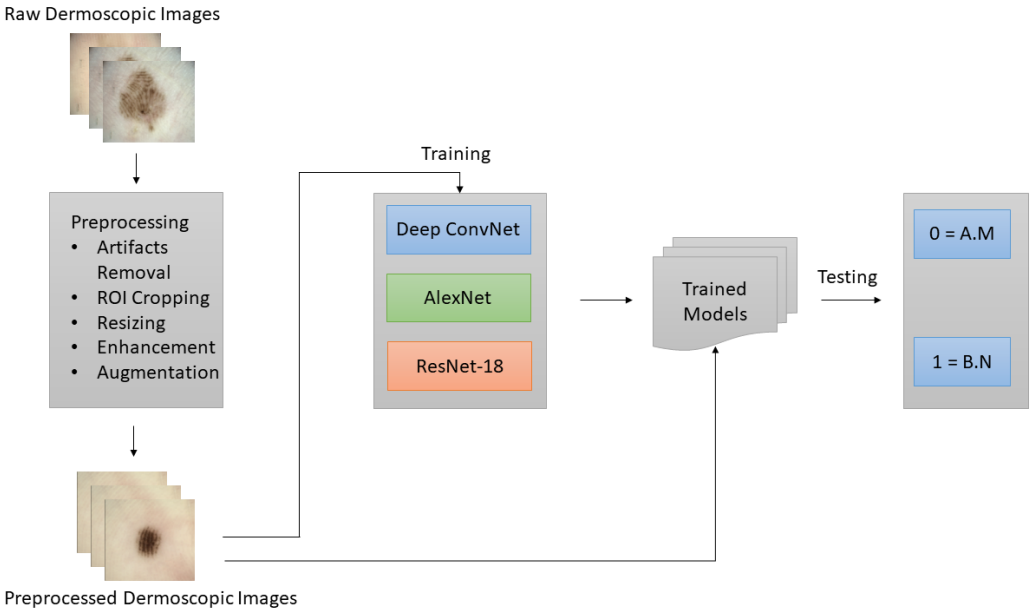


FIGURE 6.1: Proposed Methodology for AM Detection

In deep learning algorithms the features are automatically extracted and improved accuracy and performance can be been achieved. In deep learning process the first stage is to image data acquisition in case of image classification. Before training

the classifier, the raw images are preprocessed so only relevant features can be learnt by deep learning classifier. After the preprocessing these images are fed to classifier and ground truth labels are provided to learn the mapping from features to their corresponding labels. Final step in the prediction or inference in which the learning model is validated on unseen data. Because of the automatic feature extraction , this research targets the use of various deep learning algorithms for automatically extraction of feature as well as classification. The rest of the chapter discusses the details of various steps of research methodology.

6.1 Data Acquisition

Firstly, dermoscopic imaging [30] data related to melanoma disease has been acquired. Various databases provide large images datasets in melanoma and other skin cancer that contain imaging data related to melanoma in various modalities such as dermoscopic images, clinical images, and digital images.

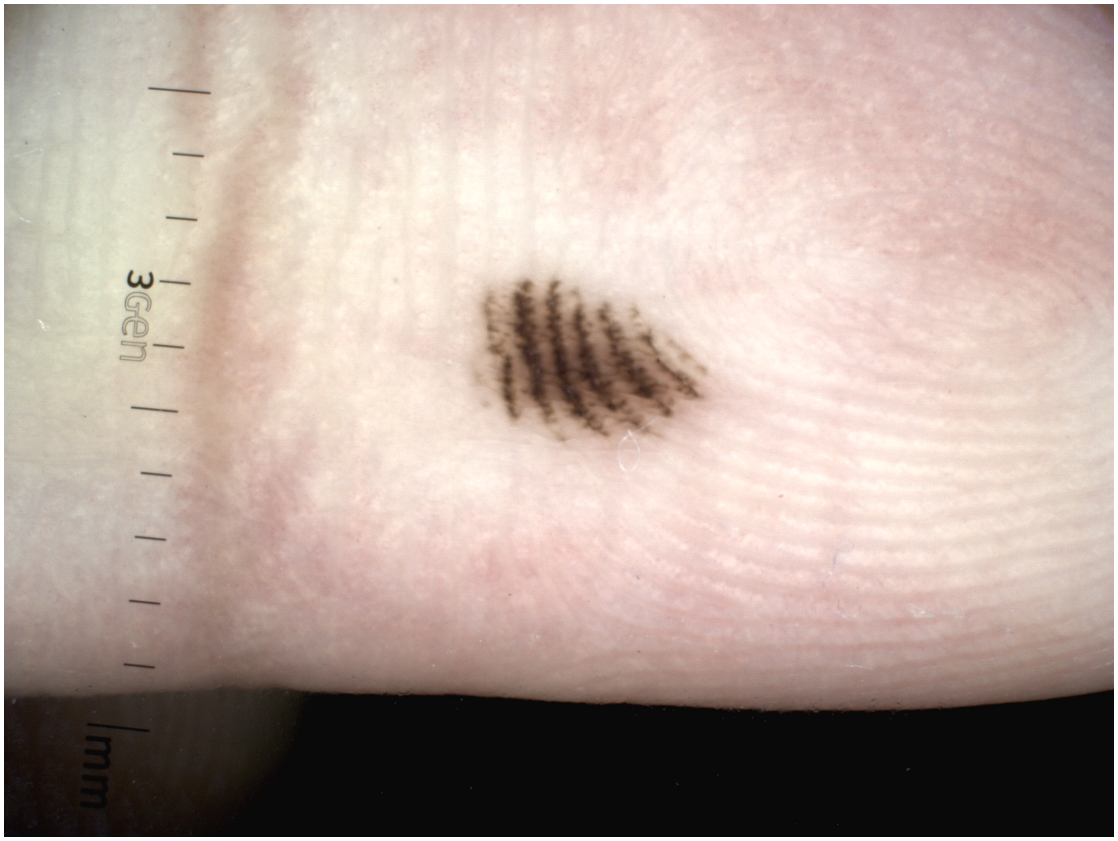


FIGURE 6.2: Benign Nevus Sample from Acral Melanoma Dataset

We utilized 724 dermoscopy images which were collected by the Severance Hospital in South Korea [39]. Out of these 724 images 350 were from acral melanoma patients and 374 images were benign nevus patients. Out of all these images

632 images were taken by the DermLite Cam and 92 images were taken through Dermlite hybrid II connected to a Japanese Nikon digital camera. All diagnoses are histopathologically confirmed.

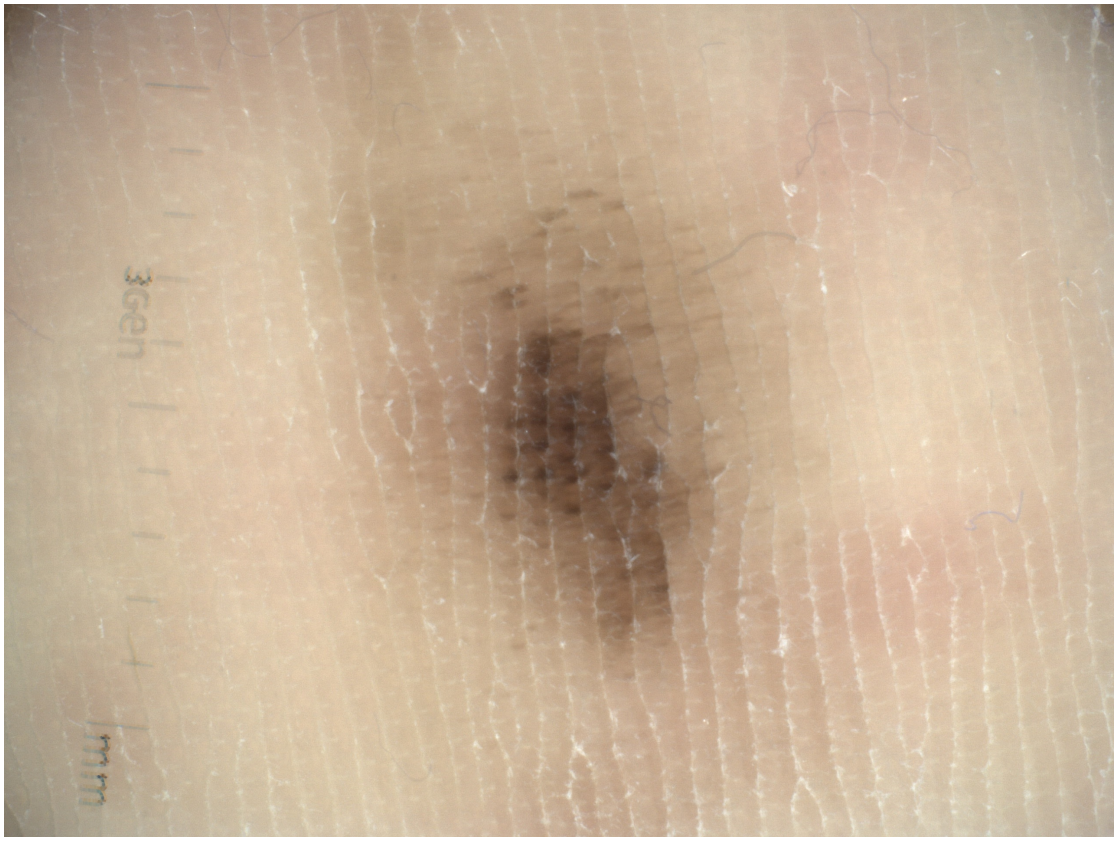


FIGURE 6.3: Acral Melanoma Sample from Acral Melanoma Dataset

6.2 Data Preparation

The dataset of images was divided into train, validation and testing where seventy percent (70%) images were used to train the algorithm and twenty percent (20%) images were used for validation while 10% images were used to test algorithm on unseen images.

6.3 Data Preprocessing

The raw dermoscopic images are high resolution images which can be very computationally expensive. To overcome this issue theses images were preprocessed so a computationally friendly algorithm can be developed. The region of interests' (ROI) were extracted using an automated python script. The images were copped from center so that skin lesion to be classified was in the image. Then these images were resized to standard size for model training.

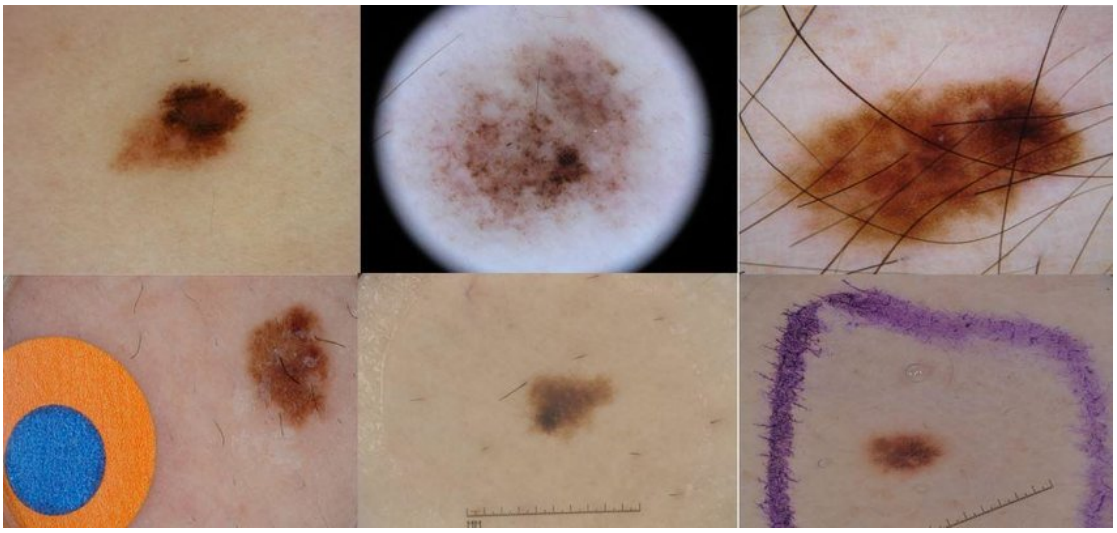


FIGURE 6.4: Dermoscopic Artifacts: Row 1: (i) normal image (ii) dark corner artifact (iii) skin hair artifact. Row 2: (i)color chart artifact (ii) ruler marker artifact (iii) ink marker artifact (Source: ISIC archive)

6.3.1 ROI Extraction

In this first step to develop a computational friendly algorithm the ROI of the skin lesion was detected. As most of the ROIs were in the center of a high-resolution image so these images were used to detect and extract the ROI using automated python script [20].

