People's Democratic Republic of Algeria Ministry of Higher Education and Scientific Research

University of KASDI MERBAH OUARGLA

Faculty of New Technologies of Information and Communication
Department of Electronic and Telecommunications

ACADEMIC MASTER'S THESIS

In: Electronic

Specialty: Electronic of Embedded Systems

By: BEKAKRA Ahmed Ramzi
LAMNIAI Mohammed Ramzi

Theme

Deep Learning Model for Skin Disease Classification

Submit Date: Tuesday 27 May 2025

Before the jury

Dr. BENSID Khaled	President	UKM Ouargla
Dr. YOUCEFA Abdelmadjid	Supervisor	UKM Ouargla
Dr. NASRI Nadjib	Examiner	UKM Ouargla

Academic Year: 2024/2025

Acknowledgement

First of all, We thank ALLAH from the bottom of our heart for having enlightened us on the right path. We would like to extend our heartfelt appreciation to our families for their unwavering support and sacrifices. Words cannot express our gratitude to our supervisor, **Prf. YOUCEFA Abdelmadjid** , for his constant availability, patience, and guide us throughout this work. His invaluable contributions have been instrumental in our success. We extend our sincere thanks to the members of the Jury for accepting the responsibility of evaluating our work : **Dr. BENSID Khaled** President. **Dr. NASRI Nadjib** Examiner. We also express our sincere thanks to all our teachers from the Department of Electronics and Telecommunications at the

University of KASDI MERBAH in OUARGLA **UKMO** for the effort they have made to ensure our education, for their skills, and especially their humility. We would also like to thank our families, our friends, as well as all those who have contributed directly or indirectly to the realization of this work.

Abstract

Skin diseases are very common health disorders, from common disorders to life-threatening cancers like melanoma. Their early and correct diagnosis is very important but typically challenging owing to similar appearances among diseases, different skin color, and a lack of experts. This thesis proposes a deep learning solution for the automatic diagnosis of skin diseases using dermatoscopic images, relying on CNNs and transfer learning. ResNet50, EfficientNetV2-B1, and Xception were trained on an improved version of the HAM10000 dataset. The results showed that Xception was the best performing at 98% accuracy, followed by EfficientNetV2-B1 at 96%, and then ResNet50 at 93%. The results show that CNN models, specifically Xception, can help dermatologists with quick and accurate diagnosis.

Keywords: Skin diseases, Deep learning, Convolutional Neural Network (CNN), Xception, EffientNetV2-B1,ResNet-50.

ملخص

أمراض الجلد اضطرابات صحية شائعة جدًا، بدءًا من الاضطرابات الشائعة وصولًا إلى أنواع السرطان المهددة للحياة مثل الورم الميلانيني. يُعد تشخيصها المبكر والدقيق أمرًا بالغ الأهمية، ولكنه عادةً ما يكون صعبًا نظرًا لتشابه أعراض الأمراض واختلاف لون البشرة ونقص الخبراء. تقترح هذه الأطروحة حلاً قائمًا على التعلم العميق للتشخيص التلقائي لأمراض الجلد باستخدام صور تنظير الجلد، بالاعتماد على الشبكات العصبية التلافيفية (CNN) والتعلم بالنقل. تم تدريب ResNet50 وEfficientNetV2-B1 وXception على نسخة مُحسّنة من مجموعة بيانات HAM10000. أظهرت النتائج أن Xception كان الأفضل أداءً بدقة 98%، يليه EfficientNetV2-B1 بنسبة 96%، ثم ResNet50 بنسبة 93%. تُظهر النتائج أن نماذج الشبكات العصبية التلافيفية (CNN)، وتحديدًا Xception، يُمكن أن تُساعد أطباء الجلد في التشخيص السريع والدقيق.

الكلمات المفتاحية: أمراض الجلد، التعلم العميق، الشبكة العصبية التلافيفية،Xception ، ResNet-50،EffientNetV2-B1.

Contents

List Of Figures

List Of Tables

Abbreviation

AI Artificial intelligence.

AK Actinic Keratoses.

BCC Basal Cell Carcinoma.

BKL Benign Keratosis-like Lesions.

CNN Convolutional Neural Network.

