SKIN LESION CLASSIFICATION USING TRANSFER LEARNING AND 1-D CONVOLUTIONAL NEURAL NETWORK

Submitted By

Md. Jibon Miah
Reg No: 2018338021

And

Md Tawsifur Rahman
Reg No: 2018338027

A dissertation submitted for the partial fulfillment of the requirements for the degree of
Bachelor of Science in Electrical and Electronic Engineering.

Department of Electrical and Electronic Engineering
Shahjalal University of Science and Technology Sylhet-3114,Bangladesh.
February, 2024

Certificate of Approval

The thesis titled "**Skin Lesion Classification using Transfer Learning and 1-D Convolutional Neural Network**" submitted by Md. Jibon Miah, Reg No. 2018338021, and Md Tawsifur Rahman, Reg No. 2018338027, Session: 2018-19, has been accepted as satisfactory in partial fulfillment of the requirement for the degree of Bachelor of Science in Electrical and Electronic Engineering on February 22,2024.

BOARD OF EXAMINERS

Supervisor

Md. Shariful Islam

Lecturer

Department of Electrical & Electronic Engineering

Shahjalal University of Science & Technology, Sylhet, Bangladesh

---

Chairman of Exam Committee

Dr. Md Rasedujjaman

Associate Professor & Head

Department of Electrical & Electronic Engineering

Shahjalal University of Science & Technology, Sylhet, Bangladesh

Declaration

It is hereby declared that this thesis titled "**Skin Lesion Classification using Transfer Learning and 1-D Convolutional Neural Network**" or any part of it has not been submitted elsewhere for the award of any degree or diploma.

_____

Supervisor
Md Shariful Islam
Lecturer
Department of EEE, SUST.

_____

Author 1
Md. Jibon Miah
Reg No: 2018338021
Dept. of EEE, SUST
Session: 2018-19

_____

Author 2
Md Tawsifur Rahman
Reg No: 2018338027
Dept. of EEE, SUST
Session: 2018-19

Dedication

To our dear loved ones, whose unwavering support and encouragement have been our guiding light throughout this journey.

Acknowledgement

We extend our heartfelt appreciation to all those who have accompanied us on our thesis expedition. A special mention goes to our parents, whose steadfast love, guidance, and encouragement have been indispensable on our path to achievement. Without their unwavering support, we would not have reached this significant milestone.

We are deeply indebted to our thesis supervisor, Md Shariful Islam, for his expert guidance and constant support. His dedication to excellence has been pivotal in shaping our thesis and fostering our critical thinking skills.

Acknowledgment is also due to the esteemed faculty members of the Department of EEE, whose exceptional teaching and research have provided us with a strong foundation of knowledge and skills essential for completing this thesis.

To our friends and classmates, we express our sincerest gratitude for their steadfast support and constructive feedback throughout this journey. Their belief in our capabilities has been invaluable. We are honored by everyone who has contributed to our thesis and cannot thank you enough for your support, guidance, and encouragement, which have made this journey possible.

Abstract

Skin cancer is a global health concern, with rising incidence rates and significant mortality rates worldwide. Early detection is crucial for effective treatment and improved patient outcomes. However, the current methods for skin lesion classification and cancer detection face challenges, including the reliance on expert dermatologists and invasive procedures such as biopsy and histopathology.

In response to these challenges, artificial intelligence, particularly deep learning, has emerged as a promising solution for skin cancer detection. By leveraging complex algorithms, deep learning models can accurately classify dermoscopic images, enabling early detection and non-invasive diagnostics.

Despite the importance of early detection, many countries, including Bangladesh, lack access to specialized professionals and face financial barriers to cancer detection. Deep learning methods offer a cost-effective and accessible solution by automating image analysis and reducing the burden on specialists.

This study utilizes a pre-trained Deep Learning framework, EfficientNetB7, and a novel Convolutional Neural Network (CNN) model to classify skin lesions accurately. The dataset, sourced from reputable databases, undergoes extensive preprocessing and augmentation before being processed by the deep learning models.

The results demonstrate the effectiveness of the proposed approach, with accuracies exceeding 90% across all sets. This study underscores the transformative potential of deep learning in revolutionizing medical diagnostics, offering hope for improved patient care and outcomes worldwide.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation of the Thesis

Skin cancer has been a serious issue worldwide, with its prevalence posing a significant threat to individuals and the healthcare systems. Along with its unstoppable growth, the depressing mortality rates pose a formidable challenge to global health systems. According to Global Cancer Statistics 2020: GLOBOCAN Estimates, the number of new skin cancer deaths in 2020 exceeded 120,000 globally [1].

Melanoma is considered to be the most dangerous form of skin cancer. It may not be the most common form of skin lesion, but the majority of deaths caused by skin cancer are due to this type of lesion [2]. Adding to that, more than 1 million new cases of non-melanoma cancers were reported by GLOBOCAN Estimates [1]. The increasing depletion of the ozone layer, which results in greater exposure to UVB rays, diminishes the chances of non-melanoma skin cancer cases decreasing anytime soon [3] .

Detecting skin cancer in the earlier stages has been a game changer [4]. It not only enhances treatment outcomes but also reduces the need for more penetrative interventions, leading to improved quality of life for patients. In fact, Melanoma exhibits an impressive 5-year survival rate of 99 percent when identified at an early stage [3]. Differentiating among the key pigmented skin lesions is therefore extremely vital in order to diagnose the proper type of skin cancer as early as possible.

Currently, various methods are utilized in order to classify skin lesions and detect. The most basic and widespread approach involves the visual examination of dermoscopic images by expert dermatologists. Biopsy, followed by histopathology, is another method which is employed for skin lesion identification, especially when a proper inference cannot be reached through dermoscopic image inspection [5].

Biopsies can be invasive and carry risks such as scarring, bleeding, and infection. Additionally, visual inspection of dermoscopic images and histopathological analysis may be time-consuming and costly, and it requires specialized expertise for accurate interpretation[6]. Moreover, the reliance on subjective human assessment in both visual examination and histopathology can introduce variability and potential errors in diagnosis.

Building upon these limitations, the application of artificial intelligence, in the form of deep learning, in skin cancer detection has emerged as a promising frontier in healthcare [7]. Dermoscopic images can be classified into their corresponding classes accurately by leveraging the power of complex algorithms of the deep learning models. These models can not only classify the dermoscopic images, but also perform efficient image segmentation and feature extraction [7]. Apart from facilitating early detection, implementing deep learning models in skin lesion image classification also provides a non-invasive and rapid approach.

Since domain expertise is imperative for skin lesion classification, deep learning systems cannot be employed on their own. Nevertheless, their implementation can potentially revolutionize early detection of skin lesions and significantly enhance the efficiency of decision-making among clinicians and dermatologists.

1.2 Thesis Objective

Aforementioned, the classification of skin lesions and detection of skin cancer as early as possible is extremely important [4]. However, due to the lack of qualified and specialized professionals, it is easier said than done in many countries in Asia, including Bangladesh [6][8]. The cost of various examinations and procedures currently involved in cancer detection makes it all the more difficult for poverty-stricken people in Bangladesh [9].

The employment of deep learning methods in this scenario can not only help pathologists and dermatologists, but ensure treatment for all patients, regardless of their financial conditions [10]. The deep learning models can automate the process of analyzing skin lesion images, reducing the burden on specialists and increasing the efficiency at the same time.

Since these models can be constructed using open-source frameworks and deployed into low-cost cloud platforms, the accessibility of cancer detection can be extended to the most remote places around the world [11], while also maintaining cost-efficiency for patients [10].

The study utilizes the power of a pre-trained Deep Learning framework, namely EfficientNetB7 [12], and a novel Convolutional Neural Network (CNN) model in order to classify skin lesions accurately and reliably [13]. The dataset, consisting of 25,335 images classified into 9 different categories, was collected from the The International Skin Imaging Collaboration (ISIC) database for 2019 [14][15][16]. ISIC builds the database of dermoscopic images by collaborating with various hospitals and cancer centers. Moreover, existing dermoscopic images from well-known databases, like HAM10000, are also included. [14]

Preceded with extensive data preprocessing and augmentation techniques, the EfficientNetB7 deep learning model was utilized to extract features from the images. After further preprocessing of the extracted features, the data was put through the proposed classification model.

The novel classifier excelled in all of the training, validation, and testing sets, scoring accuracies above 90%. Responsible implementation of deep learning methods can create a revolutionary change in the medical industry and bring about a positive change for both the patients and the doctors.

1.3 Thesis Outline

Chapter 2: Background Study

In Chapter 2, the extent of the research pertaining to the use of Artificial Intelligence (AI) in Skin Lesion Classification is presented. It encompasses the diverse research domains relevant to this study and discusses the various methodologies employed in machine learning and deep learning implementations within this field.

Chapter 3: Proposed Methodology

Within Chapter 3, a comprehensive explanation and analysis of the methodology employed throughout the study are provided. This includes a detailed discussion of the entire procedural framework utilized in the research endeavor.

Chapter 4: Results and Discussion

Chapter 4 presents the findings derived from the applications of the proposed methodology, along with discussions concerning specific parameters and attributes.

Chapter 5: Conclusion and Future Works

In Chapter 5, a concise summary of the entire study is presented, accompanied by recommendations for enhancements aimed at optimizing performance and enhancing usability.

Chapter 2

Background Study

2.1 1-D Convolutional Neural Networks

1-D CNNs are specialized for one-dimensional data and work well with raw data [71][72], thereby reducing the need for feature engineering. These networks have excelled most in signal processing tasks involving Speech Recognition, monitoring of degradation in Bearings, Rotating Machinery [74], and signals in Electrocardiograms [73].

Furthermore, leveraging features extracted from pre-trained models within a 1-D CNN framework enhances classification performance by capitalizing on the network's inherent capabilities for processing one-dimensional data. This approach reduces the dependency on manual feature engineering, streamlining the model development process. 1-D CNNs possess the distinctive ability to integrate both the process and outputs of feature extraction and classification within a unified learning framework [75]. Given the aforementioned use cases, the integration of pre-trained feature extraction with 1-D CNNs holds promise for advancing classification tasks.

2.2 Image Preprocessing

Image preprocessing is a critical component in the realm of machine learning and deep learning, serving the fundamental purpose of ensuring that models are provided with clean, relevant data devoid of noise or inaccuracies [21]. The most basic form of preprocessing is resizing the images. While it may be a simple task, it plays a significant role regarding computational costs and training efficiency [22] [23]. Models trained on images of

dimensions 256X256 will learn more data than those of 64X64 [23]. However, this will require more computational power. Therefore, a reasonable trade-off should be achieved.

The presence of hair surrounding or within lesions poses a challenge to the learning process of machine learning models [24]. Several strategies have been devised to address this issue. Thresholding techniques have been employed for border identification [25], particularly in studies focusing on Melanoma classification utilizing the ABCD technique [26]. These methodologies have found application in the segmentation of skin lesion images [27] [28].

Moreover, morphological operations have been integrated for noise reduction, border delineation, and feature extraction [29]. Operations such as morphology opening, closing, and erosion have been leveraged to refine image quality, extract pertinent features, and facilitate subsequent analysis for diagnostic or investigative purposes [30]. Additionally, the combination of Black Hat morphology with Total Variation and Inpainting methods has been explored for hair removal from skin lesion images [31] [32]. The INPAINT_TELEA function within the OpenCV library is notable for its ability to produce smooth and seamless results in inpainting tasks [33].

Furthermore, the application of DullRazor software is prevalent for image preprocessing, including hair removal and segmentation [34]. Seamless integration of algorithms such as YOLO, active contour, and JSEG fusion for image segmentation is facilitated through the utilization of DullRazor [35][36][37].

To counter improper brightness and contrast ratios in the skin lesion images, the use of various techniques for enhancement of image is prevalent. The most common method that is applied is the Contrast Limited Adaptive Histogram Equalization (CLAHE) [38][39]. Apart from that, normal adaptive histogram equalization methods [40] and Z-score transformation algorithms [41] have also been put to practice.

Data augmentation is another preprocessing approach utilized to enhance the robustness and generalizability of deep learning models trained on skin lesion images. By artificially expanding the dataset through various transformations, data augmentation mitigates overfitting and improves the model's ability to recognize diverse skin lesions under different conditions [42][43]. This technique not only augments the quantity of training data but also introduces variations that simulate real-world scenarios.

2.3 Transfer Learning and Feature Extraction

Transfer learning is a powerful technique in deep learning that leverages pre-trained models to enhance the performance of tasks with limited labeled data, such as skin lesion classification. By transferring knowledge from models trained on large datasets (e.g., ImageNet), transfer learning enables the extraction of meaningful features even when labeled datasets are scarce [44][45]. This approach accelerates model convergence, improves classification accuracy [46][47][48][49], and reduces the need for extensive computational resources and labeled data.

There are numerous pre-trained models available, each with particular strengths and weaknesses. For image classification of natural scenarios, animals, and plants, the VGG models are a step ahead of others [53], whereas DenseNet [54], MobileNet V2 [57], and EfficientNet can be used in the healthcare industry [55][56].

The EfficientNet architecture, with multiple variants, has performed relatively well compared to other state-of-the-art models, like ResNet-50 [59]. For the purpose of skin lesion classification, EfficientNet was integrated with U-Net++ for gradual changes in the convolutional neural network, leading to skin lesion segmentation, followed by classification [58]. A comparison among all the EfficientNet structures were also carried out regarding the performance in detection of skin disorders, in which EfficientNetB7

excelled over all others in terms of accuracy, while EfficientNetB0 was the most time-efficient [60].

Feature extraction using transfer learning involves fine-tuning pre-trained models to adapt them to the specific characteristics of skin lesion images, thereby capturing relevant features for accurate diagnosis and classification. Leveraging the feature extraction knowledge of the ImageNet dataset, pre-trained models perform feature extraction considerably well in all domains, including facial recognition [50], craniofacial abnormalities [51], as well as skin lesion images [52].

2.4 Overfitting Minimization

The massive discrepancy in number of images for each label causes major overfitting issues while training a deep learning model in the skin lesion image dataset. The Melanocytic Nevi (NV) class always has fifty to sixty times more images than the minority class, which is usually Dermatofibroma (DF) or Vascular Lesion (VASC).

Due to the imbalance, bias is introduced into the learning process and the deep learning algorithms tend to prioritize the majority class due as its greater in number. Therefore, the model may become overly biased towards predicting the majority class, resulting in poor generalization performance and reduced predictive accuracy, particularly for minority classes.

Various steps have been taken to try and minimize the overfitting issue while obtaining state-of-the-art model performance. One particular approach consisted of setting the number of images for each class to the number of images available in the minority class [61]. While this would minimize overfitting greatly, it would reduce the size of the training dataset, leading to the model learning lesser features.

Another approach involved performing data augmentation techniques, such as rotation, width and height variation, horizontal flips, and generating new images for the minority class [62]. This would introduce variations and somewhat equalize the number of images for all classes.

A further advancement to this approach was made by the implementation of GAN to generate synthetic images in order to reduce the class imbalance [63]. The use of GAN, while computationally expensive, has led to notable increases in model accuracies [64][65][66].

The SMOTE algorithm [69] has also been utilized to generate synthetic images for skin lesions and reduce the difference in class instances [70]. SMOTE is specialized for numerical data, so a small issue might persist when used in image data, which is generation of synthetic data that isn't corresponding to the actual image, leading to noise.

Leveraging the power of transfer learning through pre-trained models only to extract features for downstream classification later on has proven to prevent overfitting and lead to excelling accuracies [67][68].

2.5 Availability of Datasets

Deep Learning approaches trained on comprehensive datasets hold immense promise in addressing the challenge of skin lesion classification and cancer detection. However, the effectiveness of these algorithms hinges on the quality and diversity of the training data. A wide variety of datasets are available, with images collected from numerous sources and examined by experts around the world.

The HAM10000 image dataset is one of the more common ones utilized for deep learning training purposes [15]. It contains 10,015 dermoscopic images of pigmented skin lesions

with clinical diagnoses. It Includes both benign and malignant lesions, covering a wide spectrum of appearances. Information regarding seven different classes for each lesion is also available .