FIGURE 6.5: ROI Extraction Steps

6.3.2 Artifacts Removal

The raw dermoscopic images contain several artifacts which may lead to poor performance [40]. The deep learning models automatically learns the features from input data and if the data is clean the features learned by the model will

Model	Accepted Input Size	Input Size Used
ResNet-18	224	224
AlexNet	256	256
MobileNet V2	128,224	224
Deep ConvNet	128,224	224

TABLE 6.1: Image Input Size for Deep Learning Models

be good for classification otherwise the performance decreases when the input images are not clean. Some of the artifacts found in dermoscopic images are black frames in images, ruler marking, hairs etc. Our data comes from acral areas of body where there is not much hair so we did not face any problem caused by the hairs but some images contained black frames and ruler marking on the ROI of skin lesion image. So, these were removed using cropping and image processing techniques. Image inpainting algorithms found in popular computer vision library i.e. OpenCV were used to remove the artifacts on images. The unwanted pixels in skin ROI were replaced by the neighboring pixels of image ROI i.e. AM or BN. This process helped in further cleaning of images.

6.3.3 Cropping

The detected ROI was cropped so that the non-lesion area of skin can be removed from images for a better classifier. This was done using a small python script which automatically cropped all the images to obtain a clear ROI image for classification of the skin lesion. When we used the full images for classification the results were not satisfactory so we used approach as practiced by [20]. So, we first detected ROI and used python for automatically cropping of the ROI skin lesion. These images were fed to classification algorithm. add Image

6.3.4 Resizing Images

To decrease the image size and the cropped patches of images were resized to a standard size. We have utilized different pretrained deep learning models which require standard input size to input layer of classifier. For example, ResNet required input size to be a resolution of 224 by 224 where height and width of image should be 224. Without resizing these images, we cannot use transfer learning. So, for the purpose of transfer learning we resized these images to a standard acceptable input size for various pre trained deep learning models. Following table summarized the image size used for various deep learning models. These resized images were fed to deep learning classifier for classification.

6.4 Data Augmentation

The original dataset contains only 724 images belonging to two classes of AM and BN which are not enough for a robust deep learning model [27]. For development of high performance and accurate classifier a large dataset is required. To overcome this issue, we applied different image augmentation techniques. The images were rotated to different angels and flipped to artificially increase number of images. We used a rotation of 90, 180 and 270 degrees to generate new images. Besides these images transformation we also flipped image upside down and a left to right flip was also applied for increasing the dataset size. After applying different transformations, we got four times more data than the original dataset. Some of the sample images are shown in following figure.

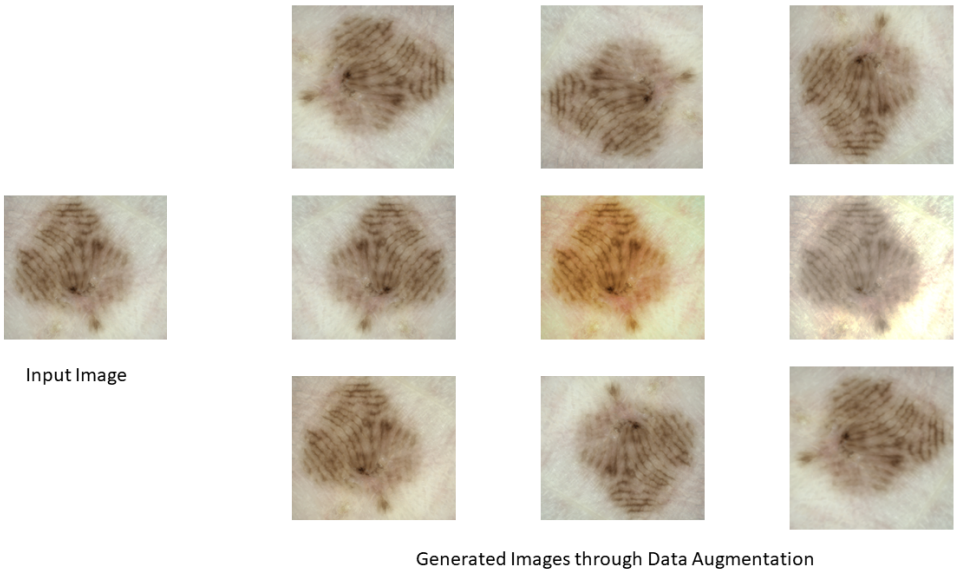


FIGURE 6.6: Images Generated through Data Augmentation

Following table summarized the original dataset and after the transformations were applied to increase the dataset.

Dataset	**Total Images**	**AM**	**BN**
Original	724	350	374
Augmented	4344	2100	2244

TABLE 6.2: Data Augmentation Details

6.5 Deep Learning Models for AM diagnosis

We used deep learning algorithms for feature extraction and classification of AM and BN. We performed our experiments on three deep learning pretrained ConvNet architectures to train and evaluate on our preprocessed datasets. First architecture is AlexNet t, second deep learning network used for this research is ResNet18 and third architecture used for this research is MobileNet v2. We have also designed and developed a custom CNN in Keras and Tensorflow for scratch and trained it on preprocessed dataset. The architectures are discussed in detail in this section.

6.5.1 AlexNet

AlexNet is the CNN, designed by Alex Krizhevsky [17] in 2012. AlexNet authors used this network and competed in the ILSVRC [6] in 2012. The model achieved a top-5 error of 15.3%, which was 10.8% lesser than the network on second position. The results of the network showed that the depth of the network was important for high performance on Image Classification. Although this was expensive but utilizing GPUs made it feasible.

The network architecture contains total eight layers five of which are convolutional layers and three fully-connected layers and a SoftMax layer for classification. This architecture used some special features which helped in improved performance of deep ConvNets. It used Max-Pooling after convolutional layers and utilized Dropout regularization. Additionally, to further train the deep network effectively the non-linear Rectified Linear Unit was used instead of tanh activation function. This architecture has total 60 million parameters and contains 65000 neurons in the overall architecture.

Some of the new features added in this network which enhanced performance of this network are discussed in following section.

6.5.1.1 RELU

AlexNet uses ReLU as a neuron output activation function. It has been shown in various studies that CNNs which are trained using gradient descent, are trained faster than using tanh. The equation represents ReLU based function.

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

6.5.1.2 Multiple GPU Suppourt

AlexNet allows for multi-GPU training. This is achieved by putting half of the model's neurons on one GPU and the other half on another GPU. This can decrease the training time and large deep networks can be trained much faster.

6.5.1.3 Overlapping Pooling

AlexNet used overlapping pooling which resulted in a drop of 0.5% error. This experiment showed that overlapping pooling can also be used to handle overfitting of large

neural networks. Although CNNs traditionally use simple pooling layers outputs with no overlapping.

6.5.1.4 Overfitting

To handle overfitting in this model the AlexNet model utilized the Dropout and Data augmentation. This results in the better network which has great generalization ability to the unseen data. In this Dropout technique each hidden neuron's output zero with a certain probability (0.5 in this case). The neurons having zero output will not contribute in forward propagation or back-propagation. The architecture of original AlexNet is shown in fig below.

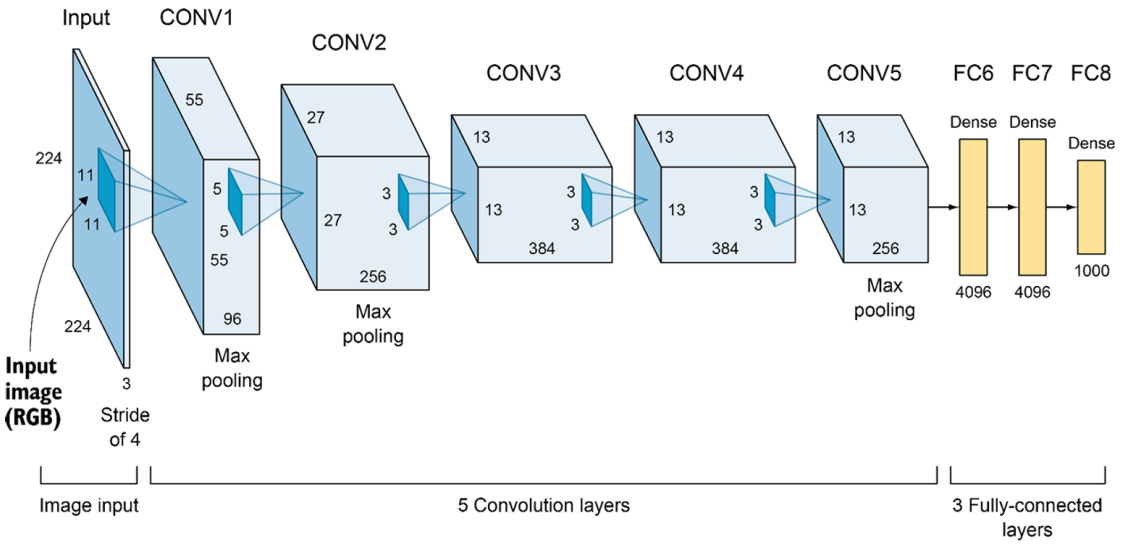


FIGURE 6.7: Original AlexNet Architecture

6.5.2 ResNet

Residual Network was proposed by the Kaiming et al. [12] back in 2016. A residual learning technique was developed to train deeper networks. As the deeper networks are hard to train so this residual method was proposed to train deep networks. The results indicate that the residual learning based deeper networks achieved better optimization and high accuracy because of the network depth. ImageNet dataset was used to train residual networks of varied depth with 152 layers. ResNet won 1st place in the ILSVRC 2015 classification competition having a top-5 error rate of 3.57%. It has been proved network depth is important to achieve better performance. However, very deep networks are difficult to train and when network depth is increased it does not result in better performance and learning. This is caused by the vanishing gradient and exploding gradient problem. When deep neural network start converging then the accuracy doesn't increase. This problem was solved by introducing a residual function. This is visually shown in following figure .