CNS central nervous system.

Conv Convolution.

DALYS Disability-Adjusted Life Years.

DF Dermatofibroma.

DL Deep Learning.

FN False Negative.

FP False Positive.

HAM10000 Human Against Machine With 10000 Training Images.

MEL Melanoma.

ML Machine Learning.

MV Melanotic nevi.

NN Neural Network.

TN True Negative.

TP True Positive.

VASC Vascular Lesions.

CHAPTER I
General Introduction

1.1 Related Work

Skin diseases represent a significant global health challenge, affecting millions of individuals across diverse demographics and geographic regions. According to the World Health Organization (WHO), skin conditions rank among the most common human ailments, with nearly 30% of the global population experiencing at least one dermatological disorder during their lifetime [11]. According to the Global Burden of Disease Study, skin and subcutaneous diseases accounted for approximately 42.9 million disability adjusted life years (DALYs) in 2019, underscoring their substantial impact on global health.

Skin diseases can arise from various causes such as poor diet, environmental changes, weakened immune systems, and UV exposure. Some conditions, if ignored, can worsen or even develop into skin cancer especially in older individuals or those with a family history.

The burden of these diseases extends beyond physical discomfort, often leading to psychological distress, social stigma, and economic hardship due to treatment costs and productivity loss. Among the most prevalent conditions are acne vulgaris (affecting 9.4% of the global population), atopic dermatitis (up to 20% in children), and life-threatening malignancies such as melanoma, whose incidence has surged by 40% over the past decade [12]. These statistics underscore the urgent need for accessible, accurate, and timely diagnostic solutions to mitigate morbidity and mortality. The main protective barrier for the human body which is the skin protects internal organs from external threats. With millions suffering from different dermatological disorders every year all around, the incidence of skin diseases is surprisingly high. Ranging from typical acne to possibly life-threatening melanoma, which has a death rate higher than many other cancer kinds, these diseases differ greatly in presentation and severity [2].

Traditional diagnosis techniques depend mostly on visual inspection by dermatologists, which demands great clinical knowledge and experience to distinguish between visually similar diseases. Among its limitations are subjective evaluation, limited access to specialisttreatment in many areas, and the possibility of diagnostic mistakes resulting from the visual similarity between various dermatological diseases. Many skin diseases have similar visual features, which makes exact differentiation difficult even for experienced doctors and so adds to the diagnostic difficulty [3].

In recent years, rapid advancements in technology particularly in Artificial Intelligence (AI) and deep learning have brought transformative changes to numerous domains, including healthcare. Among these innovations, Computer Aided Diagnosis (CAD) systems powered by AI have emerged as valuable tools for improving diagnostic accuracy, consistency, and efficiency. These systems aim to support, rather than replace, clinicians by providing a second opinion, reducing human error, and improving outcomes, especially in resource-limited settings.

The integration of AI into dermatology is part of a broader trend where machine learning algorithms are reshaping clinical workflows. In particular, deep learning a subset of machine learning inspired by the structure and function of the human brain has shown exceptional promise in medical image analysis. One of the most powerful and widely used deep learning architectures in this context is the Convolutional Neural Network (CNN), they are specifically designed to process and interpret visual data, making them ideally suited for analyzing skin lesion images and distinguishing between benign and malignant conditions with high accuracy.

Studies have demonstrated that CNN based systems can perform on par with, or even surpass, experienced dermatologists in specific tasks such as melanoma classification or psoriasis detection. These systems learn from vast datasets of labeled dermatological images, enabling them to recognize subtle patterns that might be missed during traditional visual inspection. Moreover, CNNs enable the automation of diagnostic workflows, which could enhance accessibility to dermatological care in underserved regions and support early detection crucial for reducing morbidity and mortality associated with skin diseases [13].