37000 high-quality images, with detailed descriptions and diagnoses are also present in DermNet NZ website, based in New Zealand [17]. The image data is accessible for free and covers a wide range of dermatological conditions .

A wide collection of dermoscopic images from various place around the world is combined into the PH2 dataset [18]. A total of 13,000 images consisting of clinical diagnoses and follow-up information makes it highly effective for deep learning model training. It mainly focuses on pigmented skin lesions, including melanomas and seborrheic keratosis.

The Skin Cancer Image Atlas (SCIA) [19] focuses specifically on patients diagnosed with skin cancer, with over 2000 images accurately annotated by dermatologists.

The International Skin Imaging Collaboration (ISIC) offers by far the largest dataset of them all. The organization collects skin lesion images from a wide variety of sources, including other datasets [14][15][16]. This enriches the variation and quality of the dermoscopic images present. The archives of lesion images are also updated on a regular basis, and all subsets of each year are available for use, such as ISIC 2017, ISIC 2018, ISIC 2019. A spreadsheet file is associated with the subsets, denoting not only the classification of the lesion, but also important information regarding the patients. For efficient deep learning model training, large datasets are essential [20], and the massive archive offered by ISIC is exactly what is required.

Chapter 3

Proposed Methodology

The key objective of the study is to classify skin lesion images into specific categories based on particular features corresponding to each type of lesion. In this chapter, an extensive discussion is carried out involving all the procedures involved in Image Processing, Feature Extraction, Data Manipulation, and Classification.

3.1 Dataset Description

The International Skin Imaging Collaboration (ISIC) provides a huge collection of Skin Lesion images for machine learning and deep learning purposes. All collected images are obtained through dermoscopic photography and provided in various file formats.

The particular dataset used in this study is the ISIC-2019, consisting of 25,335 images classified into 10 specific categories [14][15][16]. The categories are Melanoma (MEL), Melanocytic Nevus (NV), Basal Cell Carcinoma (BCC), Actinic Keratosis (AKIEC), Benign Keratosis-like Lesions (BKL), Dermatofibroma (DF), Vascular Lesion (VASC), Squamous Cell Carcinoma (SCC), Seborrheic Keratosis (SK), Solar Lentigo (SL). The dataset is, however, heavily imbalanced. The images are presented in '.jpg' file format. Other file formats, such as DICOM, is not available for this particular dataset.

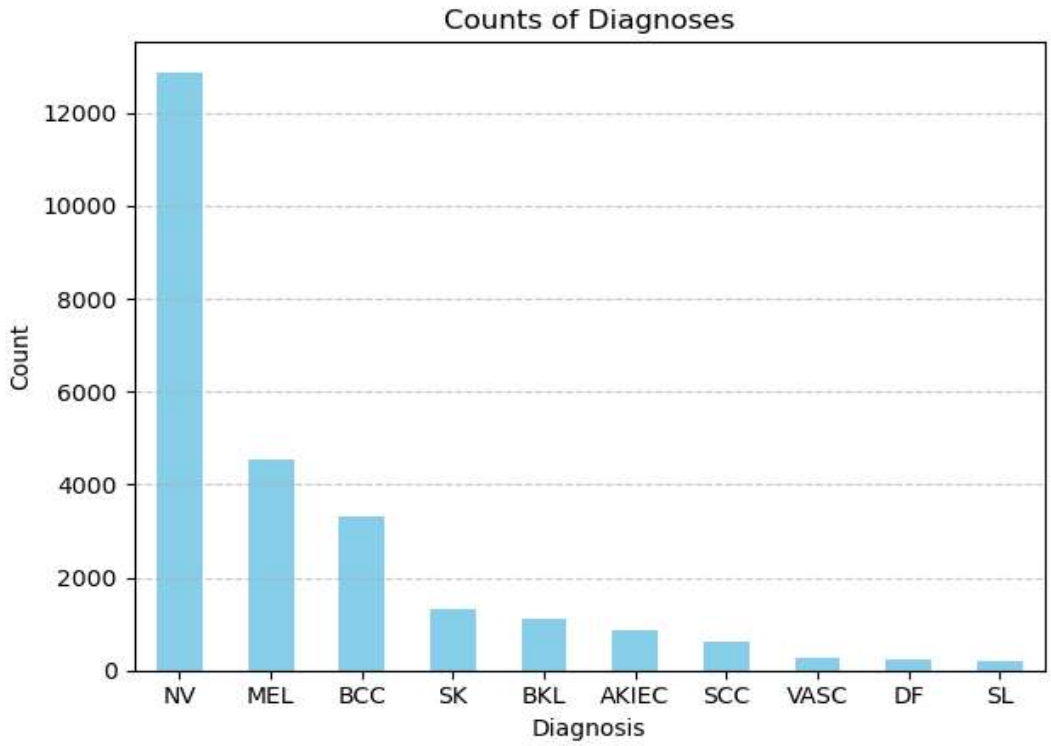

Figure 3.1: Distribution of classes in the skin lesion dataset

Figure 3.1 gives an overview of the various classes present in the dataset and the number of images belonging to each class. The dataset directory contains comprehensive patient information, including lesion image identification numbers and corresponding classifications, meticulously organized within a spreadsheet.

Concurrently, all lesion images are cataloged within a dedicated folder in the same directory. These images are readily accessible for download from the ISIC archive, facilitating easy access at no cost.

The approximate mean age of the subjects involved is 54 years. Additionally, the dataset consists of images from 13286 males and 11661 females, as denoted in Figure 3.2. The classification of the images were made by experts through various examinations such as

Histopathology, Serial Imaging, Single Image Expert Consensus, and Confocal Microscopy with Consensus Dermoscopy. The images collected cover a comprehensive range of anatomical regions, spanning from the head and mouth to the palms and torso, thereby diversifying the dataset.

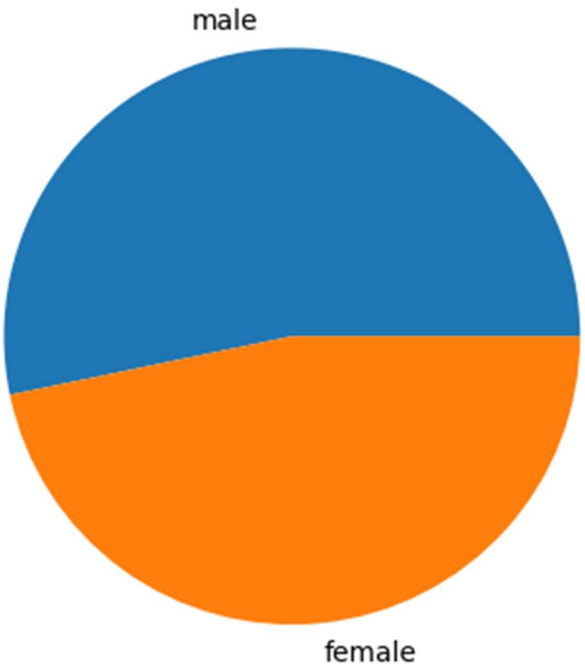

Figure 3.2: Distribution of gender in the skin lesion dataset

3.2 Data Manipulation

Upon examination of the spreadsheet, it was noticed that the category of SCC has 628 instances in the dataset. However, the instances for confirmation of diagnosis for this category were around 569. On top of that, in order to maintain relation among other popular datasets, like the HAM10000 which comprises 7 classes, and provide reliable comparison of Deep Learning models, all images categorized under SCC were eliminated.

BKLs are suggested to denote a regressive manifestation of pre-existing epidermal tumors such as SL and SK [76][77]. Keeping that in mind, all images under SL and SK were combined and brought under the BKL class, increasing the number of images labeled BKL from 1099 to 2624.

After manipulation, the total number of images in the final dataset was 24703, ranging across 7 classes.

3.3 Data Preparation

Initially, the dataset was structured as such: a directory containing all the lesion images and a corresponding spreadsheet. This spreadsheet contained crucial information such as the ID for each image paired with its corresponding label.

To streamline the dataset for analysis, a Python script was employed. This script facilitated the comparison of each image ID from the spreadsheet with the corresponding image in the directory. Based on the labels associated with each ID, the script dynamically created folders representing different categories or labels. Subsequently, each image was systematically copied or moved into the respective folder based on its assigned label. The folders were named after the corresponding labels.

This methodical approach ensured that the dataset was organized and categorized according to the specific labels or classifications assigned to each skin lesion. The final counts of each label in the dataset are depicted in Figure 3.3.

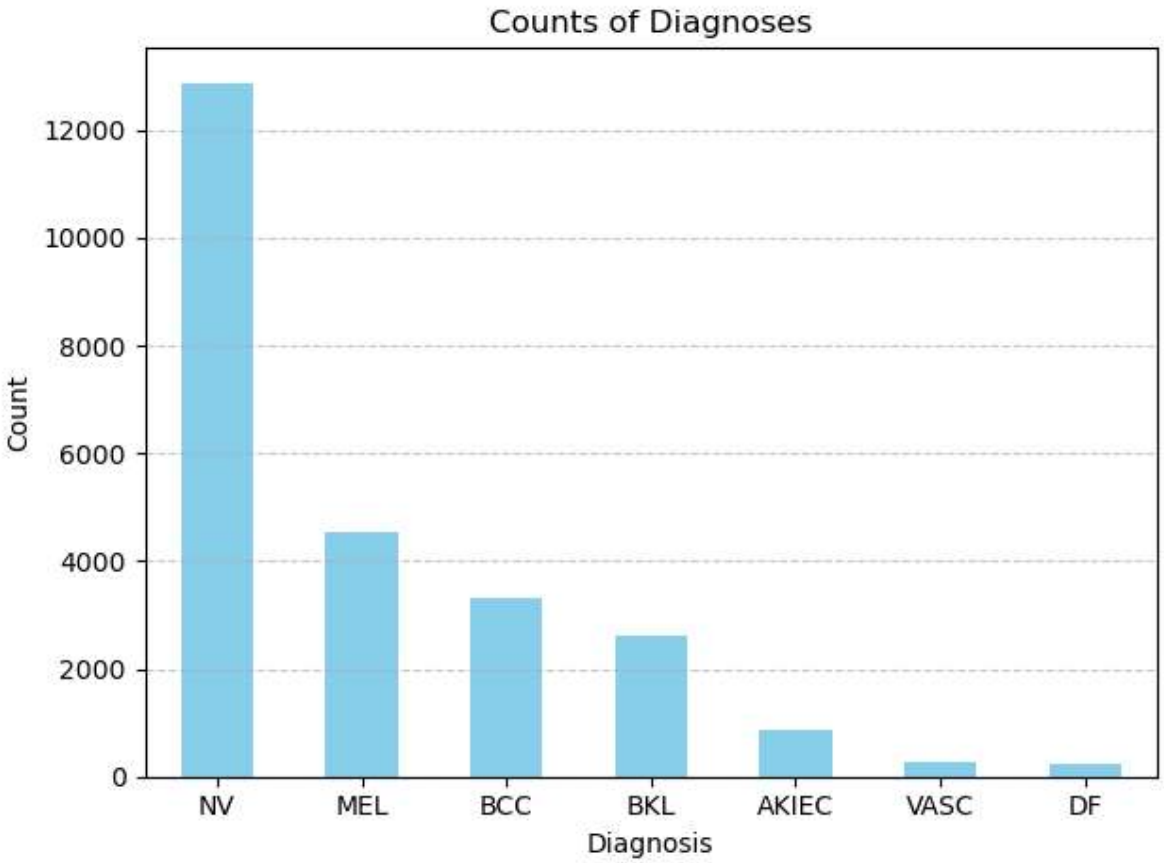

Figure 3.3: Distribution of classes in dataset after Manipulation

3.4 Image Preprocessing

Preprocessing of images is crucial for enhanced performance of any deep learning model [21]. In order to prevent the model from learning noise patterns and structures that don't pertain to the lesion, proper preprocessing must be carried out. From the initial loading of the image to the removal of hair and artifacts, each step plays a pivotal role in generating optimized images tailored for model training and ensuring the quality and efficacy of the final dataset.

3.4.1 Image Loading and Resizing

The OpenCV library is utilized to load and resize images into the desired dimensions. OpenCV, short for Open Source Computer Vision Library, is a powerful and widely-used open-source library designed for computer vision and image processing tasks. It has become a standard tool in the field of computer vision due to its extensive functionality and image processing functions.

The images are imported through folder categorization, wherein the folder names denote respective classes and are recorded alongside each image. Subsequently, each image undergoes resizing from its original dimensions to a standardized size of 224x224 pixels. This resizing strategy strikes a balance between retaining essential image details and mitigating computational resource strain [22].

3.4.2 Enhancement and Hair Removal

Hair and other unwanted structures in the skin lesion images are removed by implementing Black Hat Transformation and Inpainting. Initially, the images are converted to Grayscale, which helps simplify the image to focus solely on variations in brightness.

After that, the Black Hat transformation is applied to the grayscale image. This operation helps to highlight darker regions, such as hair and artifacts, against brighter background areas like the skin lesion. This highlighting facilitates the identification of unwanted elements for subsequent removal.

A kernel is defined with a particular shape and size, which applies the morphological blackhat operation [29] on the grayscale image using the MORPH_BLACKHAT function

of OpenCV. The Black Hat essentially highlights dark objects that correspond to the kernel, while ignoring anything larger [30].

A thresholding function [25], also belonging to OpenCV, is then applied to the Black Hat image, which converts the highlighted dark regions into a binary mask. These regions are typically associated with hair, artifacts, or other unwanted elements on the skin surface.

The thresholded image, which is black-and-white, is then passed through an inpainting function of OpenCV [31]. The INPAINT_TELEA function is well-known for its effectiveness in image restoration [32]. It works by intelligently filling gaps based on surrounding pixel information, blending seamlessly with the image.

The final image then consists of the key lesion structure of the original image, but with unwanted hair and artifacts removed. Each skin lesion image is passed through the entire process and the resulting output is saved in a different directory. Figure 3.4 offers a visual progression of the image preprocessing steps involved.

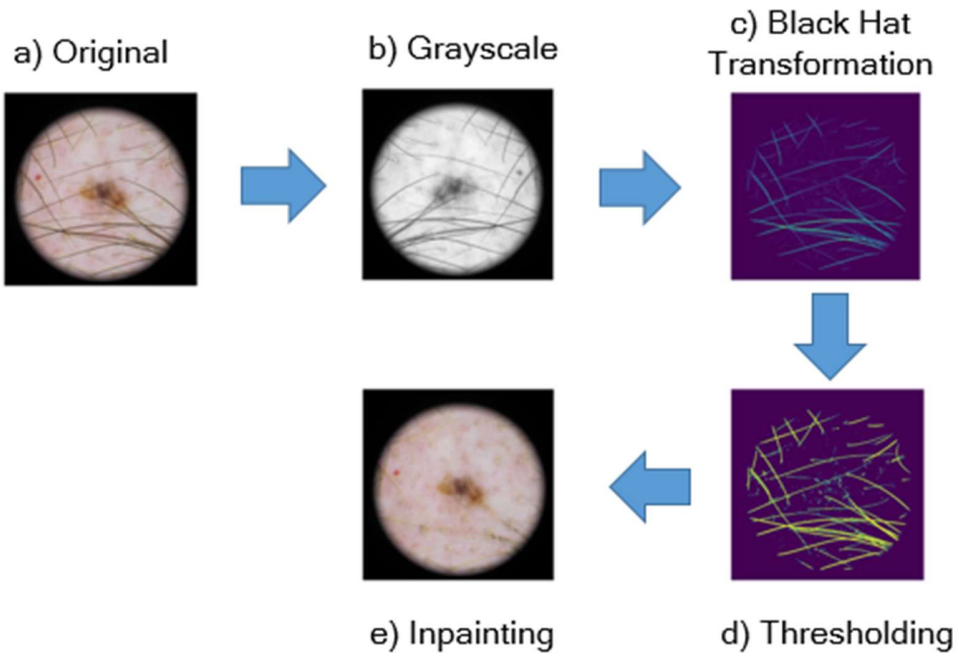

Figure 3.4: Image Preprocessing steps for Hair and Artefacts Removal

3.4.3 Data Augmentation

The robustness and generalization capabilities of the model can be greatly enhanced with the help of data augmentation [42][43]. To augment the existing dataset, the ImageDataGenerator function provided by TensorFlow is utilized. This generator applies a series of transformations, including rotation, shifting, shearing, zooming, and horizontal flipping, to the original images. These augmentations introduce variations in the preprocessed images, effectively expanding the diversity and richness of the dataset without the need for additional labeled samples.