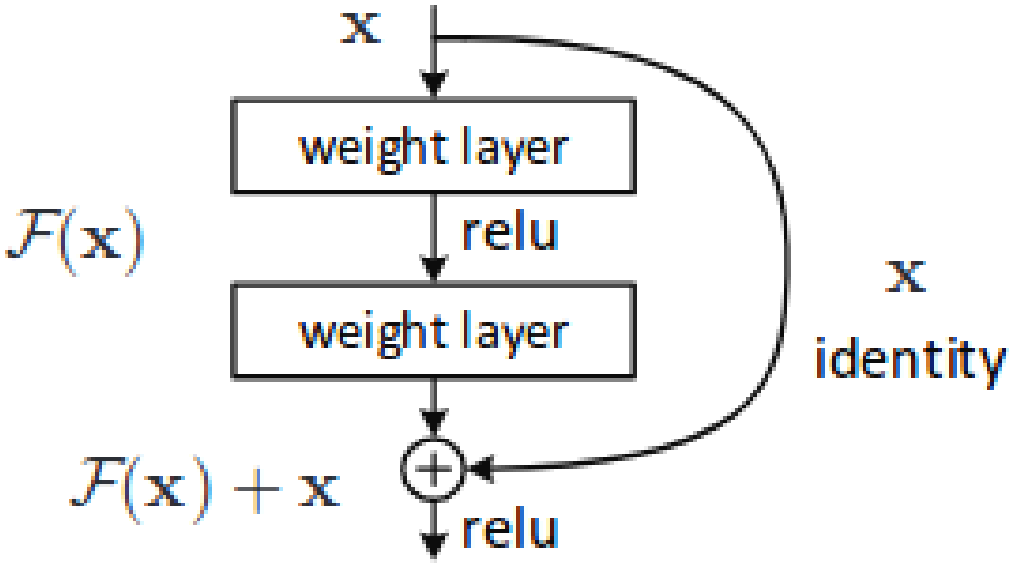


FIGURE 6.8: ResNet Building Block

In plain neural networks , network components are combined to directly learn the predicted mapping. However, in residual networks, layers are stacked in order to learn residual mapping. Figure 6.6 shows a building block of a residual learning. The mapping function, referred to as H(x), is equipped with a stacked layer where x is input. The concept of residual learning is hypothesized as if a complex mapping function could be estimated asymptotically by many nonlinear layers, then the residual function denoted as F(X) can be calculated asymptotically. The function is given by:

$$H(x) = F(x) + x$$

And residual function is:

$$F(x) = H(x) \quad x \quad (6.4)$$

Stacked layers directly learn the residual function F(x) instead of learning the original function H(x). This approach assumes that the residual mapping function is simpler to optimise than the initial mapping function. The original mapping function is determined after approximation of the residual function as H(x) = F(x) + x. This mapping function F(x)+x is realized in a feedforward neural network as a residual shortcut relation and performs element wise addition. Additionally, these networks can be trained easily using backpropagation based on SGD.

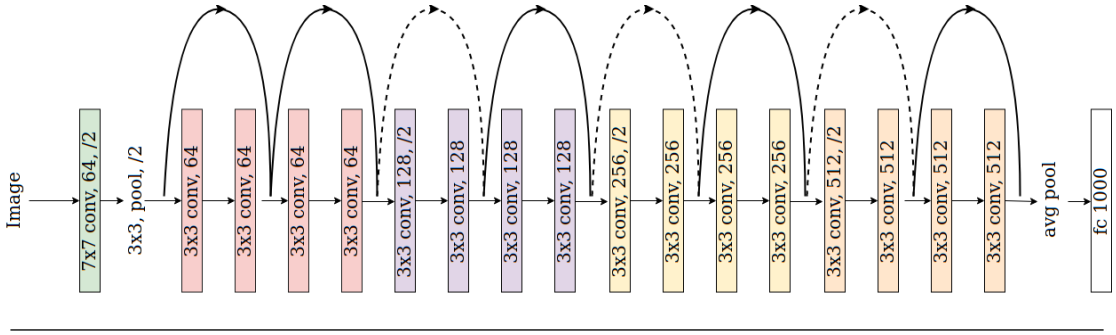


FIGURE 6.9: Original ResNet-18 Architecture

We used the 18-layer ResNet in this analysis. The network architecture is shown in Figure 6.7. The network comprises 18 layers, 17 are convolutionary layers and a fully connected layer. The convolution layers perform 3x3 filtering and the network is designed to allow the layers to have the same filters if the output feature map is the same size. Filters are doubled in the layers, however, if the output function map is halved. Phase 2 Convolution layers conduct sampling down. Finally , the average pooling is accompanied by a fully connected layer with a 2-way SoftMax layer. Residual shortcut links are put between layers in the network. There are two types of connexions. The first type of connections, denoted by solid lines, is used when the input and output dimensions are identical. The second types of connections, shown by dotted lines, can be used when the dimensions are increasing. This type of connexions still performs identity mapping but with step 2 zero padding for extended dimensions. The 18-layer ResNet contains 1.8 billion FLOPs.

6.5.3 MobileNet V1

MobileNet is more suitable architecture for mobile and embedded vision applications. MobileNet is a convolutional network architecture developed by by analysts at Google in April 2017 [14]. A few features make MobileNet awesome' small, fast, remarkably accurate, easy to tune for resource with accuracy. MobileNet requires less and can be trained very quickly as compared to traditional CNN, as well as it is simple to tune and provides more accurate classification results comparing to CNN. In this study we used image input size of 224*224* 3 with three channels RGB.

In MobileNet the traditional convolutional layers are replaced with depth wise separable convolution. The job of the convolution layer is divided into two sub-tasks: first, there is a depth-wise convolution layer that filter the data, and then a 1-by-1 (or point-wise) convolution layer that incorporates the filtered values to create new features.

The depthwise and point wise convolution together forms “Depthwise Separable Convolution”. These are more like traditional layers but are faster than the traditional layers.

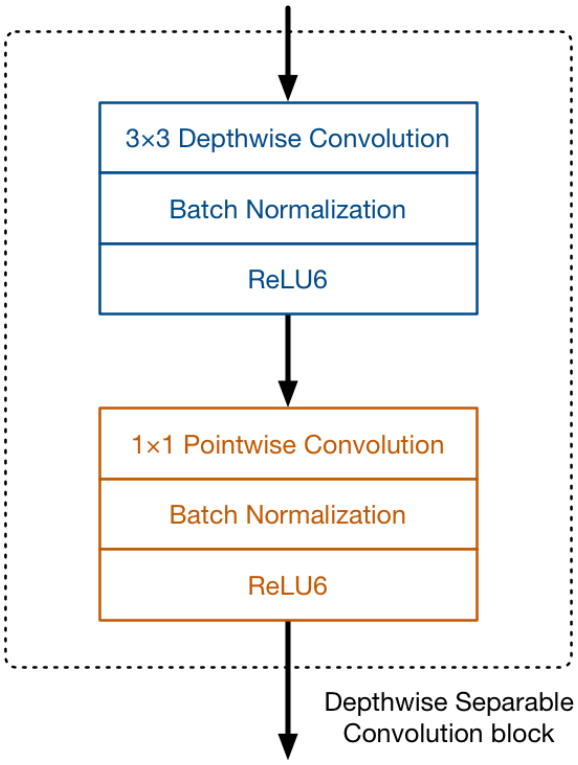


FIGURE 6.10: Depthwise Separable Convolution Block - MobileNet V1

MobileNet 's complete architecture includes a 33 convolution as the first layer, followed by a 13-fold convolution of the above building block. There are no pooling layers between these separable blocks of depth.

Conversely, a few of the depth-wise layers have a step of 2 to reduce the spatial dimensions of the data. If this occurs, the number of output channels is also multiplied by the corresponding point-wise sheet. If the input image is 224*224*3, the output of the network is 7*7*1024.

Like other modern architectures, convolution layers are accompanied by batch normalization. The activation feature used for MobileNet is ReLU6. It's just like regular ReLU, but prevents the activation from getting too large:

$$y = \min(\max(0, x), 6)$$

In MobileNet, there is indeed a global average pooling layer at the top , followed by a fully-connected classification layer or equivalent to 1 by 1 convolution, and a SoftMax layer.

6.5.4 MobileNet V2

MobileNet V2 [33] also utilizes depthwise separable convolutions, but its modified main building block have changed like this:

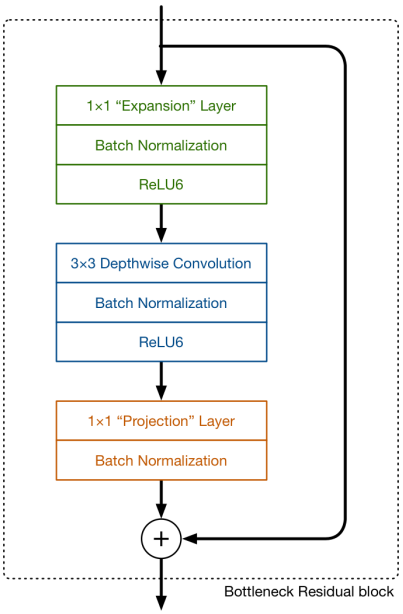


FIGURE 6.11: Bottleneck Residual Block - MobileNet V2

There are three convolutionary layers in one block. The last two are a depth-wise convolution that philtres the inputs, followed by a 11 point-wise convolution layer. But this 11 layer has a different task now.

The point-wise convolution of MobileNet V1 either holds the number of channels the same as doubled. MobileNet V2 makes the number of channels smaller. This layer is also often called a projection layer, since it projects high dimensional data onto a small dimensional tensor.

The first layer is the newly implemented block technique. This is a 11 convolution, too. It extends the channels in the data before heading to the deep convolution. This expansion layer thus includes more output channels than input channels. It's the reverse of the layer of projection. The expansion of data is regulated by the expansion factor. It's one of those hyperparameters. The default expansion factor for this is 6.

The second addition to the MobileNet V2 building block is the residual link that functions just like ResNet. This prevents the gradients from being too small or too high. Each layer has a batch normalisation and an activation feature for ReLU6. However, the activation function is not applied to the projection layer output. The complete MobileNet V2 architecture consists of 17 new building blocks consecutively followed by a standard 11 convolution, the global average pooling layer and a classification layer

6.5.5 Deep ConvNet

For further experimentation we developed a custom CNN from scratch to do know how a CNN from scratch will perform on the smaller dataset. In this section the ConvNet is

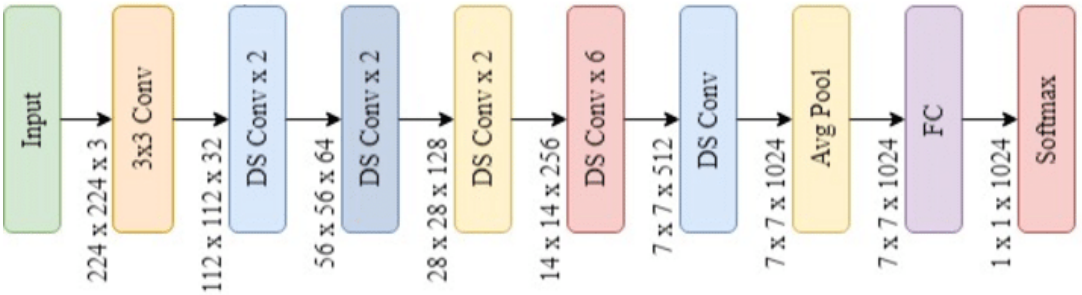


FIGURE 6.12: Original MobileNet Architecture

discussed in detail.

CNN is the variants of neural networks [18]. It derives its name from the type of hidden layers it contains. The layers of a CNN consist of convolutional layers, pooling layers, fully connected layers, and normalization layers. A typical CNN for image classification can be shown in the figure

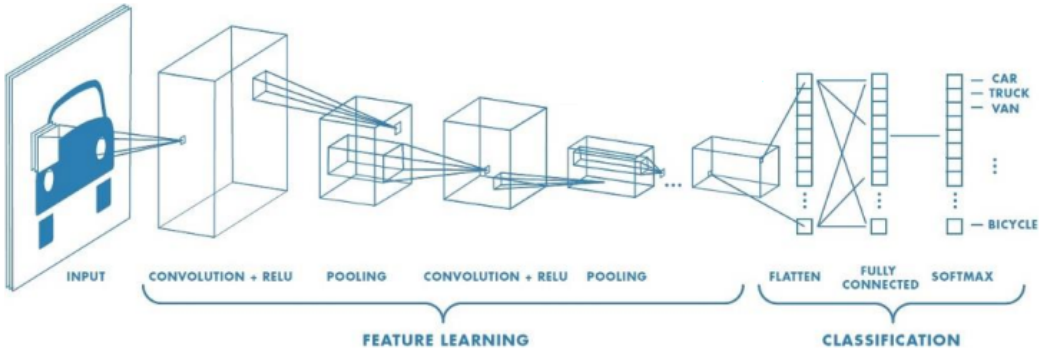


FIGURE 6.13: A typical CNN for Object Recognition

The features from images or data are extracted using convolution and pooling operations and ReLU is used as activation function. These layers are followed by the fully connected layers and classification layer.