The study by Spyridonos et al. [28] addressed the diagnostic challenge of distinguishing malignant melanomas (MM) clinically resembling seborrheic keratosis (SK-like MMs) using deep learning. They leveraged dermoscopic images from the ISIC archive (dermoscopic images Malignant Melanoma: 550 images Seborrheic Keratosis: 428 images)

Image Resolution: Ranged from 639×602 to 6720×4461 pixels. Additional Data: 200 images from BCN_20000 dataset (Hospital Clinic de Barcelona) Some images manually annotated by experts, employing transfer learning with pretrained CNNs (VGG16 and ResNet50) to extract feature embeddings, which were then classified using SVM optimized via Bayesian hyperparameter tuning. Their approach achieved 78.6% sensitivity and 84.5% specificity, highlighting the efficacy of mid-level convolutional layers for this complex task. The work underscores the potential of AI to aid in diagnosing atypical skin lesions where traditional methods often fail, while acknowledging limitations in interpretability and the need for diverse datasets.

Velasco et al. [29] developed a smartphone-based skin disease classification system using MobileNet CNN, targeting seven prevalent skin diseases (acne, eczema, psoriasis). They curated an imbalanced dataset of 3,406 images and explored sampling techniques (undersampling, oversampling) and data augmentation to optimize performance. The highest accuracy (94.4%) was achieved with oversampling and augmentation, while default preprocessing yielded 93.6%. The model was deployed as an android app, leveraging transfer learning and fine-tuning (SoftMax activation, Adam optimizer). Challenges included misclassification of psoriasis as acne, addressed via saliency maps for interpretability. The study highlights the potential of lightweight CNNs like MobileNet for mobile health applications in resource-limited settings.

Pasan Maduranga and Dilshan Nandasena. [30] utilized the HAM10000 dataset comprising 10,015 images across seven skin lesion categories (melanoma, basal cell carcinoma, etc.), split into 60% training, 20% validation, and 20% test sets, with augmentation techniques like rotation and flipping to enhance performance. After reviewing traditional methods (SVM, KNN, ANNs) and CNN architectures (Inception v4, ResNet), the authors implemented MobileNet via transfer learning for its computational efficiency and mobile compatibility, achieving 85% accuracy and an 86% weighted F1-score. While outperforming other models in efficiency, limitations like lighting sensitivity were noted. Comparative analysis of ~25,000 images across eight conditions revealed ResNet152 as the top performer among CNNs, demonstrating the potential of deep learning for scalable skin disease classification.

For Raghav Agarwal. [31] They present a comprehensive deep learning framework for automated classification of eight common skin diseases using convolutional neural networks

(CNNs) and transfer learning techniques. The researchers compiled a substantial dataset of approximately 25,000 high-quality color images representing dermatofibroma, melanocytic nevus, melanoma, squamous cell carcinoma, actinic keratosis, basal cell carcinoma, benign keratosis, and vascular lesions, sourced from multiple Kaggle repositories. The dataset was strategically partitioned into training (80%), validation (10%), and testing (10%) sets to ensure robust model evaluation. Building upon previous research that compared traditional machine learning approaches with deep learning methods, the authors conducted extensive experiments with several state-of-the-art CNN architectures including VGG16/VGG19, multiple ResNet variants (50, 101, 152), InceptionV3, DenseNet models (121, 169, 201), MobileNet, and Xception. All models were initialized with ImageNet weights and fine-tuned using consistent parameters: input images resized to 224×224 pixels, Adam optimizer with learning rate of 0.001, batch size of 32, and training over 50 epochs. Among all architectures evaluated, ResNet152 demonstrated superior performance, achieving 74.24% validation accuracy and 73.01% test accuracy, along with strong precision (~75.45% validation, ~75.30% test) and recall (~73.04% validation, ~71.71% test) metrics. The study's findings not only confirm the effectiveness of transfer learning for medical image analysis but also provide valuable insights into model selection for skin disease classification tasks, with ResNet152 emerging as the optimal architecture balancing accuracy and computational efficiency for this application.

1.2 Problem Statement

Despite the significant global burden of skin diseases particularly life-threatening conditions like melanoma early and accurate diagnosis remains a critical challenge in dermatology.

Traditional diagnostic approaches primarily rely on visual examination by dermatologists, which introduces several limitations:

- **Subjectivity and Variability:** Human mistake can occur during visual diagnosis, particularly when differentiating between lesions that appear visually similar (malignant melanomas versus benign moles). Variability between and between observers may result in incorrect diagnoses, delaying life-saving interventions.