The augmentation process exposes the model to a more comprehensive range of scenarios and variations commonly encountered in real-world scenarios. By simulating diverse conditions and viewpoints in real-time, the model can become more resilient to unseen data. Furthermore, data augmentation mitigates overfitting by introducing regularization, thereby improving the model's ability to generalize to unseen data [43].

Various augmentation techniques, along with the range over which these techniques were employed on the images, are shown in Table 3.1 below.

Parameter	Value
Rotation Range	20
Width Shift Range	0.2
Height Shift Range	0.2
Zoom Range	0.2
Shear Range	0.2
Horizontal Flip	Yes

Table 3.1: Data Augmentation Parameters

Rotation was applied up-to a maximum of 20 degrees. Some images have lesions situated in a particular corner of the image. Therefore, rotating too much can cause part of the lesion to be lost.

In order to mimic occasional imperfections in real-life dermoscopic images, the Width and Height Shifts introduce slight changes in the images that would result due to different camera positions.

The Shear range parameter brings about tiny distortions that may be the work of imperfection in lenses. Furthermore, random zooming of images helps the model to train on scaled images. Horizontal Flipping mirrors each image and trains the model to effectively recognize mirrored objects as the same class.

These varied image manipulations led to a more diverse and representative image set, allowing the deep learning model to learn robust features generalizable to unseen data.

3.5 Feature Extraction

Feature Extraction in this study is implemented by leveraging the power of Transfer Learning [52]. The choice of the model for Transfer Learning and construction for the feature extraction process is discussed in this section.

3.5.1 Transfer Learning

Transfer learning presents a powerful tool for tackling the challenges associated with small image datasets [44][45]. Its ability to leverage pre-trained knowledge significantly reduces training time, improves performance, and mitigates overfitting, paving the way for effective utilization of sophisticated deep learning architectures even with limited data [46][47][48][49]. Complex deep architectures often require vast amounts of data for effective training. Transfer learning allows utilizing powerful deep architectures even with small datasets, unlocking their potential in various applications.

Pre-trained models have learned intricate relationships between pixels and semantic concepts through training in massive datasets. Leveraging these pre-extracted features, often from the early layers of the model, provides a rich representation of the input image, bypassing the need to learn these features from scratch with limited data.

In the context of the number of images available, the dataset used in this study may not be termed as ‘small’, but it’s not enough either. Moreover, the huge imbalance in classes has notably hampered the performance of machine learning and deep learning models. Keeping these factors in mind, the transfer learning approach was chosen for the extraction of features.

3.5.2 Choice of Model : EfficientNetB7

Choosing the pre-trained model for feature extraction purposes plays a critical role in this study. The model's capacity to capture complex patterns and nuances present in skin lesion images will directly impact the final classification results. It's crucial to consider the depth, width, and the presence of an adaptable architectural design, as these properties will ultimately play a role in the output.

The EfficientNetB7 model, with a size of 256 MB, boasts the highest Top-1 accuracy score of 84.3%. Its performance against other pre-trained models provided by Keras API [78] is visualized in Figure 3.5. The Top-1 accuracy is the strictest measure of performance provided by Keras, and it translates to how accurately a model predicts the exact class in image classification tasks.

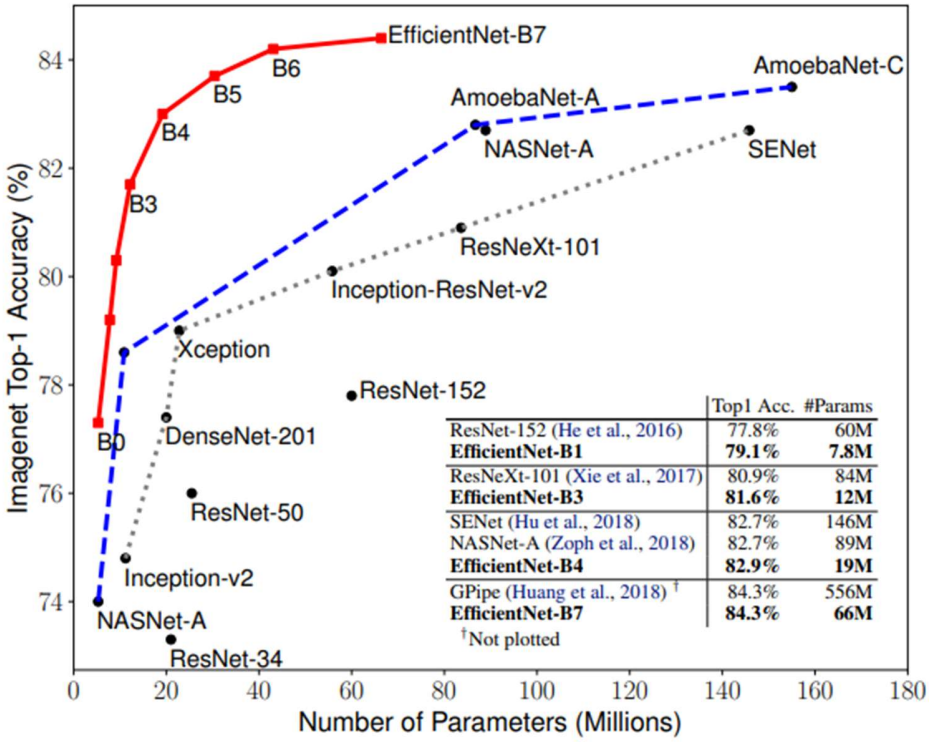

Figure 3.5: Performance of EfficientNetB7 compared to other pre-trained models [55]

EfficientNetB7 works with 66.7 million parameters, which is the one of the highest among EfficientNet architectures. More parameters generally imply a higher capacity for the model to learn intricate patterns and representations from the input data. However, it's still considerably lower than the VGG19, NASNetLarge, ConvNeXtBase. This doesn't decrease the appeal of EfficientNetB7, though, since a larger number of parameters also increases the risk of overfitting, especially in cases where data is relatively limited [79].

The depth of a neural network refers to the number of layers in the architecture, representing the depth of the feature hierarchy learned by the model. In the case of EfficientNetB7, with a depth of 438, the model encompasses a deep architecture with multiple layers of convolutional, pooling, and nonlinear activation functions. Deeper architectures allow the model to learn hierarchical representations of input data,

progressively capturing more abstract and complex features at higher layers. Figure 3.6 shows the main architecture of the EfficientNetB7 model, the '7' coming from the 7 blocks present.

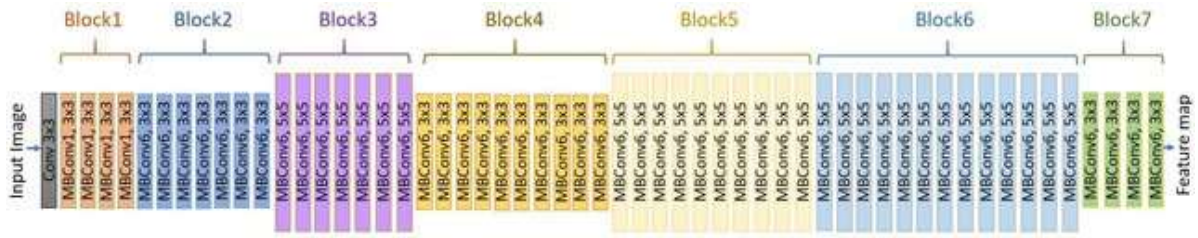

Figure 3.6: EfficientNetB7 model architecture [56]

The key to enhanced performances of the EfficientNetB7 is credited to the Mobile Inverted Bottleneck Convolution (MBConv) [56]. The MBConv architecture tackles computational challenges in convolutional neural networks by employing a bottleneck structure and leveraging depthwise and pointwise convolutions. The bottleneck structure compresses feature maps before applying depthwise convolutions. The depthwise convolutions operate on individual channels independently, extracting spatial information efficiently. Pointwise convolutions then expand channels to the desired output dimension, combining spatial information into new features. This approach enhances computational efficiency while preserving representational capacity, making MBConv well-suited for resource-constrained environments in tasks like image classification and feature extraction.

EfficientNetB7 leverages several MBConv blocks with different depths, strategically arranged in an architecture that intelligently scales parameters and computation based on specific coefficients, also known as the compound scaling approach.

3.5.3 Choice of Regularization Techniques

Employing regularization parameters following the operation of EficientNetB7 is needed to reduce model complexity, computational requirements, and overfitting issues.

GlobalAveragePooling2D [80] is chosen as the Pooling layer for this study. This pooling layer computes the average value of each feature map across all spatial locations. This means it takes the average of all the values in each feature map. Therefore, it effectively reduces the spatial dimensions of the feature maps to 1x1, collapsing them into a single value per feature map. Figure 3.7 below showcases the process visually.

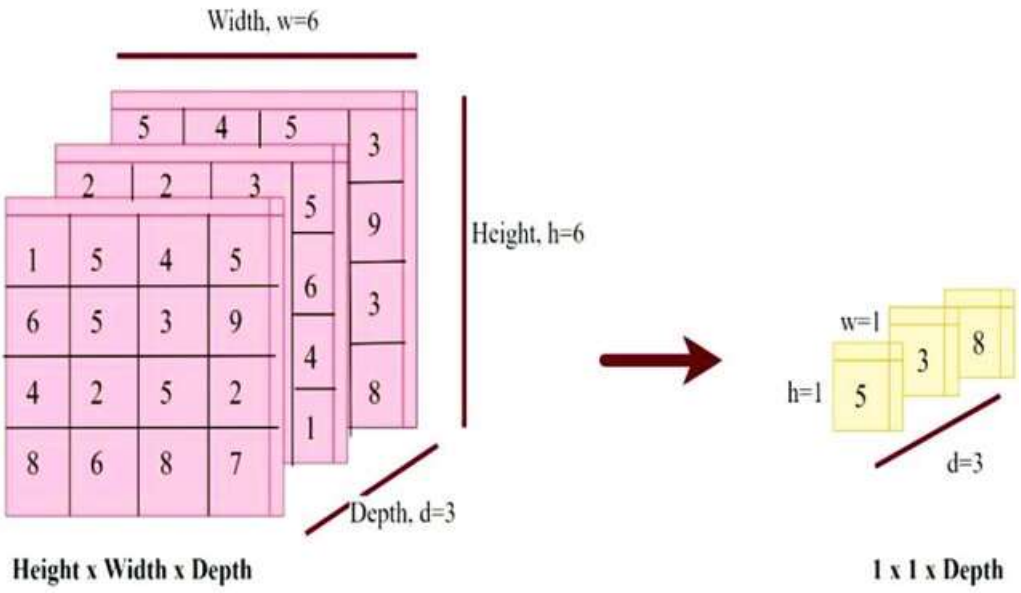

Figure 3.7: GlobalAveragePooling2D Mechanism [81]

This process helps simplify the data while retaining important information, making the model more robust to spatial translations and reducing overfitting. Additionally, it greatly

reduces the number of parameters in the model, which can help in training and inference efficiency.

GlobalAveragePooling2D excels in providing stronger regularization [82]. Since it computes the average value of each feature map, it helps in reducing overfitting by providing a smoother and less granular representation of the features.Furthermore, it's less sensitive to small spatial translations of features within the pooling regions and retains more information.

Dropout layers are also applied in conjunction with GlobalAveragePooling2D layer for more robust and generalizable features. Dropout helps prevent overfitting by randomly deactivating neurons during training. This prevents the model from relying too heavily on specific features or patterns in the training data that may not generalize well to unseen data. By encouraging the model to learn more generalizable features, Dropout can lead to better feature extraction. A general Dropout rate of 0.5 is implemented in this study.

3.5.4 Feature Extraction Architecture

At the start of the feature extraction process, the EfficientNetB7 is instantiated. The input image dimensions are set to 224X224 and the utilization of knowledge pertaining to the ImageNet dataset is also initialized. The output layer of the pre-trained model is replaced by the GlobalAveragePooling2D layer. This layer outputs single values from each feature map and concatenates or flattens them into a single vector. The flattened vector spaces are then passed through a Dropout layer with a rate of 0.5.

A fully connected layer with 1024 neurons and a ReLU activation [83] is added after the pooling and dropout layers. This layer further extracts new high-level features specific to the target classification task, building upon the pre-trained features from EfficientNetB7. A second dropout layer, also with a rate of 0.5, for further regularization.

The features extracted from every image within the dataset were systematically organized and stored within a NumPy array. The resultant array exhibited a final shape of (24703, 1024), signifying the representation of 24703 images across 1024 distinct feature spaces. Figure 3.8 visualizes the entire methodology of the feature extraction process.

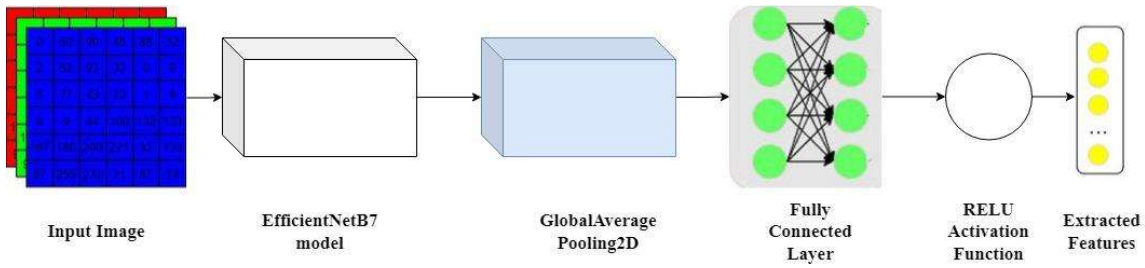

Figure 3.8: Feature Extraction Workflow [84]

3.6 Mitigation of Class Imbalance

3.6.1 Significance of Class Imbalance

There is a massive difference in the number of skin lesion images for each class in the dataset. NV has the most number of images, amounting to 12875, whereas the number of images under the category of DF is only 239.

Such an imbalance introduces bias into the learning process, as machine learning and deep learning algorithms tend to prioritize the majority class due to its higher prevalence. Consequently, the model may become overly biased towards predicting the majority class, resulting in poor generalization performance and reduced predictive accuracy, particularly for minority classes [86].

This issue is exacerbated in scenarios where the minority class contains critical or rare events that are of significant interest. In such cases, the underrepresentation of the minority class can lead to serious consequences, as the model's inability to accurately detect or classify these instances may result in missed opportunities or erroneous decisions.

Furthermore, the imbalance in class distribution can impede the learning process by hindering the model's ability to learn discriminative features associated with minority classes. As a result, the model may struggle to distinguish between different classes effectively, leading to suboptimal performance and compromised decision-making capabilities [86].

Addressing class imbalance is therefore crucial in machine learning and deep learning problems to ensure fair and accurate model predictions across all classes. Techniques such as oversampling, undersampling, and the use of specialized algorithms like SMOTE (Synthetic Minority Over-sampling Technique) [69] are commonly employed to rebalance the class distribution and mitigate the adverse effects of imbalance on model performance.

3.6.2 Choice of Class Rebalancing Technique

The specialized algorithm, SMOTE, is chosen to deal with the discrepancy of categories in this image dataset. SMOTE operates by synthesizing new instances for the minority class, thus balancing the class distribution within the dataset [69].