6.5.5.1 Convolution Layer

Mathematically, the convolution procedure is the summing up of the element-wise product of two matrices. Convolutional layers transform the input to the philtre and transfer the output to the next layer, which is close to the neuron response in the visual cortex. Each convolutionary neuron processes data for its receptive field only. Completely linked neural feed-forward networks can also be used for the extraction of features, but it is not realistic to extend this architecture to high-dimensional data such as images.There will

have to be a large amount of neurons. The convolution operation solves this problem by reducing the number of free input data parameters that allow the network to be deeper using fewer parameters. The convolution operation is given by:

equation

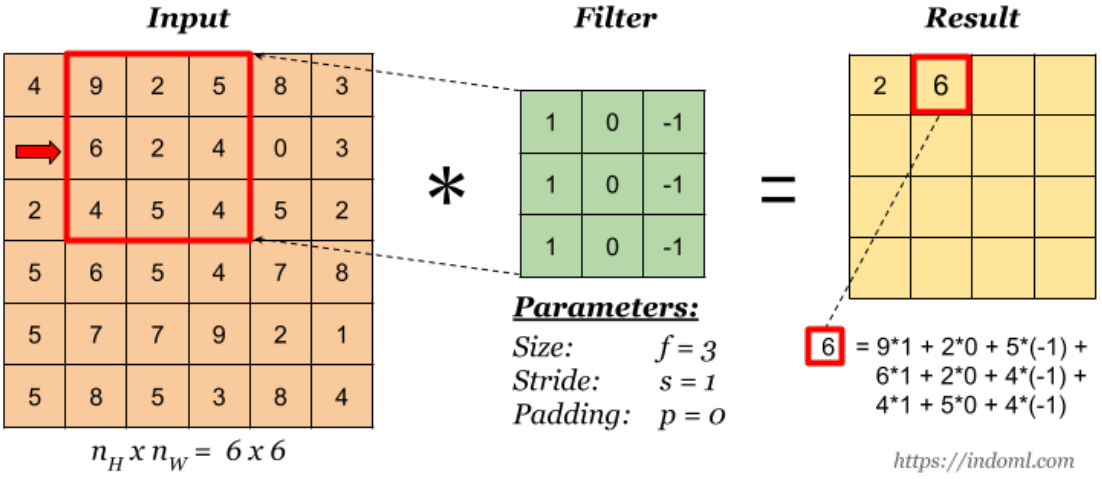


FIGURE 6.14: Convolution Operation

6.5.5.2 ReLU

ReLU stands for rectified linear unit. ReLU is most widely used neural network activation function, it is given by: f(x) = max (0, x)

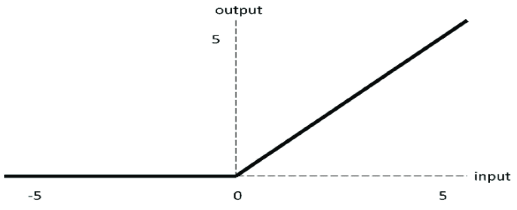


FIGURE 6.15: Rectified Linear Unit (ReLU)

6.5.5.3 Pooling Layer

The pooling layer looks at large areas of the image and collects the aggregate statistic of each region to make the n / w invariant for local transformations. The two most common aggregate functions are max pooling and average pooling,

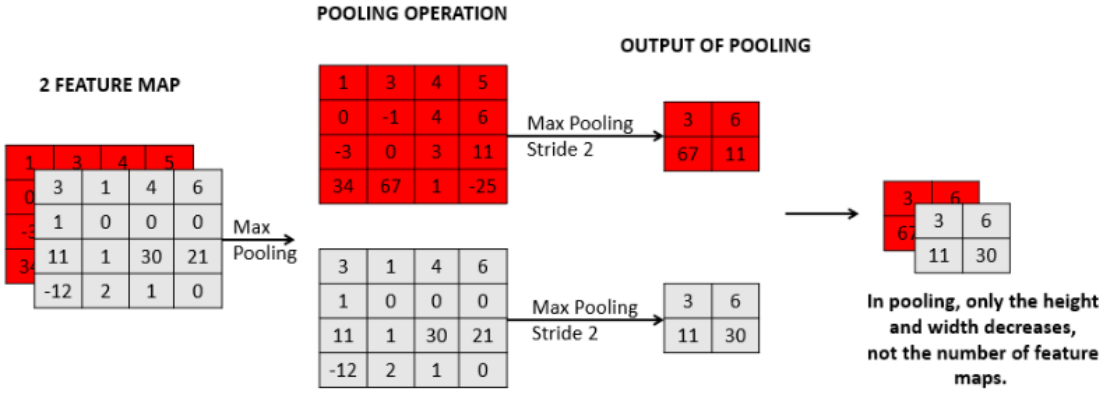


FIGURE 6.16: MaxPooling 2D Operation

6.5.5.4 Fully Connected Layer

The output of the pooling layer is flattened in order to form a large vector. Consists of a SoftMax activation feature. This layer generates a probability ranging from 0 to 1 for each classification mark.

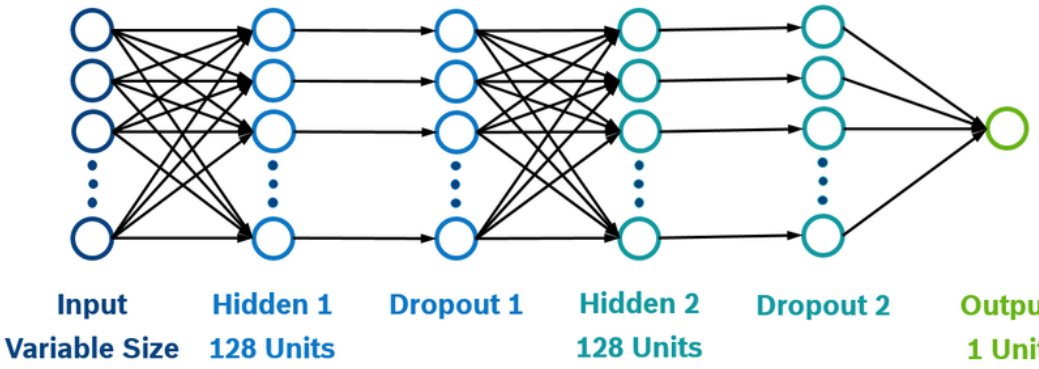


FIGURE 6.17: Fully Connected Layer

6.5.6 Custom 7 Layer Deep ConvNet

In this study we developed a Custom CNN with seven layers. Five of these layers are convolutional layers and 2 fully connected layers. We used MaxPooling after every convolutional layer. The outputs from convolution layers were normalized using Batch Normalization layers. Non-Linear activation function ReLU was used as activation function.

Input to this network is a 224*224*3 RGB image. The first layer applies eight convolutional filters with a size of 3*3. After this convolution operation Batch Normalization is applied to normalize the input. This serves as regularization sometimes and increases network learning speed. ReLU is applied after this followed by 2*2 MaxPooling. This block is repeated four times. However, we increase the number of convolutional filters to 16, 32, 64, 128 in successive layers. After the final network block the output from last convolutional block is flattened to form a 256 neuron fully connected layer followed by a 2 neuron fully connected layer. At the final layer SoftMax layer is applied to classify the inputs into predicted labels. Dropout is applied to reduce overfitting where 0.3 is used as probability value. Data augmentation is also applied to handle overfitting in the network. Our network contains total 919,346 parameters. SGD was used as optimization algorithm and Categorical Cross Entropy was used as Loss Function to be minimized during the training process. The architecture of the network is shown in following figure. image

6.5.7 Evaluation Metrics

For evaluating machine learning models in medical datasets different metrics are used such as Precision or Positive Predicted Value (PPV), Recall or True Positive Rate (TPR), Area under Receiver Operating Characteristic (ROC) curve and F1-Measure. In classification problem if dataset is balanced then accuracy can also be used for evaluating an algorithm. These have been originated from the field of Information Retrieval. They are often used binary classification problem where classifier predictions are summarized in confusion matrix. The confusion matrix contains the actual and predicted results of a classifier.

6.5.7.1 Accuracy

Accuracy [24] is calculated using following formula

$$Accuracy = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN}}$$

In this formula TP refers to true positives, FP refers to False Positives, TN refers to true negatives and FP refers to False Positives. This is ratio of true predictions and false predictions by a classifier. As in our case the dataset is balanced and no class imbalance issue is occurred so accuracy is used as one of our key evaluation metrics.

6.5.7.2 Precision (PPV)

Precision [24] is the ratio of positive predicted values that were actually positives. In other words, it is number of positive predicted cases which were actually positive in the

dataset. It is also referred as Positive Predictive Value or PPV. Precision is calculated using following formula:

$$Precision = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

6.5.7.3 Sensitivty/Recall/True Positive Rate

Recall [24] is the ratio of actual positive cases that were correctly predicted as positive. It is referred as True Positive Rate or Sensitivity and it is calculated using following formula:

$$Recall = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

6.5.7.4 F1 Measure

It is the harmonic mean of precision and recall of a classifier [24]. It is defined using following formula:

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

6.5.7.5 Area under ROC Curve

Receiver Operating Characteristics (ROC) were derived from Signal Processing in Medical Sciences [24]. It has become an assessment criterion by comparing the true positive rate (TPR) and the false positive rate (FPR). It takes values between 0 and 1 and allows you to compare the output of the classifier. If the value is closer to 1 it means the model is perfect and if it is closer to 0 the output of the model is not good. Lower classification bound is 0.5 in realistic situations where the classifier can not discriminate. It is calculated using following formula:

$$AUC = 1 - \frac{\text{FPR} + \text{FNR}}{2}$$

Chapter 7

Experiments and Results

This research was conducted to classify the subtype of melanoma using dermoscopic images dataset. We have used dermoscopic images of AM and BN from Yonsei University Hospital. We applied various preprocessing algorithms for these images. The details of the methods and deep learning algorithms have already been discussed in previous sections. In this section we have provided the experiments details and results we achieved using different deep learning architectures.

7.1 Dermoscopic Images Dataset

For the training and validation of the deep learning models the dermoscopic images data was used. The images were preprocessed augmentation were applied using various image transformation methods. This resulted in a preprocessed and large dataset which was four times than the original dataset. This dataset was divided into 70% ,20% and 10% for training, validation and testing respectively. The deep leaning models were trained using transfer learning technique and evaluated on unseen test dataset.

To train deep learning networks with this data, we need Graphical Processing Unit (GPU) for faster computation. For this, we utilized a cloud GPU service by Google known as Google Colab [4]. Google Colab is a free online cloud-based Jupyter notebook environment that allows us to train large and complex deep learning models on free GPUs, and TPUs. It provides a single 12GB NVIDIA Tesla K80 GPU with 25GBs of RAM. Keras and PyTorch [26] were used as deep learning framework. We used python as programming language. Following table summarizes the hardware and software requirement for our experimental setup.

The details of the experiments have been discussed in detail in the following section.