- **Limited Access to Specialists:** Many regions, particularly rural and low- resource areas, face a severe shortage of dermatologists, restricting timely diagnosis and increasing morbidity rates.

- **Complexity in Lesion Analysis:** Dermatological conditions demonstrate significant inter-class similarities (psoriasis versus eczema) and intra-class diversity (melanoma presenting differently across various skin tones), resulting in complicating accurate diagnosis.

- **Rising Skin Cancer Incidence:** With melanoma incidence increasing by 40% in the past decade, there is an urgent need for scalable, cost-effective tools to support early detection and reduce mortality.

While advances in Artificial Intelligence (AI), particularly Convolutional Neural Networks (CNNs), offer promising solutions for automated skin lesion analysis, existing systems still face challenges:

- **Dataset Limitations:** Many AI models are trained on non-diverse datasets, which results in biased performance across various skin tones and lesion kinds.

- **Real-World Applicability:** Most research focuses on binary classification (benign vs malignant) rather than multi-class detection of rare or complex malignancies.

This thesis aims to address these gaps by developing a robust deep learning (CNN-based) framework for the automated classification and detection of malignant skin diseases, leveraging diverse datasets and optimizing for clinical interpretability. The proposed system seeks to enhance diagnostic accuracy, reduce dependency on specialist availability, and improve early detection rates ultimately contributing to better patient outcomes in dermatology.

1.3 Contribution

In this work, we aim to achieve the following:

- Improve the accuracy of skin disease classification compared to previous studies.
- Ensure transparency by clearly explaining the key steps and decisions made during the study.
- Use different algorithms to process the dataset, compare their performance, and select the most effective one.
- Develop a Convolutional Neural Network (CNN) model with an accuracy of 90% or higher for detecting skin diseases.

1.4 Solution

To overcome the issue of low accuracy caused by manual feature extraction, this work uses a deep Convolutional Neural Network (CNN) to automatically detect and classify skin lesions. CNNs are powerful because they can learn and recognize patterns—like shapes, colors, and textures directly from the raw images without needing hand-crafted features.

By analyzing how pixel values are structured in a photo, CNNs can isolate important visual features that help in accurately identifying different skin conditions. The performance of these networks often depends on the type of computer vision task they're used for.

In our approach, we experiment with different pre-trained CNN models, such as ResNet-50, EfficientNetV2B1, and Xception, and compare their results. Among these, ResNet-50 is highlighted in prior studies for its strong performance in medical image classification, despite having a deeper and more complex structure than models like VGG-19 [27].

Our goal is to apply and fine-tune these models to achieve reliable and accurate classification of skin diseases, making diagnosis faster and more accessible.

1.5 Thesis structure

This thesis is organized as follows:

First of all, we present a general introduction about skin diseases.

- **Chapter One:** provides a general overview of skin diseases, including their types, diagnostic methods, the role of skin tone, and skin cancer facts and statistics.
- **Chapter Two:** explains the methodology and introduces the core technologies used in the study, such as deep learning, transfer learning, and the CNN architecture.
- **Chapter Three:** presents the experimental setup, compares different pre- trained CNN models, and selects the best-performing one for our task.

Finally, the **Conclusion** summarizes the key findings and discusses possible directions for future work.

CHAPTER II
Introduction to Skin Diseases Classification.

2.1 Introduction

The most common cancer in the world is skin cancer and, in many nations, both its incidence and treatment costs are continuously increasing. In this chapter, we'll start by giving an overview of the different types of skin diseases, along with some important facts and statistics about skin cancer. After that, we'll talk about the main problem this thesis is trying to solve and look at some of the work that's already been done by other researchers in this area. Then, we'll introduce our own proposed solution and explain what this thesis aims to contribute. Finally, we will conclude the chapter by providing a brief overview of the structure of the remaining thesis.

2.2 Skin diseases lesions

2.2.1 Melanocytic Nevi

Composed of melanocytes, the pigment producing cells in the skin, melanocytic nevi are benign tumors or hamartomas. Often called moles, they can show up anywhere on the body. Generally benign, these lesions may change or spread fast following skin damage including burns or severe sunburns. About 30% of melanomas start from pre-existing moles, thus monitoring them is crucial [4].