The method begins by identifying minority class instances that are in close proximity to one another in feature space. Then, synthetic instances are generated along the line segments connecting these neighboring instances.

SMOTE effectively augments the minority class by introducing synthetic examples that resemble the existing minority class instances. This process helps to alleviate the issue of class imbalance by ensuring that each class is represented more equally in the training data.

The SMOTE function, present in the imbalanced-learn open-source library, was applied to the Numpy array consisting of the extracted features. The count for each label increased up-to the instances of the majority class, 12875.

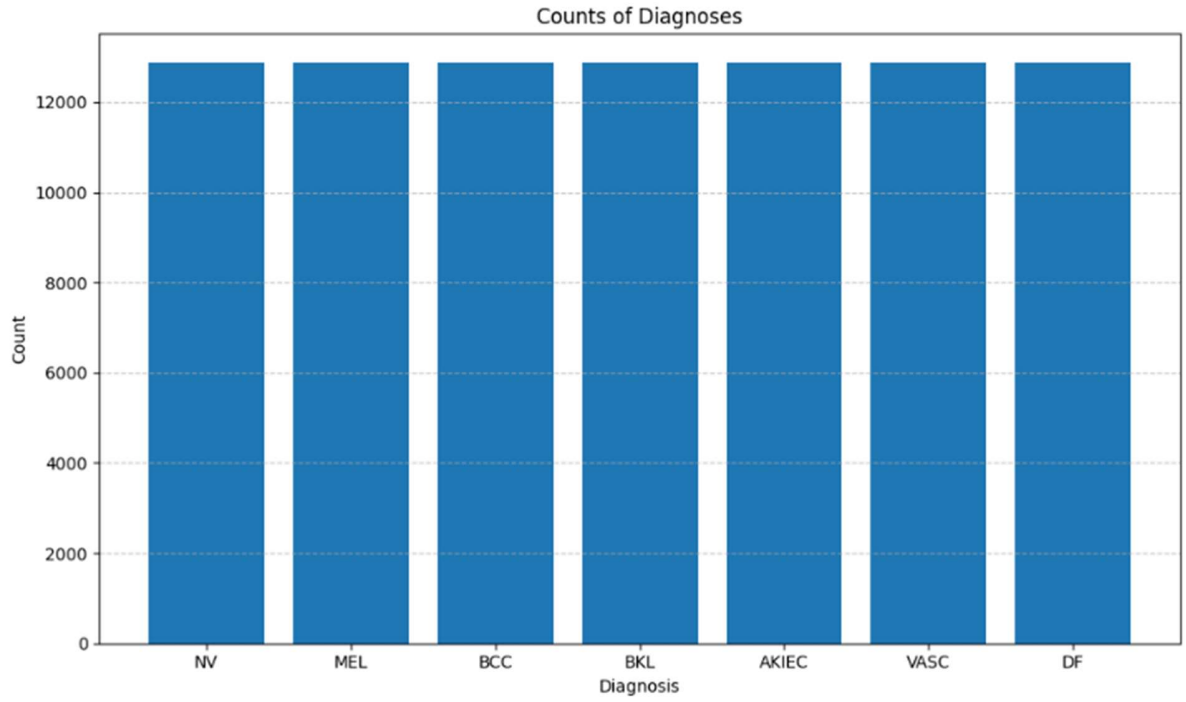

Figure 3.9: Distribution of classes in the Features Extracted dataset after SMOTE

Figure 3.9 shows the balanced features dataset after applying SMOTE. Now, it's in optimal condition for training robust machine learning and deep learning models. By leveraging this balanced dataset, models can learn from a more diverse and representative set of examples, leading to improved generalization performance and enhanced predictive accuracy across all classes.

3.7 Skin Lesion Classification

The skin lesion classification in this study is carried out using a 1-D Convolutional Neural Network [71]. The neural network is designed to capture patterns in the extracted features for proper classification. In this section, the various aspects related to the classification are discussed.

3.7.1 Convolutional Neural Network (CNN)

Convolutional Neural Networks (CNNs) [85] have the exceptional ability to automatically learn hierarchical features from raw image data by harnessing the power of convolution to extract and combine meaningful patterns, ultimately leading to robust and accurate image classification. The fundamental unit of a CNN is the convolutional layer, where numerous filters act as feature detectors. These filters, akin to sliding kernels, scan the image, capturing local information by computing element-wise multiplications and summations. By employing various filter sizes and applying them through multiple layers, the network progressively extracts increasingly high-level features. Figure 3.10 shows a very basic working of a convolutional layer.

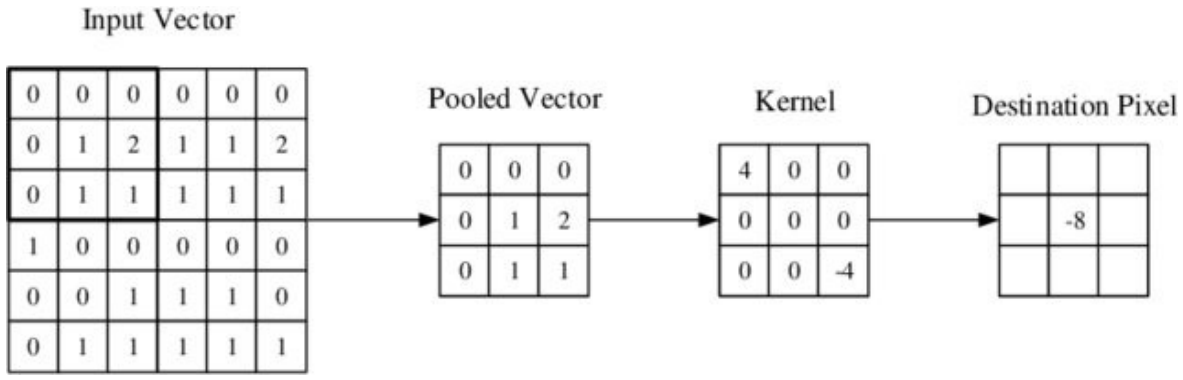

Figure 3.10: Working of a Convolutional Layer

CNN forms a hierarchy of abstractions on the initial features extracted. Subsequent layers combine the outputs of previous layers, learning more complex features like shapes, objects, and ultimately, the entire image content. This hierarchical structure allows the network to progressively capture progressively more global information, culminating in the final classification layer that assigns a label to the image. In Figure 3.11, the workflow of a typical CNN architecture is shown.

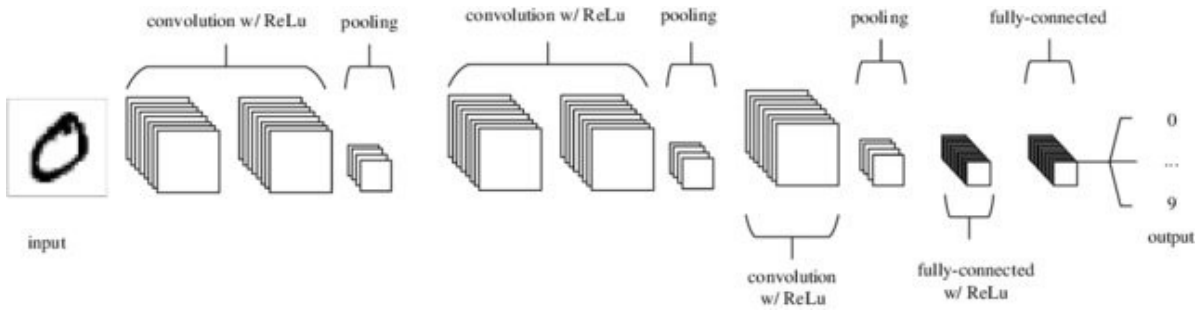

Figure 3.11: Workflow of a CNN architecture [87]

CNNs excel at analyzing 2D spatial information in images, but when the information is present in one-dimensional structure, such as the array of Features Extracted in this study, the powerful paradigm of feature extraction through convolutions extends to the 1-dimensional Convolutional Neural Networks (1D-CNNs).

1D-CNNs can capture sequential relationships within the feature array. This becomes crucial when the order of features holds inherent meaning. Therefore, 1D-CNNs usually work well with raw data and have been commonly used in Raw Signal Classification. A sample 1-D CNN structure utilized for signal classification is shown in Figure 3.12 below.

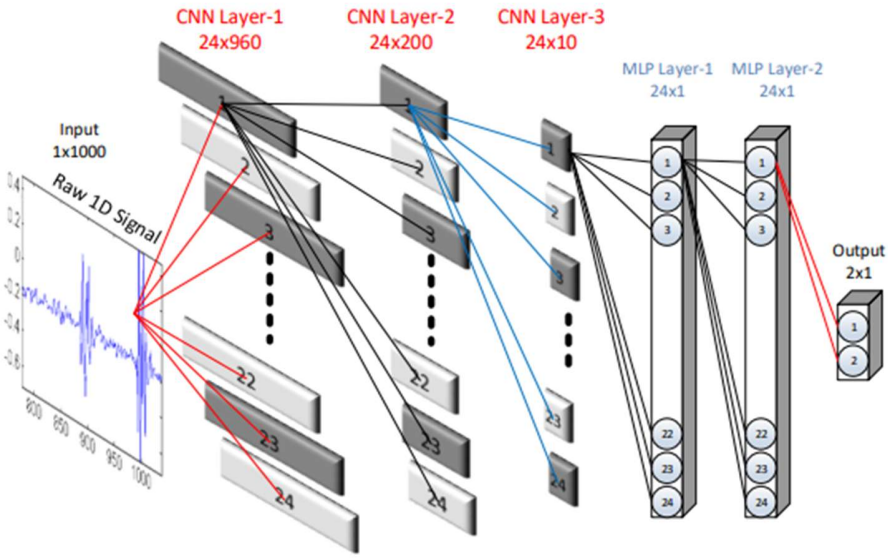

Figure 3.12: A Sample 1-D CNN Configuration [71]

The same concept has been applied to the raw data of extracted features in this study. In addition to that, the architecture was integrated with much more secondary parts for more efficient training and performance.

3.7.2 Deep Learning Model Architecture

The proposed model architecture constitutes a sequential arrangement of convolutional and dense layers, tailored for a specific task. The input data, representing sequences of length 1024 with a single channel, undergoes a series of transformations to extract meaningful features. The architecture commences with a convolutional layer, employing 32 filters of size 3, followed by a rectified linear unit (ReLU) activation function to introduce non-linearity. Subsequently, a max-pooling layer is applied to down-sample the spatial dimensions.

Another convolutional layer with 64 filters of size 3 is employed, further enriching the feature representation. Similar to the preceding layer, ReLU activation is applied to introduce non-linearity. Subsequent max-pooling further compresses the spatial dimensions, facilitating the extraction of higher-level features. The resultant feature maps are then flattened to a one-dimensional vector, preparing them for processing by densely connected layers.

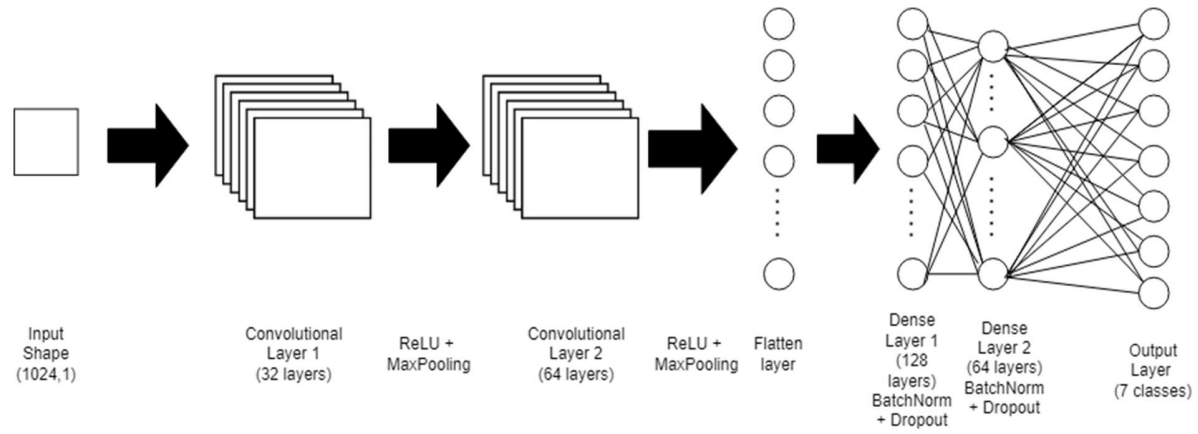

Figure 3.13: Proposed 1-D CNN Architecture

The dense layers serve to integrate the spatial information captured by the convolutional layers, promoting learning of complex patterns. A dense layer with 128 units, followed by ReLU activation, is introduced to facilitate feature transformation. Batch normalization is employed to enhance training stability and accelerate convergence by normalizing the input to each layer.

Furthermore, dropout regularization is incorporated to mitigate overfitting by randomly deactivating a fraction of neurons during training. Another dense layer, with 64 units and ReLU activation, further refines the learned representations. Batch normalization and dropout regularization are subsequently applied to this layer to sustain model generalization.

The final layer, comprising 7 units and employing the softmax activation function, outputs the probabilities of the input belonging to each of the 7 target classes. Figure 3.13 depicts the complete architecture of the classification model developed in this study. Moreover, Table 3.2 lists the model parameters, along with their output shape, and Activation Functions, where applicable.

Layer (Type)	Output Shape	Activation Function
Conv1D	(None, 1022, 32)	ReLU
MaxPooling1D	(None, 511, 32)	
Conv1D	(None, 509, 64)	ReLU
MaxPooling1D	(None, 254, 64)	
Flatten	(None, 16256)	
Dense	(None, 128)	ReLU
Batch Normalization	(None, 128)	
Dropout	(None, 128)	
Dense	(None, 64)	ReLU
Batch Normalization	(None, 64)	
Dropout	(None, 64)	
Dense	(None, 7)	Softmax

Table 3.2: Proposed 1-D CNN model parameters

The model is compiled using the Adam optimizer, minimizing the sparse categorical cross-entropy loss function, and is evaluated based on metrics such as Accuracy, F1 score, Recall score, and Precision Score.

3.7.3 Training

The dataset is divided in a 70-30 ratio, with 70% being the training set, and the 30% being the testing set. The train_test_split function of the Sklearn library was applied to make this change. Random state was set to the default value of 42.

The model is trained using the Adam optimizer, which boasts adaptive learning rate capabilities. The optimizer seeks to minimize the sparse categorical cross-entropy loss function, which is well-suited for multi-class classification tasks.

A default batch size of 32 samples per gradient update is utilized, balancing computational efficiency with gradient accuracy. Training progresses over 200 epochs, allowing the model to iteratively learn from the training data.

During training, the model's performance is evaluated using the accuracy metric, providing insights into its classification prowess.

Furthermore, the model incorporates the EarlyStopping callback mechanism, a regularization technique aimed at preventing overfitting. This strategy monitors the model's performance on a validation set and halts training if there is no improvement in the validation loss over a certain number of epochs, thus avoiding unnecessary computation and mitigating the risk of overfitting.

Moreover, for a more robust evaluation, a K-Fold cross-validation approach is employed. Specifically, the dataset is partitioned into five folds, ensuring that each fold serves as both

a training and validation set across five iterations. This process facilitates a more reliable estimation of the model's generalization performance by assessing its consistency across different data subsets. Table 3.3 lists the model training parameters involved in this study.

Parameter	Value
Optimizer	Adam
Learning Rate	0.001 (Default)
Batch Size	32 (Default)
Metrics	Accuracy, F1 Score, Recall, Precision
Epochs	200
Train-Test Split	30%
Loss Function	Sparse Categorical Crossentropy

Table 3.3: Model Training Parameters

Chapter 4

Results and Discussion

4.1 Performance Metrics

Various performance metrics were utilized in this study for a comprehensive evaluation of the model's performance in classification tasks.

In classification tasks, True Positive denotes instances where the model correctly predicts a positive outcome when the actual outcome is positive. As for True Negative, it takes place when the model correctly predicts a negative outcome when the actual outcome is negative.