Software/Hardware Requirement	**Specifications**
GPU	12GB NVIDIA Tesla K80
Deep Learning framework	PyTorch, Keras, Tensorflow
Programming Language	Python 3.7.5
Image Processing Library	OpenCV 4.1.1

TABLE 7.1: Software & Hardware Requirements

7.2 Experiments with ResNet-18

We finetuned ResNet which was originally trained on ImageNet database. The original network was trained to classify the 1000 categories of images from ImageNet. We finetuned the last convolution block of the network and trained it on our dataset along with the classifier to classify our images. The last layer was removed and a classifier was trained using following parameters. (I)FC layer with 256 units followed by ReLU activation (II) Dropout 0.4 (III)Final classification layer with 2 units and SoftMax Classifier.

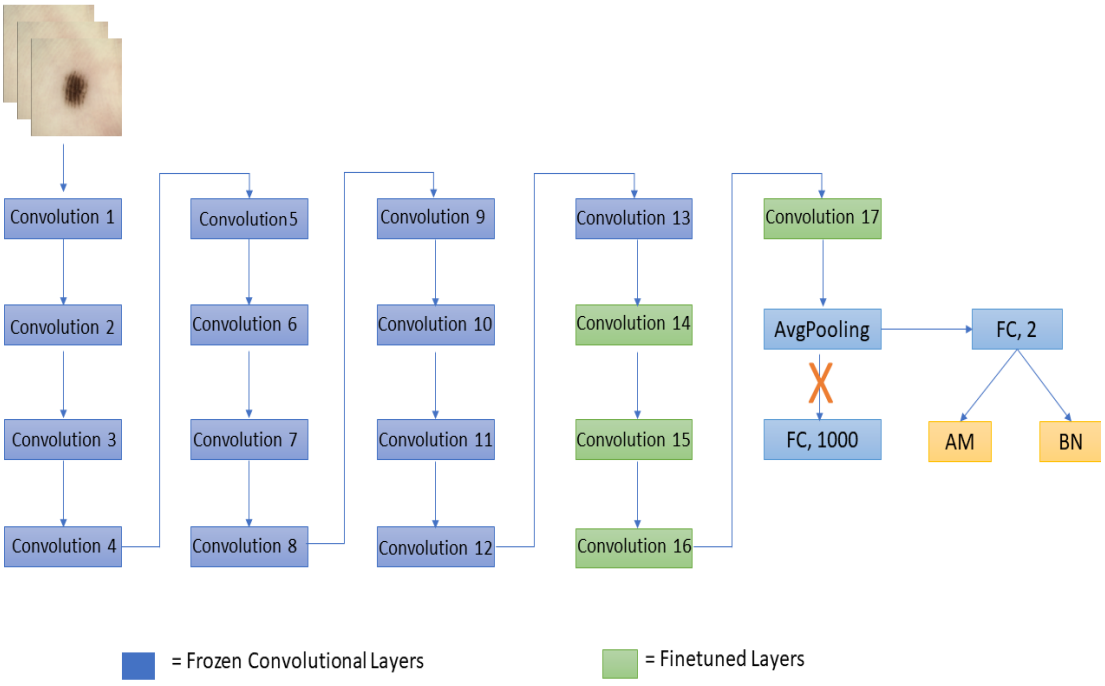


FIGURE 7.1: Modified ResNet-18 used in this study

The dermoscopic images were used to finetune our ResNet model classify different subtypes of Melanoma. Images were resized to 224x224 as ResNet accepts the standard size of this if we want to finetune the ResNet on our dataset. From these images 70% were used for training and 20% for validation and rest 10% was used for testing on

unseen data. Initial learning rate was set to 0.001. Stochastic gradient descent (SGD) with momentum was used, as optimizer where momentum was set to 0.9. Network was trained up to 25 iterations. A batch size of 32 was used during training. Following Table summarizes configuration for ResNet Classifier.

Parameters	**Values**
Input Size	224 * 224 * 3
Batch Size	32
Train/Test split	70%,20%,10%
Loss Function	Cross Entropy
Optimizer	SGD with Momentum
Learning Rate	0.001
Momentum	0.9
Epochs	25

TABLE 7.2: Training Parameters used for ResNet-18 Model

The results of the classifier are described in the following table

Class	**Accuracy**	**Precision**	**Recall**	**AROC**	**F1-Score**
AM	0.9747	0.9801	0.9701	0.9742	0.9701
BN	0.9448	0.9710	0.9801	0.9734	0.9801
Average	0.9747	0.9720	0.9703	0.9746	0.9702

TABLE 7.3: Classification Results for ResNet-18 Model

We achieved average classification accuracy of 97.47% and testing loss of 12.2% using ResNet-18. The training and validation loss and accuracy have been shown in the following figures 7.4 and 7.5. Finally, the area under the ROC curve and Confusion Matrix have been visualized in the following figures.

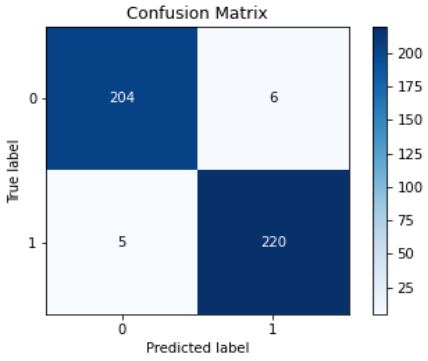


FIGURE 7.2: Confusion Matrix for our modified ResNet-18 Model

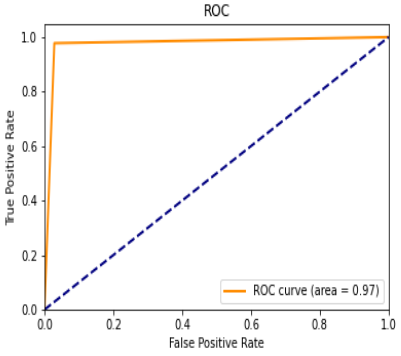


FIGURE 7.3: Area under ROC Curve for Modified ResNet-18 Model

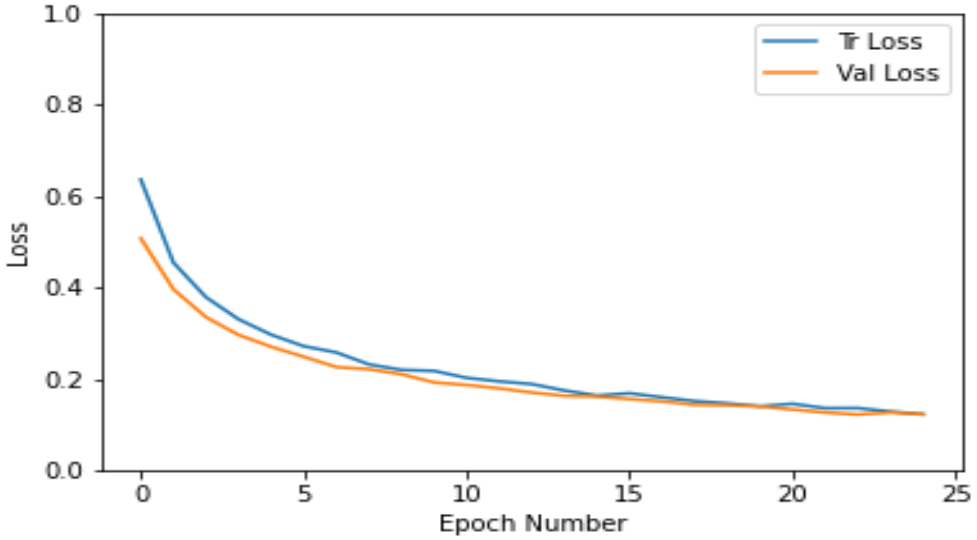


FIGURE 7.4: Training and Validation Loss Trends for modified ResNet-18 Model

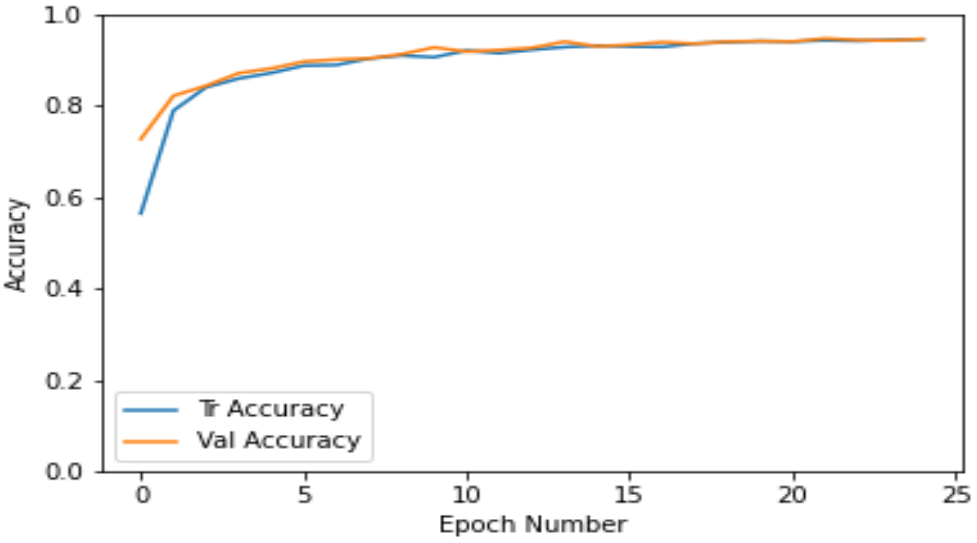


FIGURE 7.5: Training and Validation Accuracy Trends for modified ResNet-18 Model

7.3 Experiments with AlexNet

In second experiment we finetuned AlexNet which was originally trained on ImageNet database. The original network was also trained to classify the 1000 categories of images from ImageNet. The last classification layer was removed and a custom added classifier was trained using following parameters. (I)FC layer with 256 units followed by

ReLU activation (II) Dropout 0.4 (III)Final classification layer with 2 units and SoftMax Classifier.

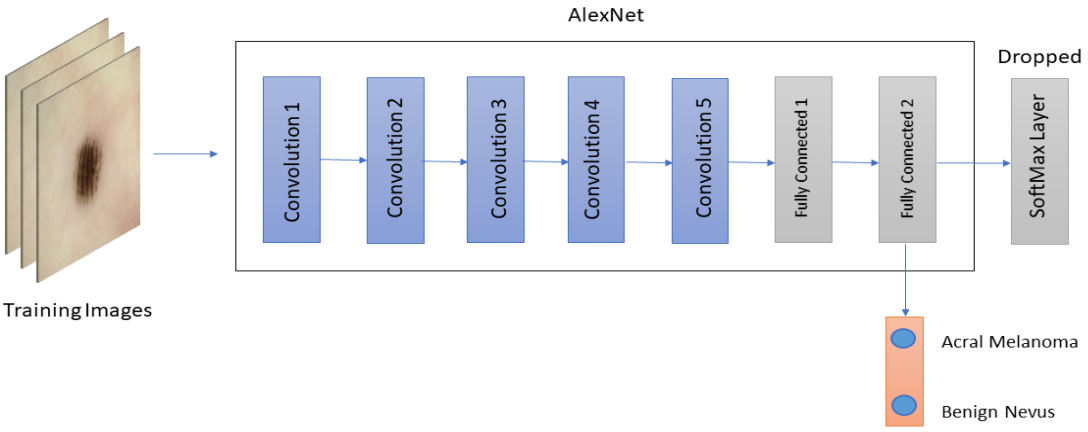


FIGURE 7.6: Modified AlexNet used in this study

The preprocessed dermoscopic images were used to finetune our AlexNet model classify different subtypes of Melanoma. Images were resized to 256x256 as AlexNet accepts the standard size of this if we want to finetune this model on our dataset. From these images 70% were used for training and 20% for validation and rest 10% was used for testing on unseen data. Initial learning rate was set to 0.001. Stochastic gradient descent (SGD) with momentum was used, as optimizer where momentum was set to 0.9. Network was trained up to 25 iterations with these hyperparameters. A batch size of 32 was used during training. Following Table summarizes configuration for AlexNet Classifier.