Figure 1:Melanotic nevi

2.2.2 Melanoma

Melanoma is the most aggressive and invasive type of skin cancer, originating from melanocytes. It can be treated very easily if detected early, but if left untreated, it can spread extremely fast to other organs. Melanomas are usually black or brown but can occasionally be pink, red, purple, or skin colored. Risk factors include fair skin, blonde or red hair, blue eyes, and excessive sun exposure [5].

This illustration shows melanoma cells extending from the surface of the skin into the deeper skin layers.

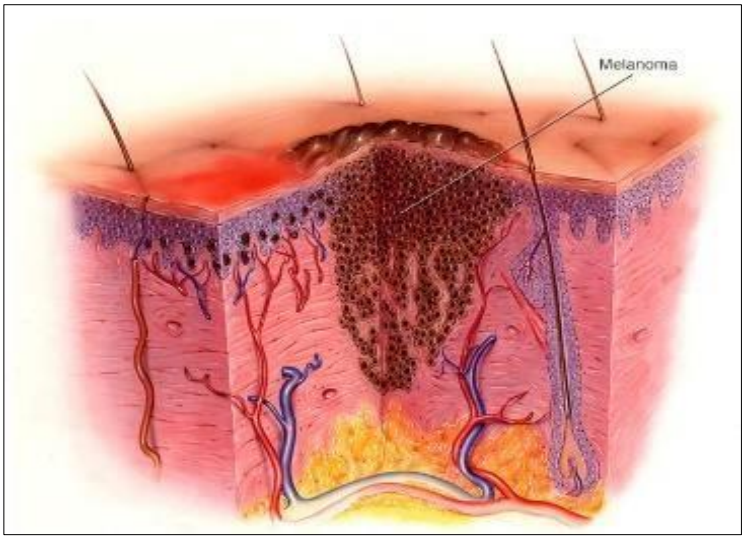

Figure 2:Melanoma

2.2.3 Benign Keratosis-like Lesions

These include seborrheic keratosis, lichen planus-like keratosis, and solar lentigo, which are benign cutaneous lesions frequently occurring in very close proximity to each other or transforming from one pattern to another. They have overlapping clinical and dermoscopic characteristics, and sometimes it may be challenging to differentiate among them. "Benign keratosis" is a very good term to treat combined or undifferentiated lesions in this entity [6].

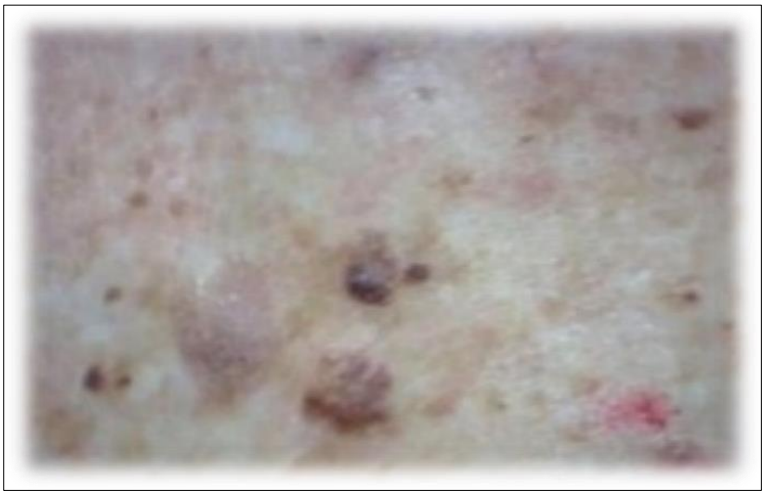

Figure 3:Bengin keratosis Lesions

2.2.4 Basal Cell Carcinoma (BCC)

Basal cell carcinoma is a skin cancer of basal keratinocytes. It is the most common skin cancer and becomes more common with advancing age. BCC is typically slow growing and has a low metastatic potential but can lead to extensive local tissue destruction. Risk factors are chronic sun exposure, fair skin, immunosuppression, and exposure to carcinogens such as arsenic [7]. Treatment is generally by surgical excision, with newer therapies like Hedgehog pathway inhibitors for metastatic disease. BCC is classified into low risk and high-risk types based on recurrence risk.