Furthermore, situations where the model incorrectly predicts a positive outcome when the actual outcome is negative is classified as False Positive, and where the model incorrectly predicts a negative outcome when the actual outcome is positive is classed as False Negative.

The Accuracy score metric represents the proportion of correctly classified samples over the total number of samples. It provides a general measure of the model's overall correctness in predicting the class labels.

Precision score is the ratio of true positive predictions to the total number of positive predictions. It measures the accuracy of positive predictions and is particularly relevant in scenarios where the cost of false positives is high.

Recall, also known as sensitivity or true positive rate, is the ratio of true positive predictions to the total number of actual positive instances in the dataset. It quantifies the model's ability to correctly identify all positive instances, without missing any. Figure 4.1 consists of the formula to calculate Precision and Recall scores.

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

$$\text{Recall} = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

Figure 4.1: Precision and Recall Formula

F1 score is the harmonic mean of precision and recall. It provides a balance between precision and recall, taking into account both false positives and false negatives. F1 score is especially useful when dealing with imbalanced datasets. It can be calculated by the formula denoted in Figure 4.2.

$$\text{F1 Score} = \frac{2}{\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Figure 4.2: F1 Score Formula

A confusion matrix is a table that visualizes the performance of a classification model by presenting the actual and predicted class labels. It provides insights into the model's performance across different classes, highlighting true positives, true negatives, false positives, and false negatives. A basic structure of the confusion matrix is visualized in Figure 4.3.

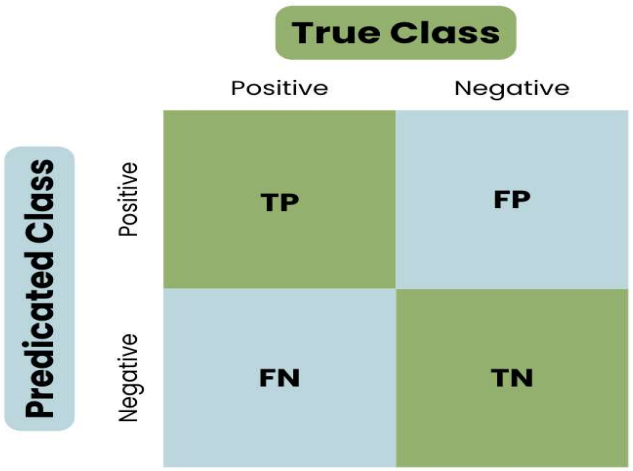

Figure 4.3: Confusion Matrix

Training-validation accuracy and loss curves are plots that illustrate the model's learning dynamics during the training process. The accuracy curve shows the model's accuracy on

the training and validation sets over successive epochs, while the loss curve depicts the corresponding training and validation loss values. These curves help in assessing the model's convergence, overfitting, and generalization ability.

4.2 Model Performance

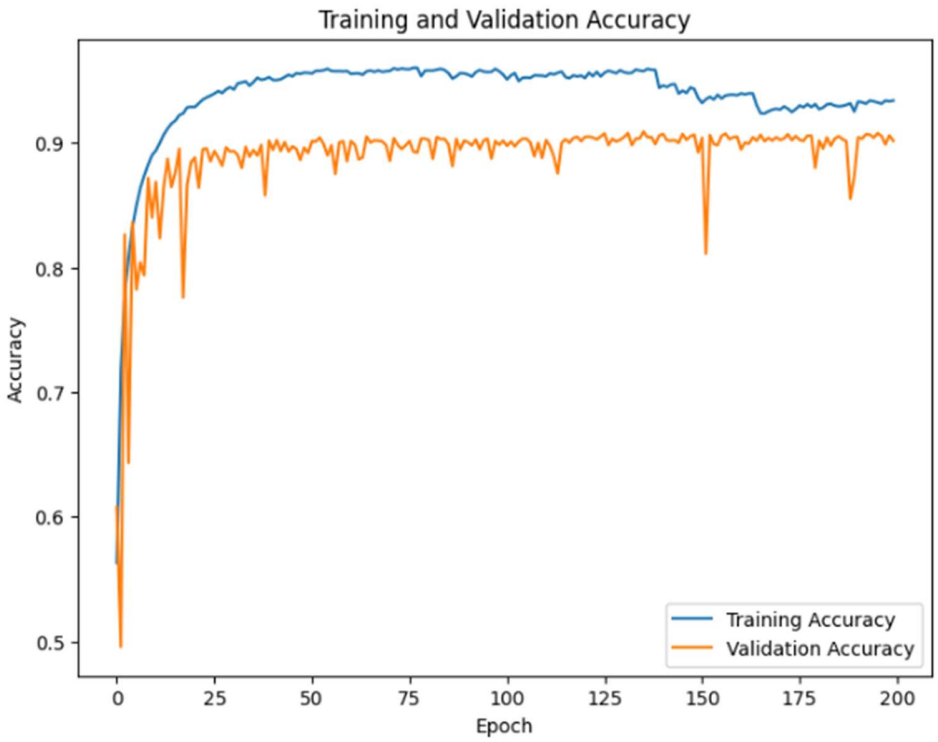

Figure 4.4: Training and Validation Accuracy of Proposed Classification Model

In Figure 4.4 above, the variation of Training and Validation accuracies are shown with increase in Epochs. The final training and validation accuracies achieved were 93.13% and 90.2%, respectively. The variation between both curves is also very minimal, suggesting that the model generalized well to unseen information and didn’t overfit in the training data.

The testing dataset consisted of 27201 samples and the model classified 24481 of them accurately, amounting to 90% accuracy score. The classification report for each category of lesion and the associated metrics are shown in the image below.

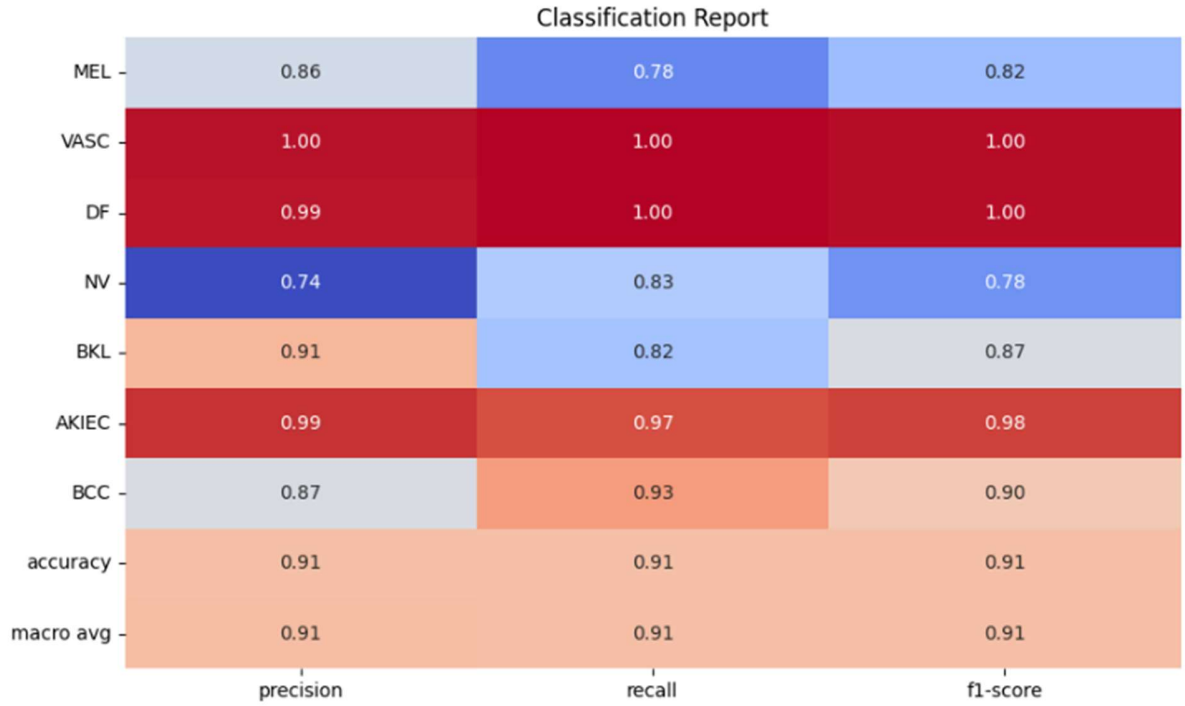

Figure 4.5: Classification Report of Proposed Classification Model

Based on the F1 score in Figure 4.5, which is the harmonic mean of both Precision and Recall scores, it can be observed that the model performed best for VASC, DF, and AKIEC classes, correctly predicting almost all the samples. There's no huge discrepancies of scores between the Precision and Recall metrics, strengthening the reliability of the model's results.

The confusion matrix for each label, shown in Figure 4.6 below, gives a further understanding of the model's performance compared to each other.

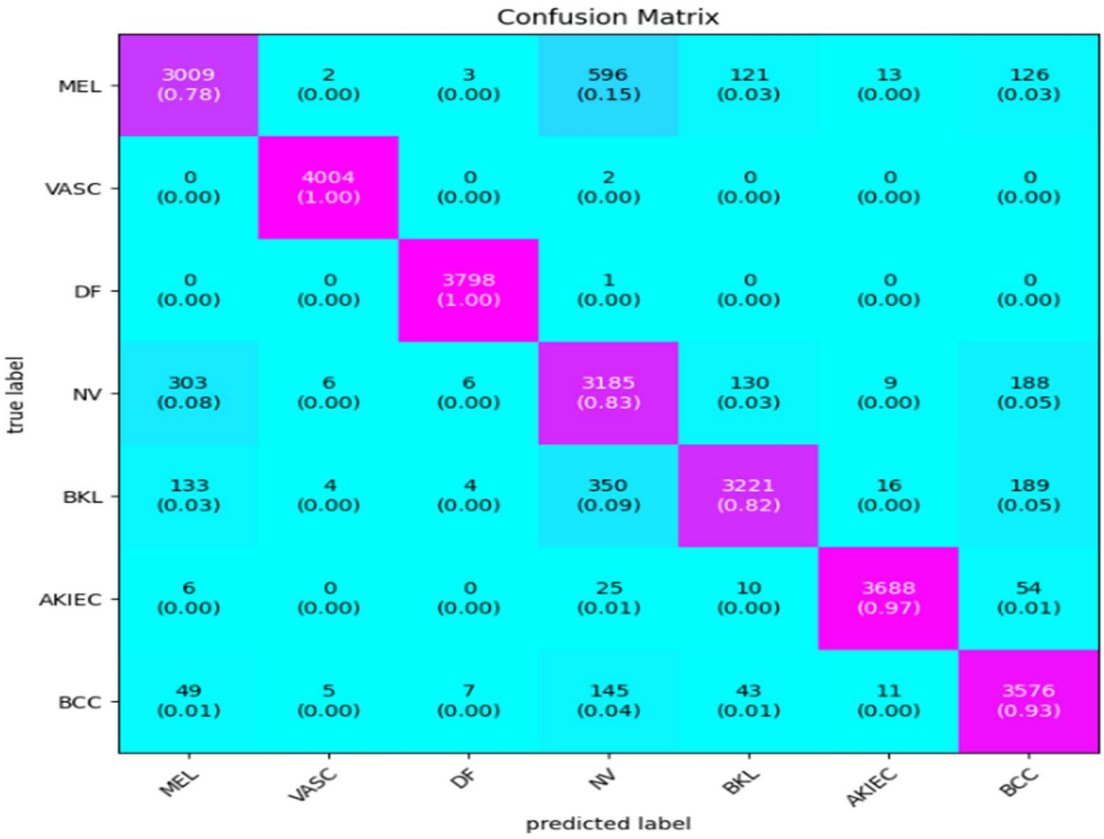

Figure 4.6: Confusion Matrix of Proposed Classification Model

The confusion matrix in the figure above suggests almost accurate performance of the model to classify lesions into DF and VASC. Although no serious discrepancies were noticed, NV is most commonly misclassified. Majority of the unseen, testing samples of the model were classified correctly.

The performance of the convolutional neural network (CNN) model in a multiclass classification task was evaluated further using a 5-fold cross-validation approach. Due to computational limitations, early stopping was employed, which is why even though 200 epochs were set for each fold, the maximum number of epochs spanned by the folds was 4.

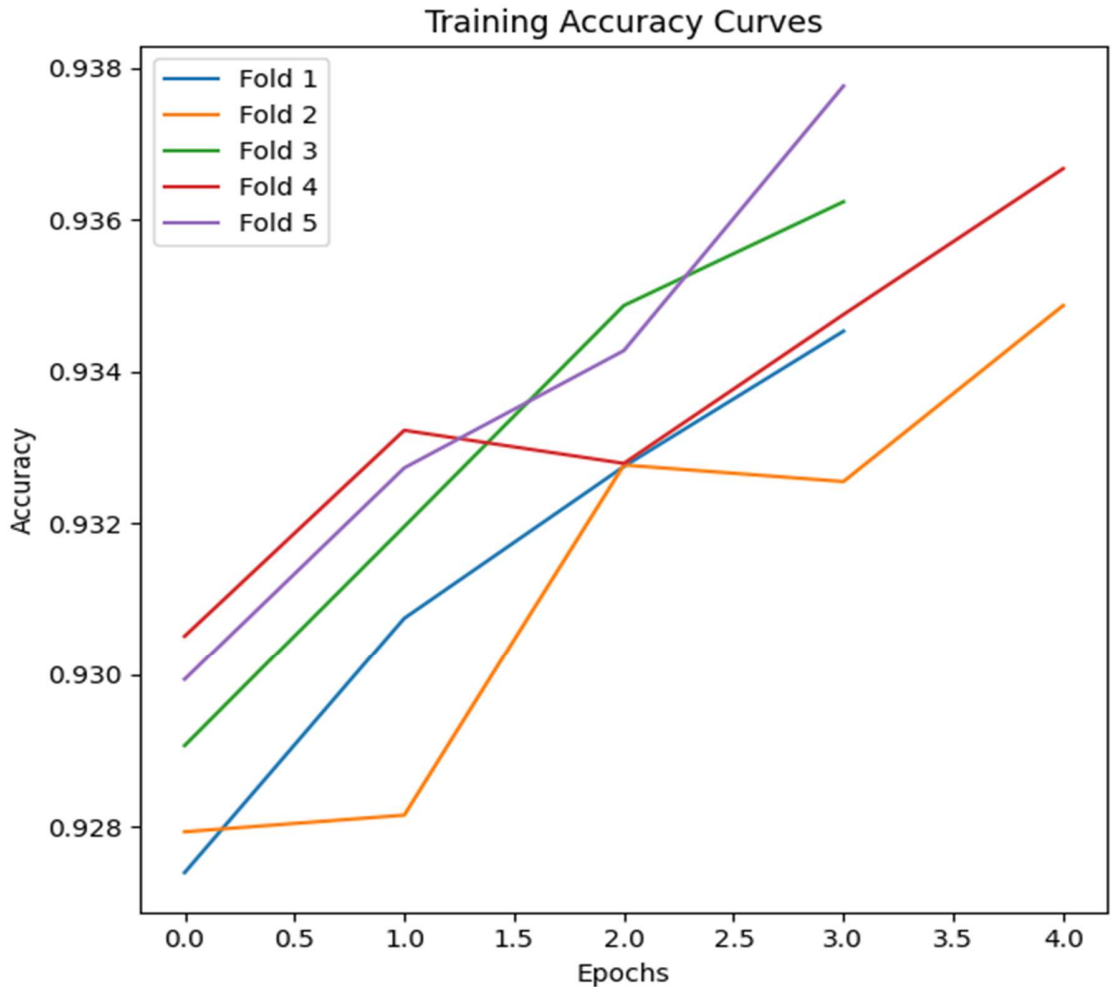

Figure 4.7: Training Accuracy Curves of K-Fold Cross Validation

It can be observed from Figure 4.7 that in the first fold, early stopping occurred in Epoch 4, with a training accuracy of 93.45% and a testing accuracy of 92.00%. Similar early stopping patterns were observed in subsequent folds, with training accuracies ranging from 92.61% to 95.71%.