Parameters	**Values**
Input Size	256 * 256 * 3
Batch Size	32
Train/Test split	70%,20%,10%
Loss Function	Cross Entropy
Optimizer	SGD with Momentum
Learning Rate	0.001
Momentum	0.9
Epochs	25

TABLE 7.4: Training Parameters used for modeified AlexNet Model

The results of the classifier are described in the following table

Class	Accuracy	Precision	Recall	AROC	F1-Score
AM	0.9619	0.9948	0.9190	0.9573	0.9554
BN	0.9555	0.9294	0.9955	0.9573	0.9613
Average	0.9587	0.9621	0.9573	0.9587	0.9584

TABLE 7.5: Classification Results for AlexNet Model

We achieved average classification accuracy of 95.87% and testing loss of 12.3% using this AlexNet pretrained ConvNet. The training and validation loss and accuracy have been shown in the figures 7.9 and 7.10. Finally, the area under the ROC curve and Confusion Matrix have been visualized in the following figure.

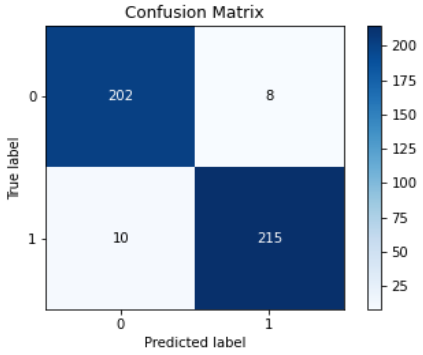


FIGURE 7.7: Confusion Matrix for Modified AlexNet Model

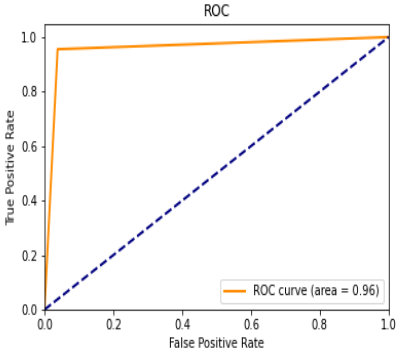


FIGURE 7.8: Area under ROC curve for Modified AlexNet Model

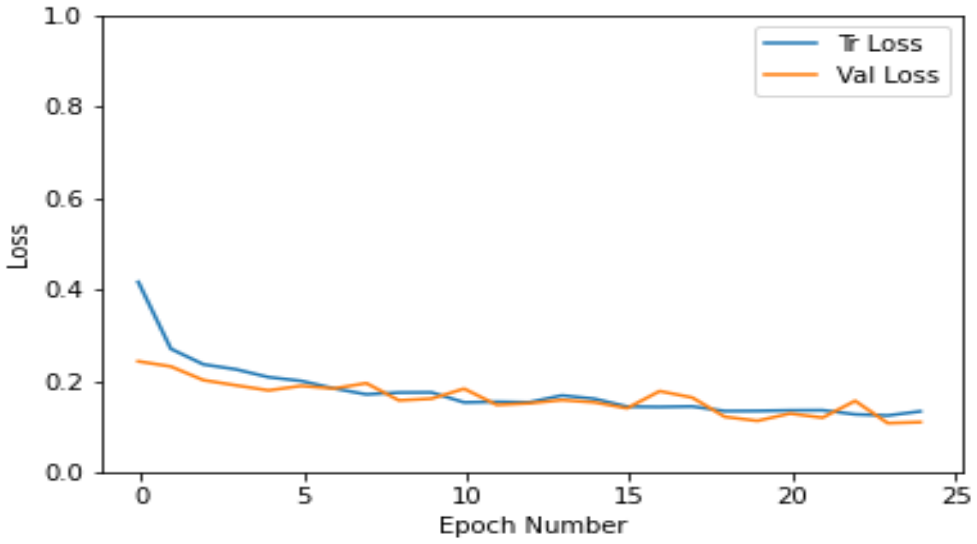


FIGURE 7.9: Training and Validation Loss Trends for modified AlexNet Model

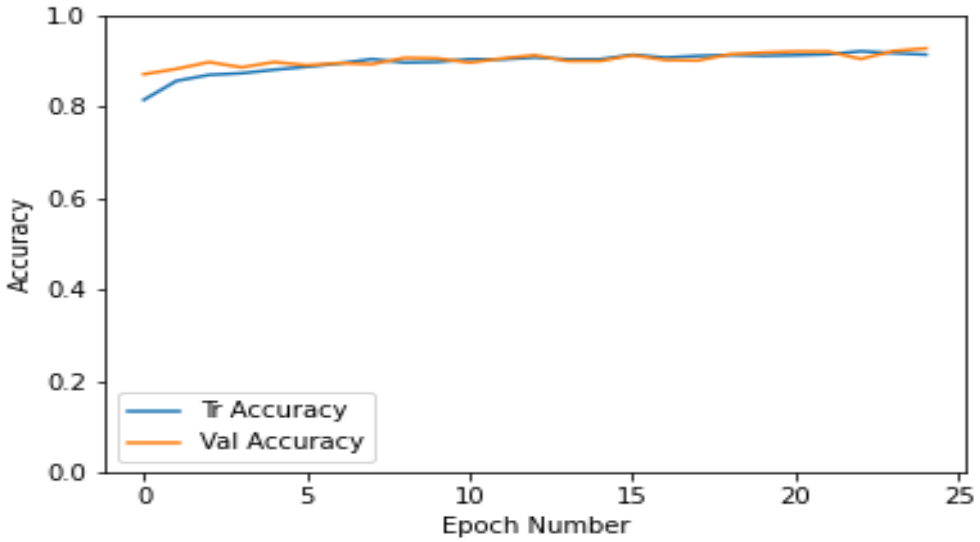


FIGURE 7.10: Training and Validation Accuracy Trends for modified AlexNet Model

7.4 Experiments with MobileNet V2

For our third experiment we used MobileNet V2 for our transfer learning process which was originally trained on ImageNet database. in this experiment for transfer learning process, the last layer was removed and a classifier was added and trained using following parameters. (I)FC layer with 256 units followed by ReLU activation (II) a Dropout of 0.4 (III)Final classification layer with 2 units and SoftMax Classifier.

The preprocessed dermoscopic images were used to train this Finetuned MobileNet V2 Model to classify different subtypes of Melanoma. Images were resized to 224x224 as ResNet accepts the standard size of this if we want to finetune the ResNet on our dataset. From these images 70% were used for training and 20% for validation and rest 10% was used for testing on unseen test data. Initial learning rate was set to 0.001. Stochastic gradient descent (SGD) with momentum was used, as optimizer where momentum was set to 0.9. Network was trained up to 25 iterations. A batch size of 32 was used during training. For training the AlexNet and MobileNet same parameters were used but in case of ResNet-18 we also trained final convolution block. Following Table summarizes configuration for MobileNet V2 Classifier.

The results of the classifier are described in the following table 7.7

Parameters	**Values**
Input Size	224 * 224 * 3
Batch Size	32
Train/Test split	70%,20%,10%
Loss Function	Cross Entropy
Optimizer	SGD with Momentum
Learning Rate	0.001
Momentum	0.9
Epochs	25

TABLE 7.6: Training Parameters used for MoileNet-V2 model

Class	**Accuracy**	**Precision**	**Recall**	**AROC**	**F1-Score**
AM	0.9190	0.9948	0.9190	0.95	0.9554
BN	0.9555	0.9294	0.9955	0.95	0.9613
Average	0.9581	0.9619	0.95	0.9579	0.9584

TABLE 7.7: Classification Results for MobileNet-V2 Model

We achieved average classification accuracy of 95.47% using MobileNet V2. The training and validation loss and accuracy trends have been shown in the following figure.

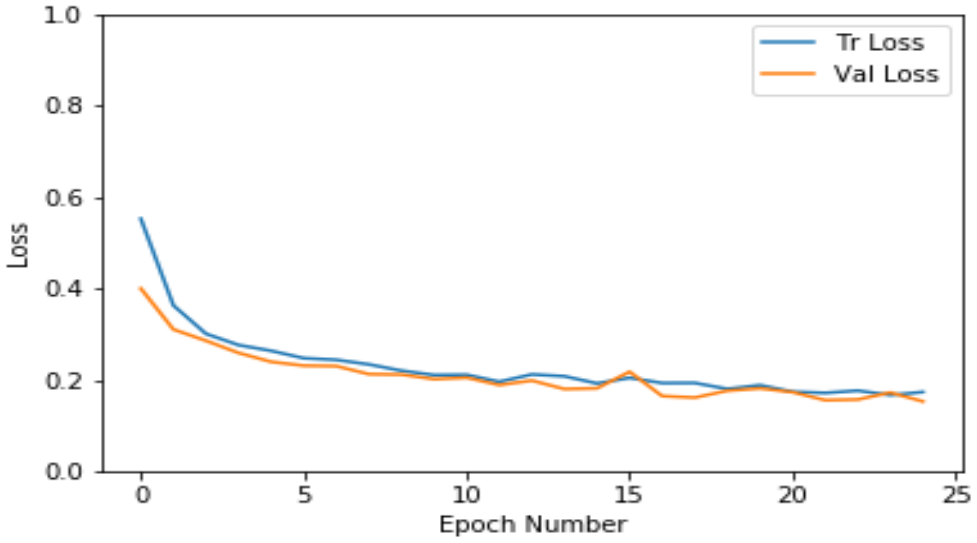


FIGURE 7.11: Training and Validation Loss Trends for modified MobileNet Model

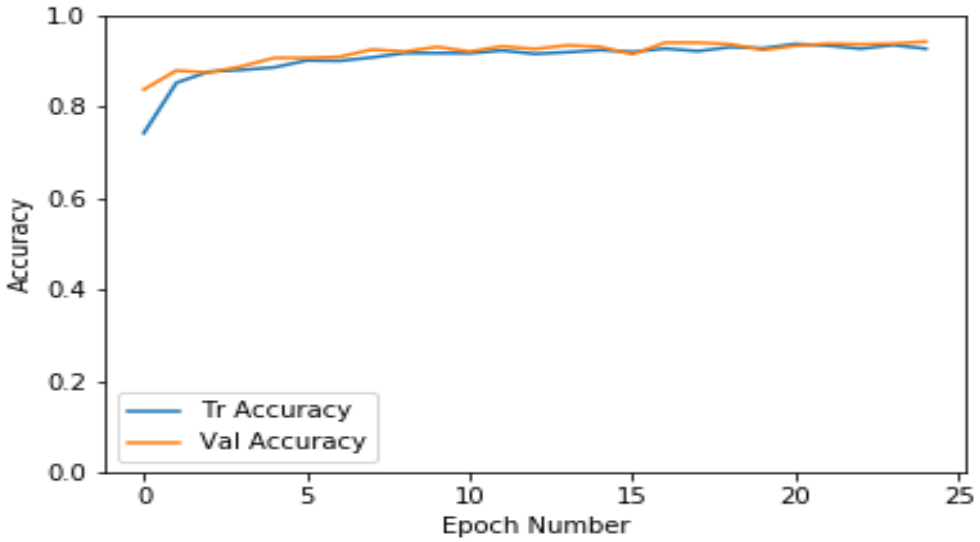


FIGURE 7.12: Training and Validation Accuracy Trends for modified MobileNet Model

Finally, the area under the ROC curve and Confusion Matrix have been visualized in the following figure.