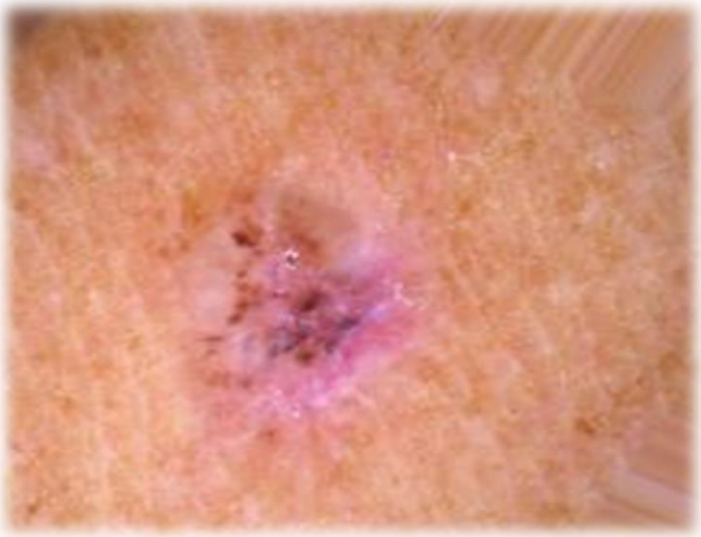

Figure 4:Basal Cell Carcinoma.

2.2.5 Actinic keratosis

Actinic keratoses are ultraviolet radiation-induced dysplastic proliferations of keratinocytes. They present as scaly, rough macules or plaques in sun exposed skin and have a potential to progress to squamous cell carcinoma. They are largely diagnosed clinically and, histopathology being left for uncertain cases, actinic keratoses are pre-cancerous lesions and need to be followed and treated to prevent malignant transformation [8].

Figure 5:Actinic Keratosis

2.2.6 Vascular Lesions

Vascular lesions represent a diverse collection of cutaneous neoplasms with abnormal proliferation or vascular malformation. Various vessel patterns such as arborizing, dotted, linear-irregular, and polymorphous vessels have been identified by dermoscopic studies that facilitate the differentiation of melanocytic and nonmelanocytic neoplasms. The lesions derive their clinical importance in terms of their role in differential diagnosis, spanning from benign hemangiomas to angiosarcomas [9].

Figure 6:Vascular Lesions

2.2.7 Dermatofibroma

A common benign fibrous skin tumor, dermatofibroma can develop on its own or because of trauma like insect bites. Usually on the extremities, it manifests as firm, elevated nodules. Multiple eruptive dermatofibromas are benign, but they may be linked to systemic conditions or immune dysfunction. The diagnosis is confirmed by histopathology, and although these lesions are usually benign, they might need to be removed if they are symptomatic or for aesthetic purposes [10].

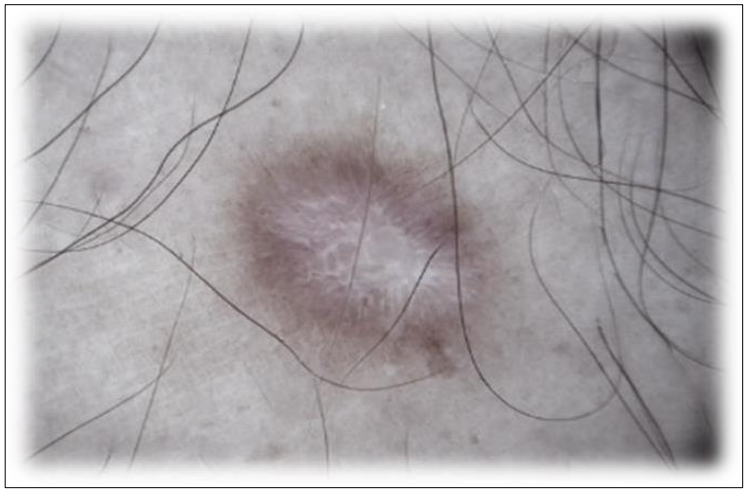

Figure 7:Dermatofibroma

In summary

Table 1:Skin lesions classification Table

Lesion Type	Nature
Melanocytic Nevi	Benign (can turn malignant)
Melanoma	Malignant
Benign Keratosis-like Lesions	Benign
Basal Cell Carcinoma (BCC)	Malignant
Actinic Keratoses	Benign with malignant potential
Vascular Lesions	Varies (Benign or Malignant)
Dermatofibroma	Benign

2.3 Sampling Techniques

2.3.1 Skin biopsy

The process of taking skin samples or cells out of your body for laboratory testing is called a skin biopsy. A skin biopsy is frequently used to identify or rule out specific illnesses and skin disorders.