The training accuracy curves show an increase in accuracy with epochs. For the second fold, a slight decrease was noticed while moving from Epoch 2 to Epoch 3. However, it was made up for in the following epochs, when the accuracy surpassed 93.4%.

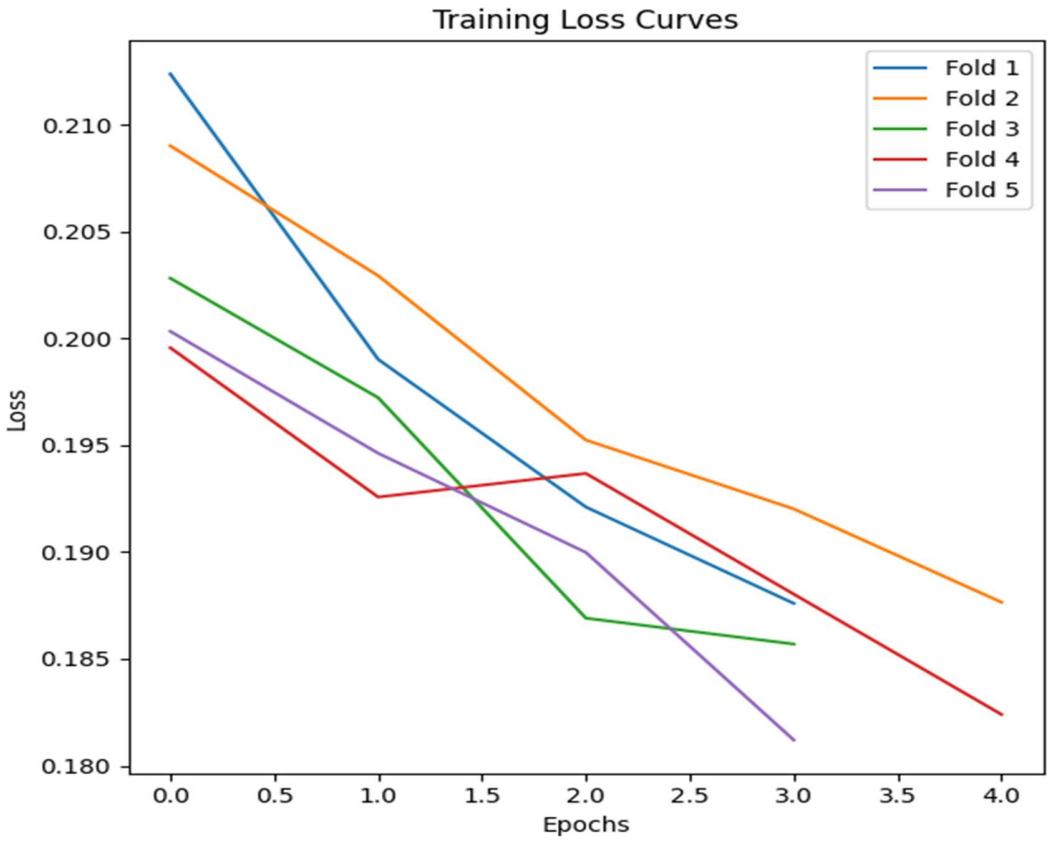

Figure 4.8: Training Loss Curves of K-Fold Cross Validation

The training loss curves across each fold in Figure 4.8 also denote a clear decrease in loss with increase in epochs. Further analysis was carried out on the test set performance of the folds. The mean F1 score for each fold is visualized in Figure 4.9 below.

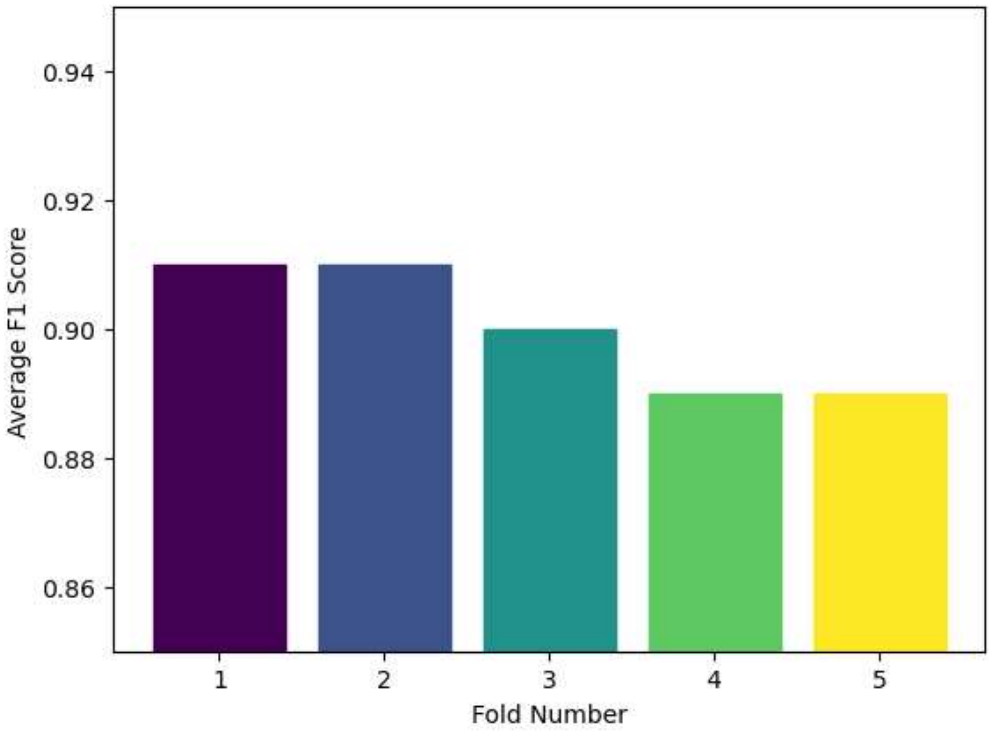

Figure 4.9: Average F1-Score per Fold

The scores vary around 90% and are very close to each other. The consistency of the F1-scores across multiple folds indicates that the model generalizes well to unseen data and is not overly sensitive to variations in the training data.

4.3 Further Explorations

Each and every layer and parameter initiated in the model architecture and training impacts the final classification performance of the model. Apart from the model, the preceding preprocessing techniques heavily impact how the model learns training data as well, ultimately playing a major role in the final output. The impact of certain changes is explored in this section.

4.3.1 Altering the Number of Layers

The impact of altering the 1-D convolutional layer architecture used for classification in this study was monitored. The aim to assess how removing or adding a convolutional layer influences the learning dynamics and overall accuracy of the model.

To achieve this, two separate 1-D CNN architectures were constructed with the original model serving as the baseline. The max-pooling, dropout, and dense layers were kept the same. Subsequently, two modified versions were created: one with one fewer convolutional layer and another with one additional convolutional layer than the main model.
Each model was trained using the same dataset and hyperparameters, and their training and validation accuracies were monitored over 80 epochs.

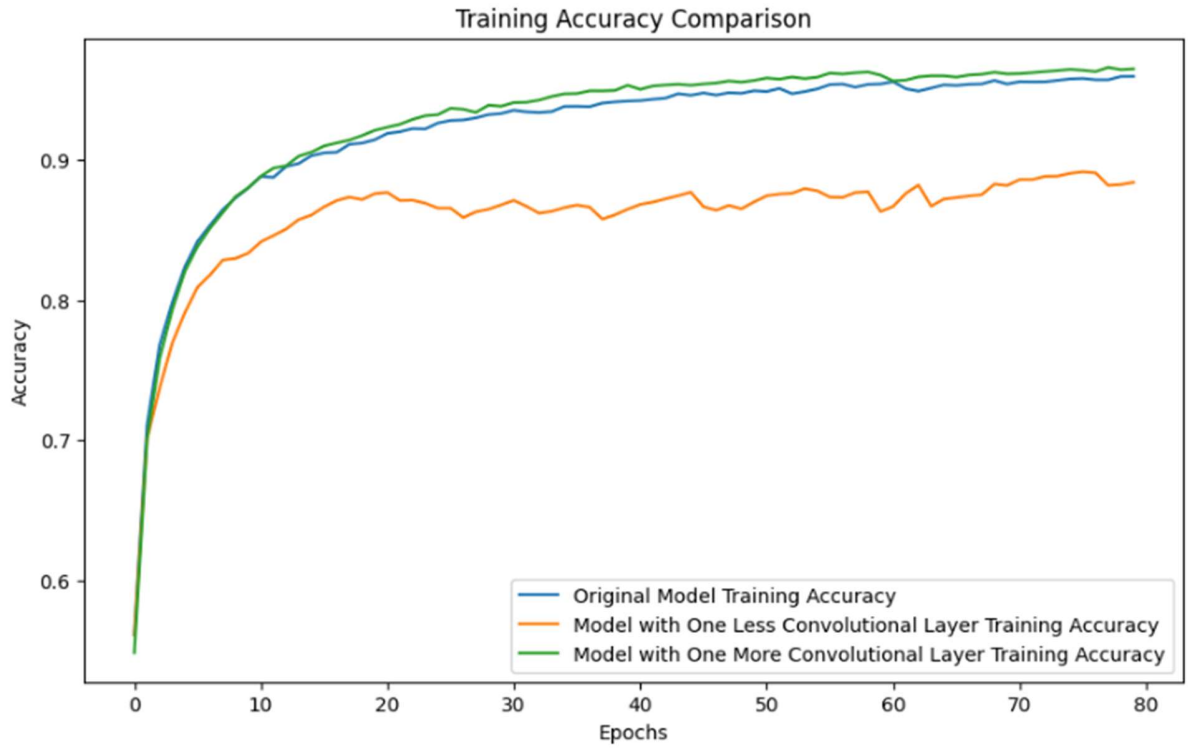

Figure 4.10: Training Accuracy Comparison for Different Layers

It can be easily concluded from Figure 4.10 that the model's performance is significantly hampered when one convolutional layer is removed. Furthermore, the original model and the model with one more convolutional layer are on par with each other, but the original model edges it in the final set of epochs.

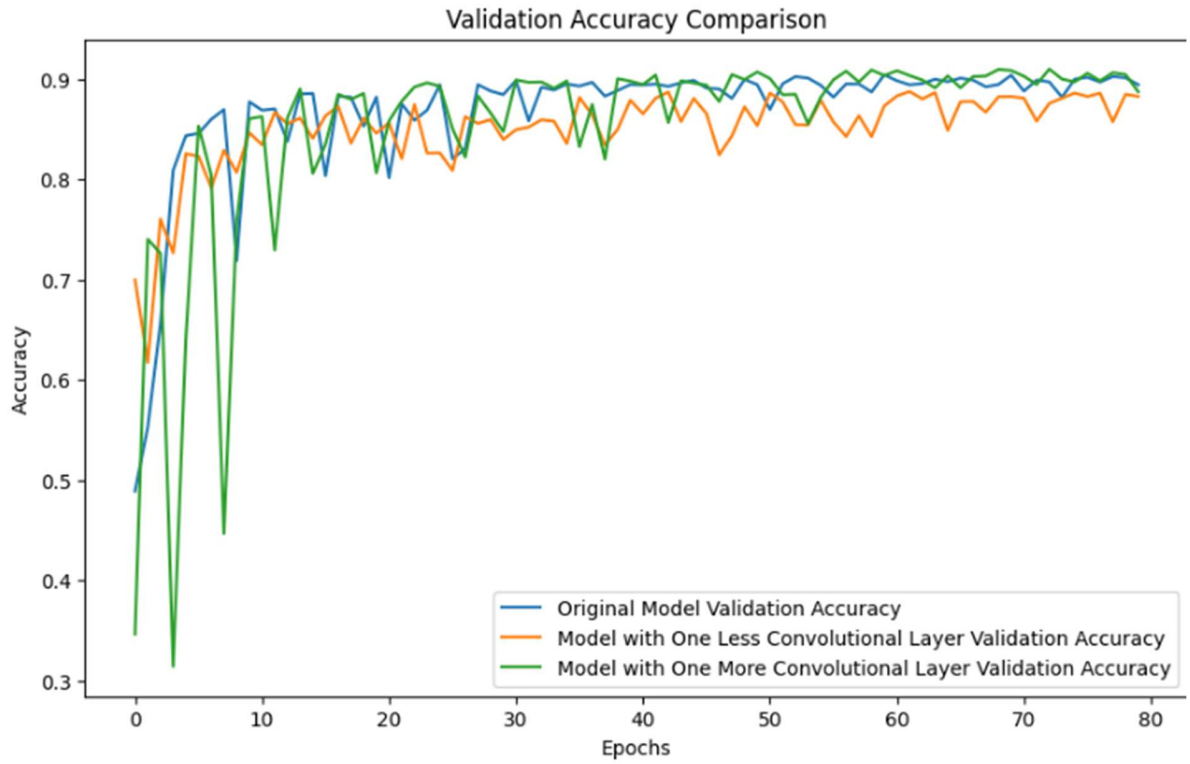

Figure 4.11: Validation Accuracy Comparison for Different Layers

The validation accuracy curves in Figure 4.11 show similar results. The model with one less convolutional layer has a stably increasing validation accuracy, but it's lesser than the other two. The performance of the model with one more convolutional layer than the original model is very unstable, constantly moving to and from extremities. Ultimately, the original model excels, boasting greater accuracy while maintaining stability and lower computation power.

4.3.2 Changing the Number of Filters

The classification power of CNN is impacted by the number of filters in each convolutional layer. These play a crucial role in balancing model capacity, performance, and efficiency.

While increasing filters leads to richer feature extraction and potentially better accuracy, it also brings challenges like longer training times and the risk of overfitting. Optimal filter selection hinges on various factors, including dataset size, desired model complexity, hardware constraints, and network architecture.

Akin to the original model, two new models are constructed. One of the models incorporates a higher number of filters in the first convolutional layer, while the other model increases the number of filters in the second convolutional layer. The number of filters in subsequent layers of each model are kept the same as the original.

The models are compiled using the Adam optimizer and sparse categorical cross-entropy loss function. With 20% data reserved for validation, the models are trained for 80 epochs.

The validation accuracies of the models in Figure 4.12 suggest that they exhibit comparable levels of effectiveness in classifying data.

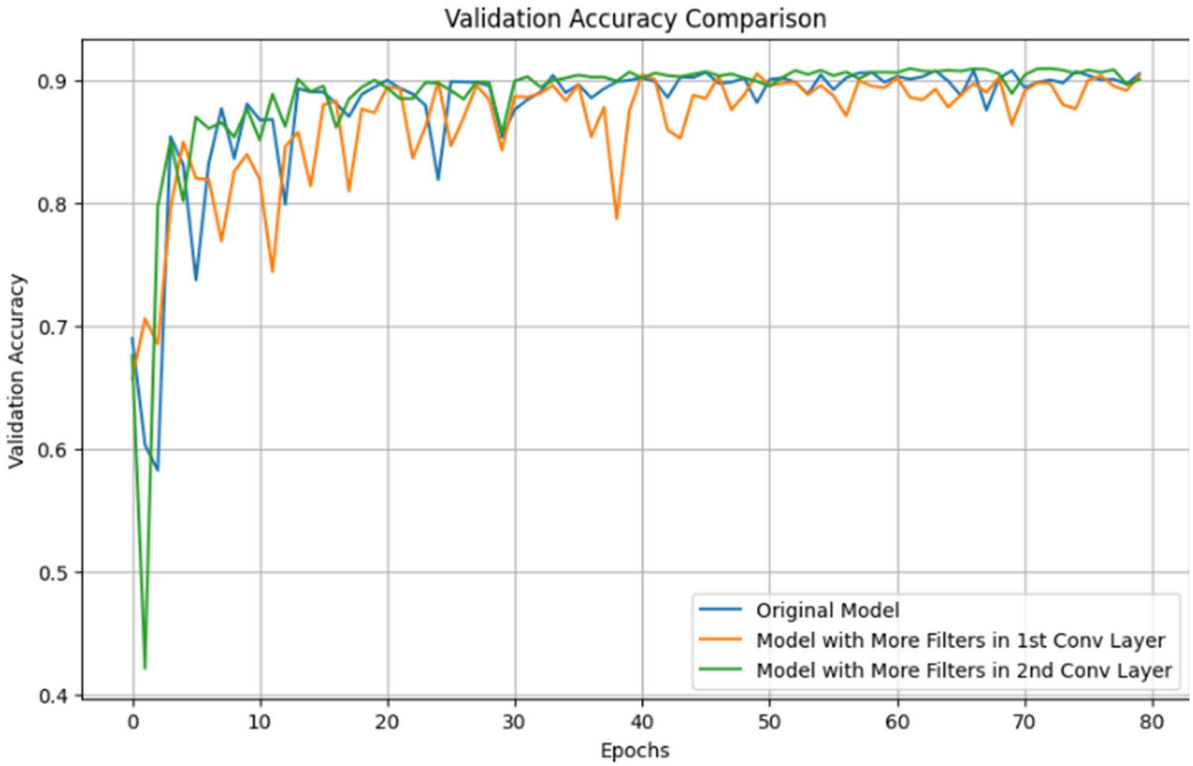

Figure 4.12: Validation Accuracy Comparison for Different Filter Constructions

A closer scrutiny of the validation accuracy graphs reveals differences in their learning behaviors. The model with more filters in the first convolutional layer displays a relatively higher degree of instability, characterized by spikes and fluctuations in accuracy throughout. In contrast, the model with more filters in the second convolutional layer demonstrates greater stability, with a smoother and more consistent trajectory of accuracy improvements over epochs.