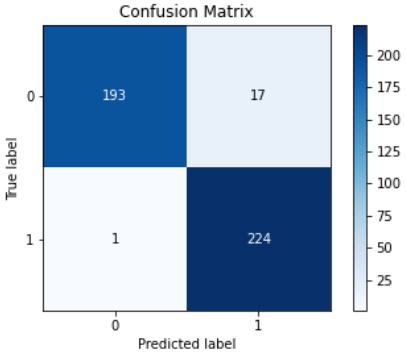


FIGURE 7.13: Confusion Matrix for MobileNet Model

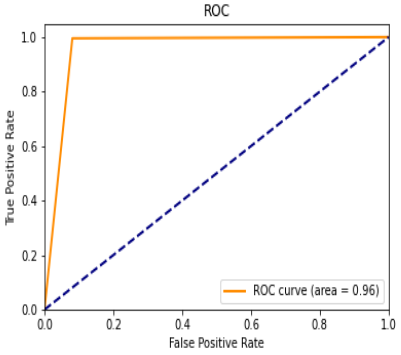


FIGURE 7.14: Area under ROC cruve for MobileNet Model

7.5 Experiments with 7-layered Deep ConvNet

For our last experiment we designed and developed a ConvNet and trained it from scratch on preprocessed dermoscopic images dataset. The architecture and parameters used in this model development is described in great detail in previous section. Here in this section we will discuss the parameters and experiments that we performed. For development of this mode we used Keras and Tensorflow [1]. For our transfer learning we used PyTorch for model design and development.

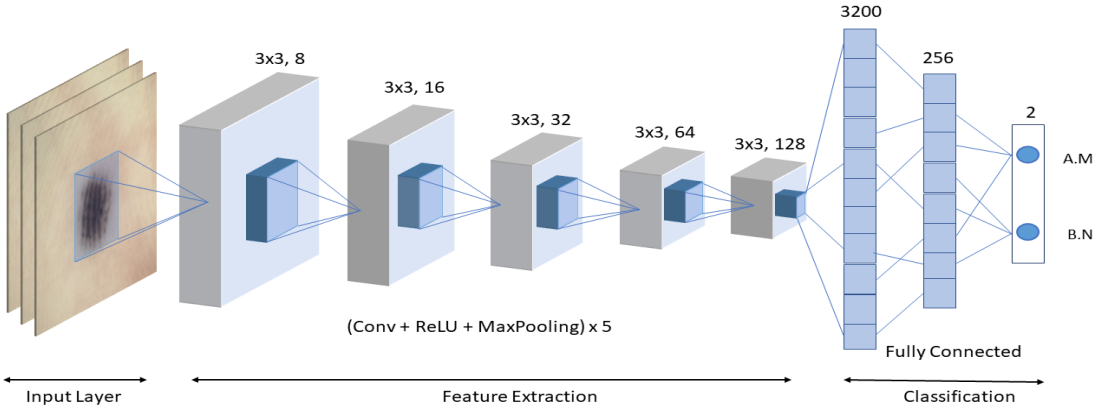


FIGURE 7.15: A 7-layers Deep ConvNet used in this study

The preprocessed dermoscopic images were used to train this custom model to classify different subtypes of Melanoma. Images were resized to 224x224 where each image was a 2D image with three channels i.e. RGB. From these preprocessed images 70% were used for training and 20% for validation and rest 10% was used for testing on unseen test data. Initial learning rate was set to 0.0001 and weight decay factor was reduced by dividing the learning rate to number of epochs after every epoch. Stochastic gradient descent (SGD) with momentum was used, as optimizer where momentum was set to 0.9. Network was trained up to 100 epochs. A batch size of 16 was used during training. Following Table summarizes overall configuration for our custom-built classifier

Parameters	**Values**
Input Size	224 * 224 * 3
Batch Size	16
Train/Test split	70%,20%,10%
Loss Function	Cross Entropy
Optimizer	SGD with Momentum
Learning Rate	0.0001
Momentum	0.9
Epochs	100

TABLE 7.8: Training Parameters used for Deep ConvNet Model

The results of the classifier are described in the table 7.9. We used PyCm [9] for evaluating results of all models. We achieved average classification accuracy of 91.03% using our Deep ConvNet. The training and validation loss and accuracy have been shown in the figures 7.16 and 7.17.

Class	Accuracy	Precision	Recall	AROC	F1-Score
AM	0.9100	0.9211	0.8904	0.9096	0.9055
BN	0.9110	0.9008	0.9288	0.9096	0.9146
Average	0.9102	0.9109	0.9096	0.9091	0.9100

TABLE 7.9: Classification Results for Deep ConvNet Model

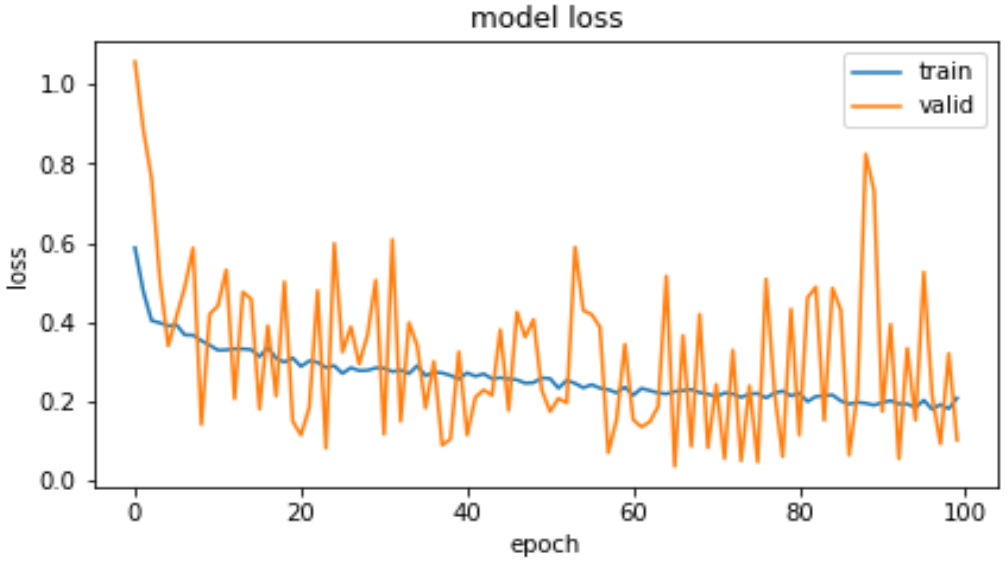


FIGURE 7.16: Training and Validation Loss Trends for our Deep ConvNet Model

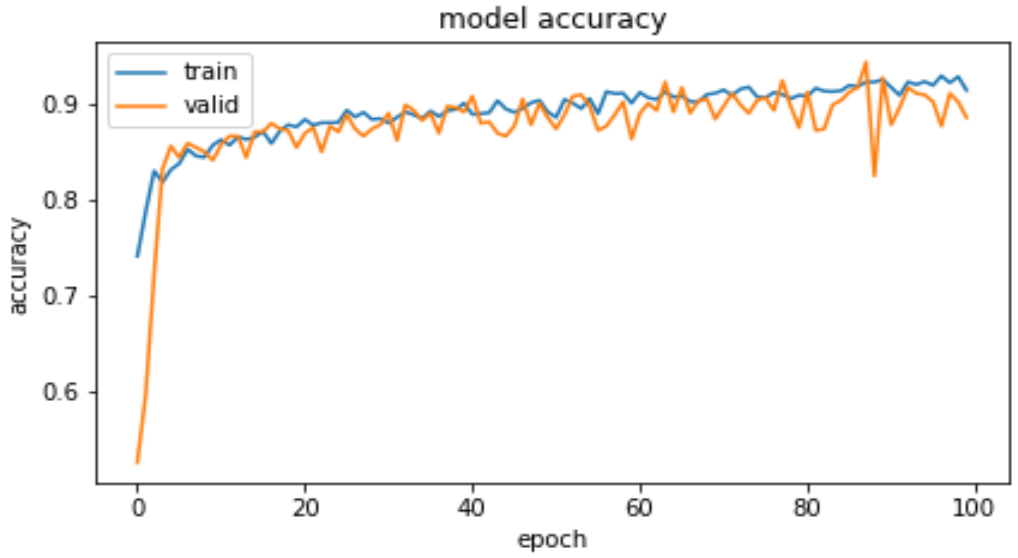


FIGURE 7.17: Training and Validation Accuracy Trends for our Deep ConvNet Model

Finally, the area under the ROC curve and confusion matrix have been visualized in the figures 7.18 and 7.19 .

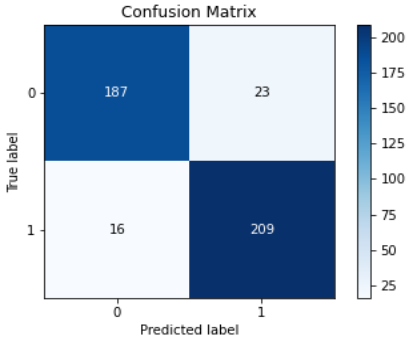


FIGURE 7.18: Confusion Matrix for Deep ConvNet Model

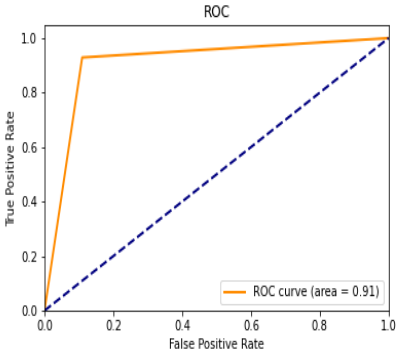


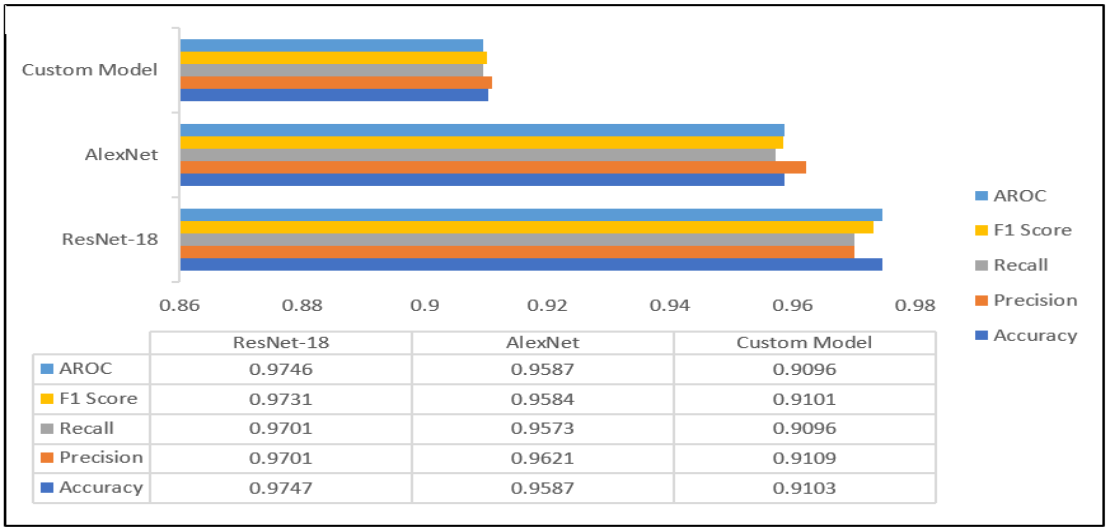
FIGURE 7.19: Area under ROC for Deep ConvNet Model

7.6 Comparative Analysis of Developed Deep Learning Models

In this section we have compared the results obtained from all four deep learning models. Transfer learning delivers state of the arts results in image classification as the models are already trained on large ImageNet database so model had learned useful low-level features which can be transferred to train these models on new datasets.