When a solid diagnosis is needed or dermoscopy results are unclear, this method is employed [14].

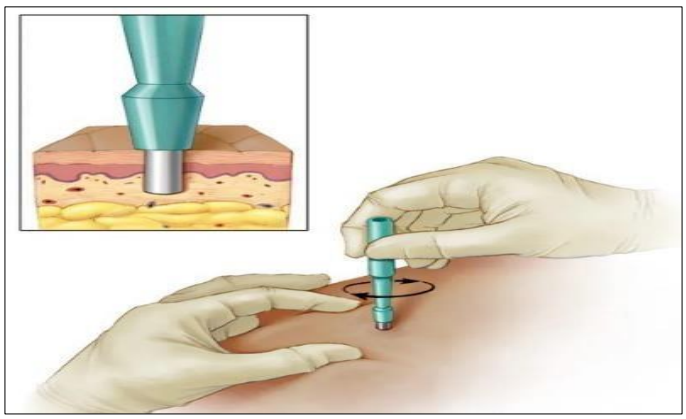

Figure 8:skin biopsy Technique

2.3.2 Dermoscopy

Dermoscopy is a non-invasive method of skin examination that makes use of a portable instrument called a dermatoscop, it is sometimes referred to as dermatoscopy or epiluminescence microscopy, this device's enhanced magnifying lens and integrated lighting enable medical professionals to see the skin's deeper features [16].

2.4 Relation between skin tone and skin disease

The color of our skin really affects how it responds to sunlight and can also influence the appearance or diagnosis of certain skin conditions. Individuals with darker skin, which is prevalent in various African regions, possess higher levels of melanin, the pigment responsible for skin color. Melanin does more than just affect how we look it also serves as a natural defense against the harmful rays of the sun.

It helps keep the skin safe from sunburn, early ageing, and even certain kinds of skin cancer [17].

Melanin provides natural protection by absorbing and scattering ultraviolet (UV) rays, and it also benefits the body by functioning as an antioxidant. With darker skin, it means that less UV radiation penetrates. For instance, research indicates that black skin permits significantly less UV radiation to penetrate compared to lighter skin, providing about 500 to 1000 times more protection against specific skin cancers [18].

It is interesting to note that lighter skin developed in regions with lower UV levels, such as those nearer to the poles, where the body had to produce more vitamin D due to the limited sunlight available. On the flip side, darker skin is more adapted to tropical areas where there is a lot of UV exposure [17].

Even though individuals with darker skin tones are at a lower risk for sun-related skin issues, it does not mean they are completely protected. It's really important to remember that certain skin conditions might be tougher to notice on darker skin tones, and this can sometimes lead to a delay in getting a diagnosis.

2.5 Skin Cancer Facts & Statistics

Skin cancer is the most common cancer worldwide, with incidence and treatment costs rising steadily in many countries. The main types include basal cell carcinoma (BCC), squamous cell carcinoma (SCC), and melanoma [20].

2.5.1 Incidence and Prevalence

- In England (2019), BCC incidence rate was 282.36 per 100,000 person-years, SCC was 85.24, and melanoma was 27.246 [19].

- In the United States, over 5 million adults had treatment for skin cancer per year from 2007 to 2011 [21].

- In South Africa, Basal cell carcinoma (BCC) was the most prevalent type (54.8%), followed by squamous cell carcinoma (SCC) (18.9%), squamous cell

carcinoma in-situ (8.0%), Kaposi's sarcoma (6.7%), and malignant melanoma (6.1%), according to a Western Cape study that found that 68% of clinically suspected cases were confirmed as malignant [22].