While all models achieve similar validation accuracy in the end, their learning processes differ in terms of stability and convergence behavior. The instability observed in the model with more filters in the first convolutional layer indicates potential challenges in training dynamics, possibly due to increased complexity or overfitting tendencies. In contrast, the

model with more filters in the second convolutional layer exhibits a more robust and stable learning trajectory.

4.3.3 Overfitting Issues

SMOTE serves as a pivotal tool for addressing overfitting concerns, particularly in scenarios where class imbalances exist. It's used in this study to deal with the massive imbalance in numbers among classes.

The consequences of not employing SMOTE is investigated by systematically preparing a dataset without the application of SMOTE. The dataset undergoes all preceding preprocessing techniques. Subsequently, a 1-D CNN architecture is trained on this dataset without the application of SMOTE. The training process spans 200 epochs, utilizing the Adam Optimizer and employing the Sparse Categorical Crossentropy loss function.

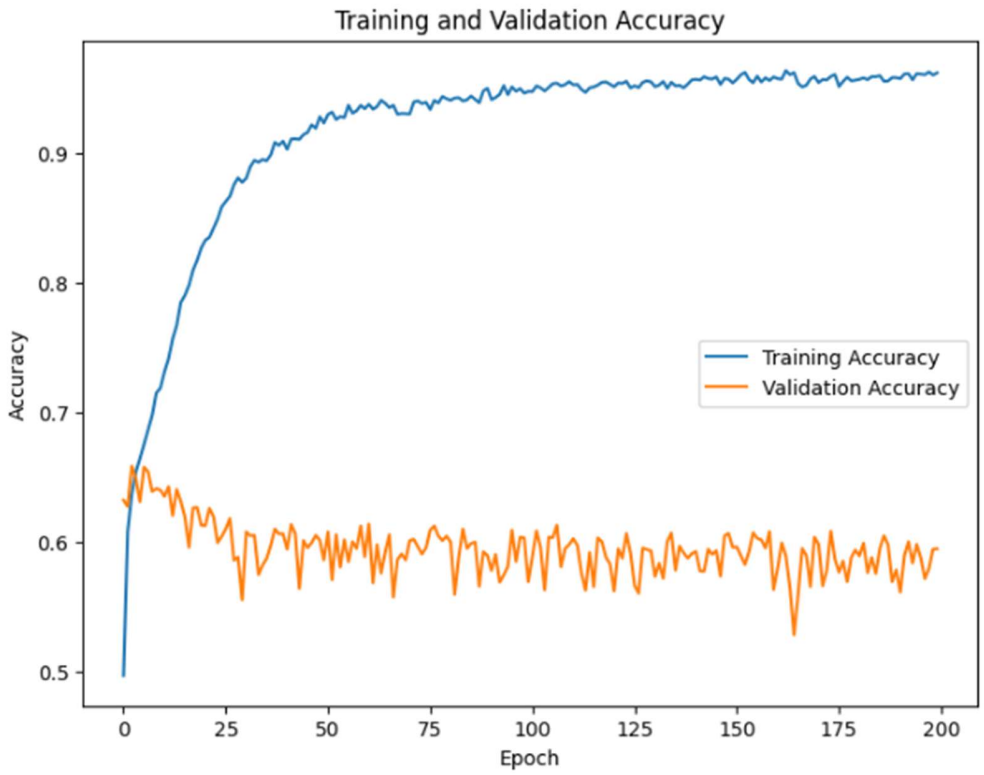

Figure 4.13: Training and Validation Accuracy in the absence of SMOTE

The huge difference between the training and validation performance is apparent from the Accuracy and Loss curves in Figure 4.13 and Figure 4.14. The high training accuracy of 96.20% suggests that the model has become excessively tuned to the nuances present within the training dataset. Consequently, when evaluated on previously unseen validation data, the model struggles to generalize effectively, resulting in a significantly lower validation accuracy of 59.40%.

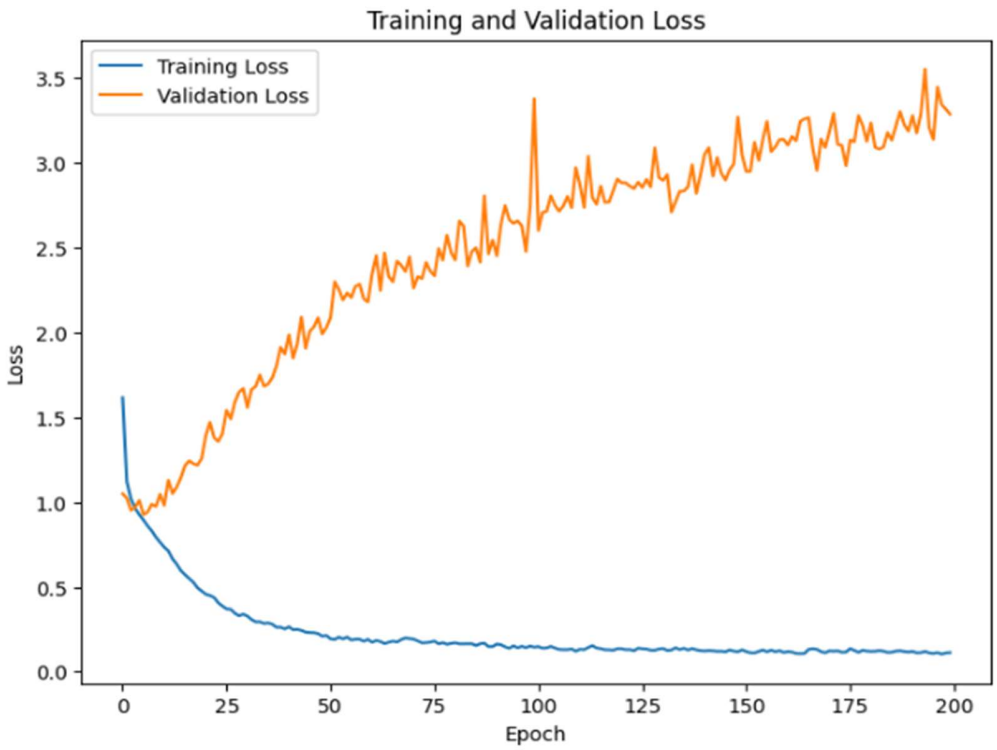

Figure 4.14: Training and Validation Loss in the absence of SMOTE

The significance of overfitting lies in its potential to undermine the model's predictive capabilities on new, real-world data. An overfitted model may exhibit poor generalization, leading to inaccurate predictions and reduced reliability in practical applications. Therefore, mitigating overfitting is crucial for ensuring the robustness and effectiveness of machine learning models.

Chapter 5

Conclusion and Future Works

A comprehensive exploration of dermatological image analysis, aiming to develop an effective framework for lesion classification was carried out in this study. Meticulous image preparation of the ISIC-2019 skin lesion image dataset, followed by image preprocessing techniques were applied to enhance image quality, including removal of unwanted artifacts such as hair. Data augmentation methods were employed to enrich the dataset and ensure comprehensive representation of lesion characteristics.

The inherent capabilities of a pre-trained model, EfficientNetB7, was harnessed to extract meaningful features from the skin lesion images. The feature extraction process was optimized through fine-tuning parameters such as dropout rates and pooling mechanisms to enhance feature discernibility.

Moreover, to address the challenge of class imbalance and mitigate the risk of overfitting, the Synthetic Minority Over-sampling Technique (SMOTE), a powerful algorithm for augmenting minority class samples, was applied to the extracted features. By rebalancing the dataset and promoting equitable representation of all classes, SMOTE played a crucial role in enhancing the robustness and generalization capabilities of our model.

Finally, a custom deep learning model, a 1D convolutional neural network (CNN), tailored specifically for lesion classification was created. This model, comprising two convolutional layers, dropout regularization, max-pooling, and dense layers, served as the cornerstone of the predictive framework, enabling accurate classification of dermatological lesions based on the extracted features.

Moving forward, future endeavors will focus on further refining the model architecture, exploring novel techniques for feature extraction, and expanding the scope of the analyses to encompass a broader array of dermatological conditions.

The implementation of optimization algorithms on selecting the most sensible features from the extracted array could also be explored in future research endeavors, offering potential avenues for improving model robustness and predictive accuracy. Ensemble classification techniques are another opportunity to extend the study into broader spectrums.

By continuing to innovate and iterate, the advancement of dermatological image analysis and its applications in clinical practice can create a revolutionary change in the healthcare community.

Bibliography

[1] Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021, February 4). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. https://acsjournals.onlinelibrary.wiley.com/doi/10.3322/caac.21660

[2] Melanoma of the Skin — Cancer Stat Facts. (n.d.). SEER Cancer. https://seer.cancer.gov/statfacts/html/melan.html

[3] Skin Cancer Facts & Statistics. (n.d.). The Skin Cancer Foundation. https://www.skincancer.org/skin-cancer-information/skin-cancer-facts/

[4] Wu, L., & Qu, X. (2015). Cancer biomarker detection: recent achievements and challenges.. *Chemical Society reviews*, 44 10, 2963-97 . https://doi.org/10.1039/c4cs00370e.

[5] Patil, C., Kirshnamoorthi, H., Ellis, D., Leeuwen, T., & Mahadevan-Jansen, A. (2011). A clinical instrument for combined raman spectroscopy-optical coherence tomography of skin cancers. Lasers in Surgery and Medicine, 43. https://doi.org/10.1002/lsm.21041.

[6] Kanchanachitra, C., Lindelow, M., Johnston, T., Hanvoravongchai, P., Lorenzo, F., Huong, N., Wilopo, S., & Rosa, J. (2011). Human resources for health in southeast Asia: shortages, distributional challenges, and international trade in health services. *The Lancet*, 377, 769-781. https://doi.org/10.1016/S0140-6736(10)62035-1.

[7] Li, Y., & Shen, L. (2017). Skin Lesion Analysis towards Melanoma Detection Using Deep Learning Network. *Sensors (Basel, Switzerland)*, 18.

https://doi.org/10.3390/s18020556.

[8] Ahmed, S., Hossain, M., RajaChowdhury, A., & Bhuiya, A. (2011). The health workforce crisis in Bangladesh: shortage, inappropriate skill-mix and inequitable distribution. *Human Resources for Health*, 9, 3 - 3. https://doi.org/10.1186/1478-4491-9-3.

[9] Molla, A., & Chi, C. (2020). How Much Household Healthcare Expenditure Contributes to Poverty? Evidence from the Bangladesh Household Income and Expenditure Survey, 2010. Journal of Poverty, 24, 627 - 641. https://doi.org/10.1080/10875549.2020.1742269.

[10] Dargan, S., Kumar, M., Ayyagari, M., & Kumar, G. (2020). A Survey of Deep Learning and Its Applications: A New Paradigm to Machine Learning. Archives of Computational Methods in Engineering, 1-22. https://doi.org/10.1007/S11831-019-09344-W.

[11] George, Y., Zayed, H., Roushdy, M., & Elbagoury, B. (2014). Remote Computer-Aided Breast Cancer Detection and Diagnosis System Based on Cytological Images. IEEE Systems Journal, 8, 949-964. https://doi.org/10.1109/JSYST.2013.2279415.

[12] Tan, M. & Le, Q.. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. Proceedings of the 36th International Conference on Machine Learning, in Proceedings of Machine Learning Research 97:6105-6114 Available from https://proceedings.mlr.press/v97/tan19a.html.

[13] Erickson, B., Korfiatis, P., Akkus, Z., Kline, T., & Philbrick, K. (2017). Toolkits and Libraries for Deep Learning. Journal of Digital Imaging, 30, 400 - 405.

https://doi.org/10.1007/s10278-017-9965-6.

[14] Tschandl P., Rosendahl C. & Kittler H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Sci. Data 5, 180161 doi.10.1038/sdata.2018.161 (2018)

[15] Noel C. F. Codella, David Gutman, M. Emre Celebi, Brian Helba, Michael A. Marchetti, Stephen W. Dusza, Aadi Kalloo, Konstantinos Liopyris, Nabin Mishra, Harald Kittler, Allan Halpern: "Skin Lesion Analysis Toward Melanoma Detection: A Challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), Hosted by the International Skin Imaging Collaboration (ISIC)", 2017; arXiv:1710.05006.

[16] Marc Combalia, Noel C. F. Codella, Veronica Rotemberg, Brian Helba, Veronica Vilaplana, Ofer Reiter, Allan C. Halpern, Susana Puig, Josep Malvehy: "BCN20000: Dermoscopic Lesions in the Wild", 2019; arXiv:1908.02288.

[17] New Zealand Dermatological Society. DermNet NZ. https://dermnetnz.org/

[18] T. Mendonça, P. M. Ferreira, J. S. Marques, A. R. S. Marcal and J. Rozeira, "PH2 - A dermoscopic image database for research and benchmarking," 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Osaka, Japan, 2013, pp. 5437-5440, doi: 10.1109/EMBC.2013.6610779. keywords: {Lesions;Databases;Malignant tumors;Image segmentation;Skin;Image color analysis;Manuals},

[19] https://www.cancer.org/cancer/types/skin-cancer/skin-cancer-image-gallery.html

[20] LeCun, Yann & Bengio, Y. & Hinton, Geoffrey. (2015). Deep Learning. Nature. 521. 436-44. 10.1038/nature14539.

[21] Gordo, A., Almazán, J., Revaud, J., & Larlus, D. (2016). End-to-End Learning of Deep Visual Representations for Image Retrieval. International Journal of Computer Vision, 124, 237-254. https://doi.org/10.1007/s11263-017-1016-8.

[22] Mahbod, A., Schaefer, G., Wang, C., Ecker, R., Dorffner, G., & Ellinger, I. (2020). Investigating and Exploiting Image Resolution for Transfer Learning-based Skin Lesion Classification. 2020 25th International Conference on Pattern Recognition (ICPR), 4047-4053. https://doi.org/10.1109/ICPR48806.2021.9412307.

[23] Mahbod, A., Schaefer, G., Wang, C., Dorffner, G., Ecker, R., & Ellinger, I. (2020). Transfer learning using a multi-scale and multi-network ensemble for skin lesion classification. Computer methods and programs in biomedicine, 193, 105475 . https://doi.org/10.1016/j.cmpb.2020.105475.