For medical image analysis large publicly available datasets are limited to some specific domains. As the melanoma is rare cancer which affects the human skin, so large dataset is not available to train the automated systems. In addition, our specific task in this research was to design and develop a ConvNet for AM classification and due to rare and infrequent occurrences of this subtype of cancer a large public dataset is not available. However, we have tried to use data augmentation to enlarge our dataset and develop an effective automated classifier. Although we trained a ConvNet from scratch we got promising results without using transfer learning.

Following figures compares the results from our classifier and the pretrained ConvNets. We used Precision, Recall, Accuracy, F1-Score and AROC for evaluation of the models.



	ResNet-18	AlexNet	Custom Model
AROC	0.9746	0.9587	0.9096
F1 Score	0.9731	0.9584	0.9101
Recall	0.9701	0.9573	0.9096
Precision	0.9701	0.9621	0.9109
Accuracy	0.9747	0.9587	0.9103

FIGURE 7.20: Comparison of Classification Results of Deep Learning Models

7.7 Comparison with state-of-the-art

A limited work is done in subtype classification of acral melanoma in literature due to limited and small dataset as discussed in previous sections. However, we found only two research projects which targeted the subtype of melanoma detection. One study used same dataset as we did [39] while other study [19] used over 1000 dermoscopic images of AM and BN. The dataset used by authors [19] is not publicly available so results cannot be reproduced or further analysis cannot be done. However, we have compared our results with these two research works and we found that our results were comparable to the work done by these authors and can be used for effective classification of melanoma subtypes.

Study	**Method**	**Classification**	**Acc %**
Hiam et al(2017)	SVM	Melanoma	92.1
Kumar et al(2018)	ANN	Melanoma	97
Yu et al(2018)	Transfer Learning	Acral Melanoma	83
Adegun et al(2019)	Autoencoder	Melanoma	95
This Study	ResNet-18	**Acral Melanoma**	**97**

TABLE 7.10: Results Comparison with state-of-the-art Methods

Chapter 8

Conclusion and Future Directions

Melanoma is a caused by abnormal growth of the skin cells which affect the majority of population in the world. It has many subtypes which can be fatal if not treated early. Once it increases the malignant tumors can affect the nearby healthy cells which results in metastatic melanoma disease where survival rate of patients decreases sharply. Due to great visual similarity between a benign and a malignant tumor on skin, it becomes difficult to diagnose the disease in early stages. This visual similarity increases the chances of misdiagnosis resulting in death of the patient. The traditional methods to automatically diagnose melanoma are limited to only binary class classification in which a melanoma is differentiated from a non-melanoma skin lesion. However, the melanoma has four further subtypes which have poor diagnosis such as acral melanoma.

A computational model for feature extraction, melanoma subtype classification has been proposed in our research. Dermoscopic images of AM were acquired from Yonsei university dataset to distinguish the acral melanoma (AM) from an acral nevus (BN). Different image preprocessing techniques were used for image enhancement and to make images ready for deep learning model. The preprocessed models were fed to deep learning models for classification. These images were used to train three deep learning models which are AlexNet, ResNet and MobileNet v2. AlexNet achieved 95% accuracy whereas MobileNet achieved 93% accuracy on unseen test data. ResNet outperformed these two models and got an accuracy of 97% on test dataset. Additionally, we designed a deep learning model and trained it from scratch without using transfer learning and achieved an accuracy of 91% on test data. Analysis of results also shown that classification and prediction of acral melanoma using dermoscopic images data is challenging but it has promising results.

In future work, these models can be utilized for further analysis. As we have a small images dataset, so acquiring a large dataset will help in developing a better classifier. In medical field, the data sources are usually limited due to privacy concerns, we can

work on developing deep learning models that can perform better on limited data. Additionally, other latest and powerful techniques for data augmentation can be used such as generating images using Generative Adversarial Networks (GANs). This research only used 724 images, however more data will result in a high performing automated system. For further future research, all of the subtypes of melanoma can be targeted for automatic diagnosis using deep learning methods. This research will help in future if someone wants to develop an automated diagnosis system for melanoma subtypes. The future system can have a single deep learning network which will diagnose all the subtypes of melanoma skin cancer by utilizing an end to end deep learning model.

References

[1] Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, et al. Tensorflow: A system for large-scale machine learning. In *12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16)*, pages 265–283, 2016.

[2] Hiam Alquran, Isam Abu Qasmieh, Ali Mohammad Alqudah, Sajidah Alhammouri, Esraa Alawneh, Ammar Abughazaleh, and Firas Hasayen. The melanoma skin cancer detection and classification using support vector machine. In *2017 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT)*, pages 1–5. IEEE, 2017.

[3] A Antony, A Ramesh, A Sojan, B Mathews, and TA Varghese. Skin cancer detection using artificial neural networking. *Skin*, 4(4), 2016.

[4] Ekaba Bisong. Google colaboratory. In *Building Machine Learning and Deep Learning Models on Google Cloud Platform*, pages 59–64. Springer, 2019.

[5] Titus J Brinker, Achim Hekler, Alexander H Enk, Joachim Klode, Axel Hauschild, Carola Berking, Bastian Schilling, Sebastian Haferkamp, Dirk Schadendorf, Tim Holland-Letz, et al. Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task. *European Journal of Cancer*, 113:47–54, 2019.

[6] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255. Ieee, 2009.

[7] Andre Esteva, Brett Kuprel, Roberto A Novoa, Justin Ko, Susan M Swetter, Helen M Blau, and Sebastian Thrun. Dermatologist-level classification of skin cancer with deep neural networks. *nature*, 542(7639):115–118, 2017.

[8] C Fink and HA Haenssle. Non-invasive tools for the diagnosis of cutaneous melanoma. *Skin Research and Technology*, 23(3):261–271, 2017.

[9] Sepand Haghighi, Masoomeh Jasemi, Shaahin Hessabi, and Alireza Zolanvari. Pycm: Multiclass confusion matrix library in python. *Journal of Open Source Software*, 3(25):729, 2018.

[10] Seung Seog Han, Myoung Shin Kim, Woohyung Lim, Gyeong Hun Park, Ilwoo Park, and Sung Eun Chang. Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm. *Journal of Investigative Dermatology*, 138(7):1529–1538, 2018.

[11] Seung Seog Han, Ik Jun Moon, Woohyung Lim, In Suck Suh, Sam Yong Lee, Jung-Im Na, Seong Hwan Kim, and Sung Eun Chang. Keratinocytic skin cancer detection on the face using region-based convolutional neural network. *JAMA dermatology*, 156(1):29–37, 2020.

[12] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[13] Khalid M Hosny, Mohamed A Kassem, and Mohamed M Foaud. Classification of skin lesions using transfer learning and augmentation with alex-net. *PloS one*, 14 (5):e0217293, 2019.

[14] Andrew G Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*, 2017.

[15] Hitoshi Iyatomi, Hiroshi Oka, M Emre Celebi, Koichi Ogawa, Giuseppe Argenziano, H Peter Soyer, Hiroshi Koga, Toshiaki Saida, Kuniaki Ohara, and Masaru Tanaka. Computer-based classification of dermoscopy images of melanocytic lesions on acral volar skin. *Journal of Investigative Dermatology*, 128(8):2049–2054, 2008.

[16] Sara Hosseinzadeh Kassani and Peyman Hosseinzadeh Kassani. A comparative study of deep learning architectures on melanoma detection. *Tissue and Cell*, 58: 76–83, 2019.

[17] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[18] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[19] S Lee, YS Chu, SK Yoo, S Choi, SJ Choe, SB Koh, KY Chung, L Xing, B Oh, and S Yang. Augmented decision-making for acral lentiginous melanoma detection using deep convolutional neural networks. *Journal of the European Academy of Dermatology and Venereology*, 2020.

[20] Yuexiang Li and Linlin Shen. Skin lesion analysis towards melanoma detection using deep learning network. *Sensors*, 18(2):556, 2018.

[21] Eleni Linos, Kenneth A Katz, and Graham A Colditz. Skin cancer—the importance of prevention. *JAMA internal medicine*, 176(10):1435–1436, 2016.

[22] Justine Mayer. Systematic review of the diagnostic accuracy of dermatoscopy in detecting malignant melanoma. *The Medical Journal of Australia*, 167(4):206–210, 1997.

[23] A Murugan, S Anu H Nair, and KP Sanal Kumar. Detection of skin cancer using svm, random forest and knn classifiers. *Journal of medical systems*, 43(8):269, 2019.

[24] Terry Ngo. Data mining: practical machine learning tools and technique, by ian h. witten, eibe frank, mark a. hell. *ACM SIGSOFT Software Engineering Notes*, 36 (5):51–52, 2011.

[25] Kristjan Orthaber, Matevž Pristovnik, Kristijan Skok, Barbara Perić, and Uroš Maver. Skin cancer and its treatment: novel treatment approaches with emphasis on nanotechnology. *Journal of nanomaterials*, 2017, 2017.

[26] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. In *Advances in neural information processing systems*, pages 8026–8037, 2019.

[27] Fábio Perez, Cristina Vasconcelos, Sandra Avila, and Eduardo Valle. Data augmentation for skin lesion analysis. In *OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis*, pages 303–311. Springer, 2018.

[28] MS Poornima and K Shailaja. Detection of skin cancer using svm. *Int. Res. J. Eng. Technol*, 4, 2017.

[29] GD Praveenkumar and M Dharmalingam. Detection of skin cancer using artificial neural network classifier.

[30] SM Rajpara, AP Botello, J Townend, and AD Ormerod. Systematic review of dermoscopy and digital dermoscopy/artificial intelligence for the diagnosis of melanoma. *British Journal of Dermatology*, 161(3):591–604, 2009.

[31] Mi Ryung Roh, Jihyun Kim, and Kee Yang Chung. Treatment and outcomes of melanoma in acral location in korean patients. *Yonsei medical journal*, 51(4):562–568, 2010.

[32] Julie Ann A Salido and C Ruiz. Using deep learning to detect melanoma in dermoscopy images. *International Journal of Machine Learning and Computing*, 8(1), 2018.

[33] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4510–4520, 2018.

[34] Fatma Sherif, Wael Ahmed Mohamed, and Ashraf Mohra. Skin lesion analysis toward melanoma detection using deep learning techniques. *International Journal of Electronics and Telecommunications*, 65(4):597–602, 2019.

[35] Rebecca L Siegel, Kimberly D Miller, Ann Goding Sauer, Stacey A Fedewa, Lynn F Butterly, Joseph C Anderson, Andrea Cercek, Robert A Smith, and Ahmedin Jemal. Colorectal cancer statistics, 2020. *CA: a cancer journal for clinicians*, 2020.

[36] Eric Robert Tkaczyk. Innovations and developments in dermatologic non-invasive optical imaging and potential clinical applications. *Acta dermato-venereologica*, page 5, 2017.

[37] Halil Murat Ünver and Enes Ayan. Skin lesion segmentation in dermoscopic images with combination of yolo and grabcut algorithm. *Diagnostics*, 9(3):72, 2019.

[38] Sejung Yang, Byungho Oh, Sungwon Hahm, Kee-Yang Chung, and Byung-Uk Lee. Ridge and furrow pattern classification for acral lentiginous melanoma using dermoscopic images. *Biomedical Signal Processing and Control*, 32:90–96, 2017.

[39] Chanki Yu, Sejung Yang, Wonoh Kim, Jinwoong Jung, Kee-Yang Chung, Sang Wook Lee, and Byungho Oh. Acral melanoma detection using a convolutional neural network for dermoscopy images. *PloS one*, 13(3):e0193321, 2018.

[40] Howard Zhou, Mei Chen, Richard Gass, James M Rehg, Laura Ferris, Jonhan Ho, and Laura Drogowski. Feature-preserving artifact removal from dermoscopy images. In *Medical Imaging 2008: Image Processing*, volume 6914, page 69141B. International Society for Optics and Photonics, 2008.