[24] Majtner, T., Lidayová, K., YILDIRIM, S., & Hardeberg, J. (2016). Improving skin lesion segmentation in dermoscopic images by thin artefacts removal methods. 2016 6th European Workshop on Visual Information Processing (EUVIP), 1-6. https://doi.org/10.1109/EUVIP.2016.7764580.

[25] Emre Celebi M., Kingravi H.A., Iyatomi H., Alp Aslandogan Y., Stoecker W.V., Moss R.H., Malters J.M., Grichnik J.M., Marghoob A.A., Rabinovitz H.S. Border detection in dermoscopy images using statistical region merging. Ski. Res. Technol. 2008;14:347–353. doi: 10.1111/j.1600-0846.2008.00301.x.

[26] Nachbar, F., Stolz, W., Merkle, T., Cognetta, A., Vogt, T., Landthaler, M., Bílek, P., Braun-falco, O., & Plewig, G. (1994). The ABCD rule of dermatoscopy. High prospective value in the diagnosis of doubtful melanocytic skin lesions.. Journal of the American Academy of Dermatology, 30 4, 551-9 . https://doi.org/10.1016/S0190-9622(94)70061-3.

[27] Guarracino M.R., Maddalena L. SDI+: A novel algorithm for segmenting dermoscopic images. IEEE J. Biomed. Health Inform. 2019;23:481–488. doi: 10.1109/JBHI.2018.2808970.

[28] Jahanifar M., Tajeddin N.Z., Gooya A., Asl B.M. Segmentation of lesions in dermoscopy images using saliency map and contour propagation. arXiv. 20171703.00087v2

[29] Pennisi A., Bloisi D.D., Nardi D., Giampetruzzi A.R., Mondino C., Facchiano A. Skin lesion image segmentation using Delaunay Triangulation for melanoma detection. Comput. Med. Imaging Graph. 2016;52:89–103. doi: 10.1016/j.compmedimag.2016.05.002.

[30] Nida N., Irtaza A., Javed A., Yousaf M.H., Mahmood M.T. Melanoma lesion detection and segmentation using deep region based convolutional neural network and fuzzy C-means clustering. Int. J. Med. Inform. 2019;124:37–48. doi: 10.1016/j.ijmedinf.2019.01.005.

[31] Khan, A., Iskandar, D., Al-Asad, J., & El-Nakla, S. (2021). Classification of Skin Lesion with Hair and Artifacts Removal using Black-hat Morphology and Total Variation. International Journal of Computing and Digital Systems. https://doi.org/10.12785/IJCDS/100157.

[32] Guillemot, C., & Meur, O. (2014). Image Inpainting : Overview and Recent Advances. IEEE Signal Processing Magazine, 31, 127-144. https://doi.org/10.1109/MSP.2013.2273004.

[33] Voronin, V., Marchuk, V., Sherstobitov, A., & Egiazarian, K. (2012). Image inpainting using cubic spline-based edge reconstruction. , 8295. https://doi.org/10.1117/12.905791.

[34] Lee, T., Ng, V., Gallagher, R., Coldman, A., & McLean, D. (1997). Dullrazor®: A software approach to hair removal from images. Computers in biology and medicine, 27 6, 533-43 . https://doi.org/10.1016/S0010-4825(97)00020-6.

[35] Ünver H.M., Ayan E. Skin lesion segmentation in dermoscopic images with combination of YOLO and grabcut algorithm. Diagnostics. 2019;9:72. doi: 10.3390/diagnostics9030072.

[36] Garnavi R., Aldeen M., Celebi M.E., Varigos G., Finch S. Border detection in dermoscopy images using hybrid thresholding on optimized color channels. Comput. Med. Imaging Graph. 2011;35:105–115. doi: 10.1016/j.compmedimag.2010.08.001

[37] Javed R., Rahim M.S.M., Saba T., Rashid M. Region-based active contour JSEG fusion technique for skin lesion segmentation from dermoscopic images. Biomed. Res. 2019;30:1–10.

[38] Okuboyejo D.A., Olugbara O.O., Odunaike S.A. CLAHE inspired segmentation of dermoscopic images using mixture of methods. In: Ao H.K.K.S.-I., Amouzegar M.A., editors. Transactions on Engineering Technologies. Springer; Berlin/Heidelberg, Germany: 2014. pp. 355–365

[39] Ibraheem M.R., Elmogy M. A non-invasive automatic skin cancer detection system for characterizing malignant melanoma from seborrheic keratosis; Proceedings of the 2nd International Conference on Computer and Information Sciences (ICCIS); Manta, Ecuador. 27–29 July 2020; pp. 1–5.

[40] Hoshyar A.N., Jumaily A.A., Hoshyar A.N. Pre-processing of automatic skin cancer detection system: Comparative study. Int. J. Smart Sens. Intell. Syst. 2014;7:1364–1377. doi: 10.21307/ijssis-2017-710.

[41] Rajput A.S., Tanwar V.K., Raman B. Score-based secure biomedical model for effective skin lesion segmentation over eHealth cloud. ACM Trans. Multimed. Comput. Commun. Appl. (TOMM) 2021;17:1–19. doi: 10.1145/3430806.

[42] Perez, F., Vasconcelos, C., Avila, S., & Valle, E. (2018). Data Augmentation for Skin Lesion Analysis. , 303-311. https://doi.org/10.1007/978-3-030-01201-4_33.

[43] Pham, T., Luong, C., Visani, M., & Hoang, V. (2018). Deep CNN and Data Augmentation for Skin Lesion Classification. , 573-582. https://doi.org/10.1007/978-3-319-75420-8_54.

[44] Iqbal Hussain, M.A.; Khan, B.; Wang, Z.; Ding, S. Woven Fabric Pattern Recognition and Classification Based on Deep Convolutional Neural Networks. Electronics 2020, 9, 1048.

[45] Jeon, H.K.; Kim, S.; Edwin, J.; Yang, C.S. Sea Fog Identification From GOCI Images Using CNN Transfer Learning Models. Electronics 2020, 9, 311.

[46] Menegola, A., Fornaciali, M., Pires, R., Bittencourt, F., Avila, S., & Valle, E. (2017). Knowledge transfer for melanoma screening with deep learning. 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), 297-300. https://doi.org/10.1109/ISBI.2017.7950523.

[47] Lopez, A., Giró-i-Nieto, X., Burdick, J., & Marques, O. (2017). Skin lesion classification from dermoscopic images using deep learning techniques. 2017 13th

IASTED International Conference on Biomedical Engineering (BioMed), 49-54. https://doi.org/10.2316/P.2017.852-053.

[48] Georgakopoulos, S., Kottari, K., Delibasis, K., Plagianakos, V., & Maglogiannis, I. (2017). Detection of Malignant Melanomas in Dermoscopic Images Using Convolutional Neural Network with Transfer Learning. , 404-414. https://doi.org/10.1007/978-3-319-65172-9_34.

[49] Valle, E., Fornaciali, M., Menegola, A., Tavares, J., Bittencourt, F., Li, L., & Avila, S. (2017). Data, depth, and design: Learning reliable models for skin lesion analysis. Neurocomputing, 383, 303-313. https://doi.org/10.1016/j.neucom.2019.12.003.

[50] Phornchaicharoen, A., & Padungweang, P. (2019). Face Recognition using Transferred Deep Learning for Feature Extraction. 2019 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT-NCON), 304-309. https://doi.org/10.1109/ECTI-NCON.2019.8692306.

[51] Agarwal, S., Hallac, R., Mishra, R., Li, C., Daescu, O., & Kane, A. (2018). Image Based Detection of Craniofacial Abnormalities using Feature Extraction by Classical Convolutional Neural Network. 2018 IEEE 8th International Conference on Computational Advances in Bio and Medical Sciences (ICCABS), 1-6. https://doi.org/10.1109/ICCABS.2018.8541948.

[52] Almaraz-Damian, J., Ponomaryov, V., Sadovnychiy, S., & Castillejos-Fernandez, H. (2020). Melanoma and Nevus Skin Lesion Classification Using Handcraft and Deep Learning Feature Fusion via Mutual Information Measures. Entropy, 22. https://doi.org/10.3390/e22040484.

[53] Wang, L., Guo, S., Huang, W., & Qiao, Y. (2015). Places205-VGGNet Models for Scene Recognition. ArXiv, abs/1508.01667.

[54] Zhou, T., Ye, X., Lu, H., Zheng, X., Qiu, S., & Liu, Y. (2022). Dense Convolutional Network and Its Application in Medical Image Analysis. BioMed Research International, 2022. https://doi.org/10.1155/2022/2384830.

[55] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in Proc.PMLR, California, USA, pp. 6105–6114, 2019.

[56] Baheti, Bhakti & Innani, Shubham & Gajre, Suhas & Talbar, Sanjay. (2020). Eff-UNet: A Novel Architecture for Semantic Segmentation in Unstructured Environment. 1473-1481. 10.1109/CVPRW50498.2020.00187.

[57] Srinivasu, P., SivaSai, J., Ijaz, M., Bhoi, A., Kim, W., & Kang, J. (2021). Classification of Skin Disease Using Deep Learning Neural Networks with MobileNet V2 and LSTM. Sensors (Basel, Switzerland), 21. https://doi.org/10.3390/s21082852.

[58] Yang, C., Ren, J., Huang, H., Chuang, L., & Chang, P. (2021). Deep Hybrid Convolutional Neural Network for Segmentation of Melanoma Skin Lesion. Computational Intelligence and Neuroscience, 2021. https://doi.org/10.1155/2021/9409508.

[59] Miglani, V., & Bhatia, M. (2020). Skin Lesion Classification: A Transfer Learning Approach Using EfficientNets. , 315-324. https://doi.org/10.1007/978-981-15-3383-9_29.

[60] Hridoy, R., Akter, F., & Rakshit, A. (2021). Computer Vision Based Skin Disorder Recognition using EfficientNet: A Transfer Learning Approach. 2021 International Conference on Information Technology (ICIT), 482-487. https://doi.org/10.1109/ICIT52682.2021.9491776.

[61] Shetty, B., Fernandes, R., Rodrigues, A.P. et al. Skin lesion classification of dermoscopic images using machine learning and convolutional neural network. Sci Rep 12, 18134 (2022). https://doi.org/10.1038/s41598-022-22644-9

[62] Alam TM, Shaukat K, Khan WA, Hameed IA, Almuqren LA, Raza MA, Aslam M, Luo S. An Efficient Deep Learning-Based Skin Cancer Classifier for an Imbalanced Dataset. Diagnostics. 2022; 12(9):2115. https://doi.org/10.3390/diagnostics12092115

[63] Pollastri, F., Bolelli, F., Paredes, R., & Grana, C. (2019). Augmenting data with GANs to segment melanoma skin lesions. Multimedia Tools and Applications, 79, 15575 - 15592. https://doi.org/10.1007/s11042-019-7717-y.

[64] Qin, Z., Liu, Z., Zhu, P., & Xue, Y. (2020). A GAN-based image synthesis method for skin lesion classification. Computer methods and programs in biomedicine, 195, 105568 . https://doi.org/10.1016/j.cmpb.2020.105568.

[65] Frid-Adar, M., Klang, E., Amitai, M., Goldberger, J., & Greenspan, H. (2018). Synthetic data augmentation using GAN for improved liver lesion classification. 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), 289-293. https://doi.org/10.1109/ISBI.2018.8363576.

[66] Perez, F., Vasconcelos, C., Avila, S., & Valle, E. (2018). Data Augmentation for Skin Lesion Analysis. , 303-311. https://doi.org/10.1007/978-3-030-01201-4_33.

[67] Ostapenko, O., Lesort, T., Rodr'iguez, P., Arefin, M., Douillard, A., Rish, I., & Charlin, L. (2022). Foundational Models for Continual Learning: An Empirical Study of Latent Replay. ArXiv, abs/2205.00329. https://doi.org/10.48550/arXiv.2205.00329.

[68] Devi, N., & Borah, B. (2019). An Approach of Transferring Pre-trained Deep Convolutional Neural Networks for Aerial Scene Classification. , 551-558. https://doi.org/10.1007/978-3-030-34869-4_60.

[69] Lemaître, G., Nogueira, F., & Aridas, C. K. (2017). Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning. Journal of Machine Learning Research, 18(17), 1-5. Retrieved from http://jmlr.org/papers/v18/16-365.html

[70] Abayomi-Alli, O., Damaševičius, R., Misra, S., Maskeliūnas, R., & Abayomi-Alli, A. (2021). Malignant skin melanoma detection using image augmentation by oversampling in nonlinear lower-dimensional embedding manifold. Turkish J. Electr. Eng. Comput. Sci., 29, 2600-2614. https://doi.org/10.3906/elk-2101-133.

[71] Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., & Inman, D. J. (2021). 1D convolutional neural networks and applications: A survey. Mechanical systems and signal processing, 151, 107398.

[72] Ince, T., Kiranyaz, S., Eren, L., Askar, M., & Gabbouj, M. (2016). Real-Time Motor Fault Detection by 1-D Convolutional Neural Networks. IEEE Transactions on Industrial Electronics, 63, 7067-7075. https://doi.org/10.1109/TIE.2016.2582729.

[73] Kiranyaz, S., Ince, T., & Gabbouj, M. (2015). Real-time patient-specific ECG classification by 1-D convolutional neural networks. IEEE Transactions on Biomedical Engineering, 63(3), 664-675.

[74] Eren, L. (2017). Bearing fault detection by one-dimensional convolutional neural networks. Mathematical Problems in Engineering, 2017, 1-9.

[75] Eren, L., Ince, T., & Kiranyaz, S. (2019). A generic intelligent bearing fault diagnosis system using compact adaptive 1D CNN classifier. Journal of Signal Processing Systems, 91, 179-189.

[76] Groesser, L., Herschberger, E., Landthaler, M., & Hafner, C. (2012). FGFR3, PIK3CA and RAS mutations in benign lichenoid keratosis. British Journal of Dermatology, 166. https://doi.org/10.1111/j.1365-2133.2011.10788.x.

[77] Farnetani, F., Ciardo, S., & Pellacani, G. (2020). In Vivo Reflectance Confocal Microscopy for Nonmelanocytic Benign Skin Tumors. , 157-161. https://doi.org/10.1007/978-3-030-45351-0_15.

[78] https://keras.io/api/applications/

[79] Ying, X. (2019). An Overview of Overfitting and its Solutions. Journal of Physics: Conference Series, 1168. https://doi.org/10.1088/1742-6596/1168/2/022022.

[80] Lin, M., Chen, Q., & Yan, S. (2013). Network in network. arXiv preprint arXiv:1312.4400.

[81] Siddiq, S & Shailendra, Roopashree & Suha, Musfirah & Ruthvik, MRS & Divyasree, Kuruva. (2021). SIGNify – A mobile solution for Indian sign language using MobileNet architecture. 1-9. 10.1109/GCAT52182.2021.9587639.

[82] Zhou, B., Khosla, A., Lapedriza, À., Oliva, A., & Torralba, A. (2015). Learning Deep Features for Discriminative Localization. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2921-2929. https://doi.org/10.1109/CVPR.2016.319.

[83] Wang, G., Giannakis, G., & Chen, J. (2018). Learning ReLU Networks on Linearly Separable Data: Algorithm, Optimality, and Generalization. IEEE Transactions on Signal Processing, 67, 2357-2370. https://doi.org/10.1109/TSP.2019.2904921.

[84]https://www.learndatasci.com/tutorials/convolutional-neural-networks-image-classification/

[85] Yamashita, R., Nishio, M., Do, R., & Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. Insights into Imaging, 9, 611 - 629. https://doi.org/10.1007/s13244-018-0639-9.

[86] Buda, M., Maki, A., & Mazurowski, M. (2017). A systematic study of the class imbalance problem in convolutional neural networks. Neural networks : the official journal of the International Neural Network Society, 106, 249-259 . https://doi.org/10.1016/j.neunet.2018.07.011.

[87] Mukherjee, Uddipan & Pancholi, Sidharth. (2021). A Visual Domain Transfer Learning Approach for Heartbeat Sound Classification.