- According to the Global Burden of Disease (GBD) 2017 study, which analyzed data from 195 countries worldwide, findings show that squamous cell carcinoma had the highest increase in incidence among all cancers, rising by 310% between 1990 and 2017. Men had higher rates of keratinocyte carcinoma at all ages. For melanoma, women had higher rates up to around age 50 [23].

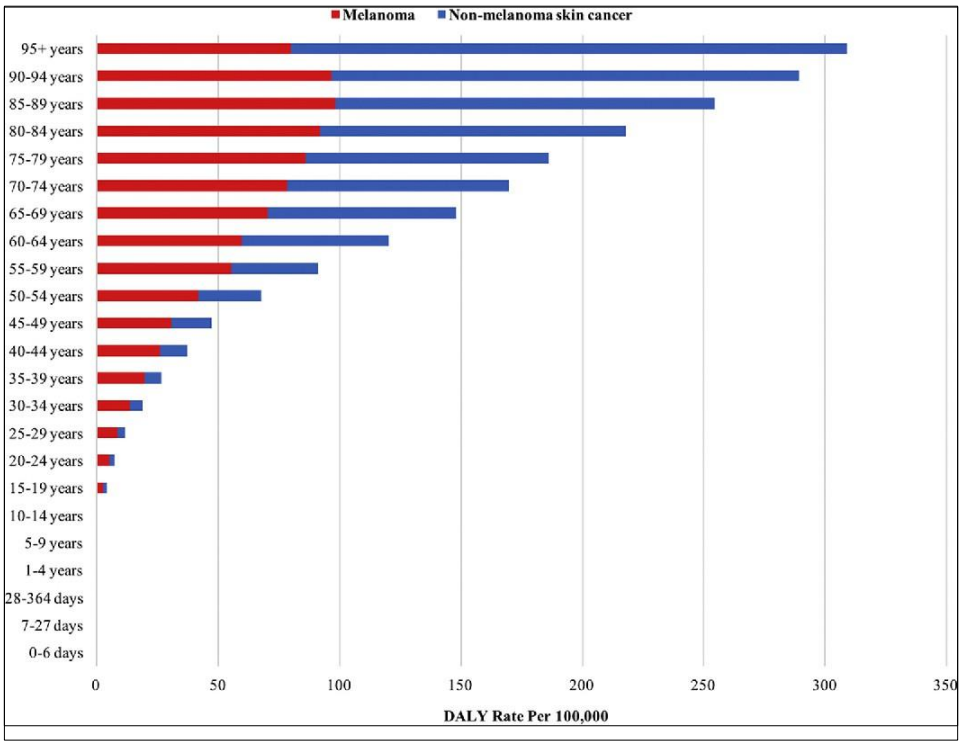

Figure 9:Melanoma and keratinocyte carcinoma disability-adjusted life years per 100,000 population at the global level by age in 2017.

2.5.2 Survival and Economic Impact

- Around 98% to 90% of people with SCC and melanoma will still be alive after 5 years, while over 100% of people with BCC will still be alive [19].

- More than 60% of people with late-stage melanoma, Merkel cell cancer, and genital SCC will still be alive after 5 years [19].

- In 2015, skin cancer treatments in South Africa cost an estimated 92.4 million ZAR (about $15.7 million) [24].

- Skin cancer treatment expenses in the United States increased by 126%, from $3.6 billion in (2002–2006) to $8.1 billion in (2007–2011) [21].

2.5.3 Prevention and Early Detection

- Skin cancer can be diagnosed up to 14 years earlier when individuals at high risk are identified through early detection programs and risk prediction tools [25].

- Prevention strategies focus on reducing UVR exposure and promoting protective behaviors such as: [26]

    - ✓ Using sunscreen with adequate SPF
    - ✓ Wearing protective clothing and hats
    - ✓ Avoiding tanning beds

2.6 Disability-Adjusted Life Years (DALYs)

DALYS is a metric used to measure the overall burden of disease. It reflects the total number of years lost due to illness, disability, or premature death within a given population [32].

DALYs are calculated as the sum of two components:

- Years of Life Lost (YLL): This measures the years lost due to premature death compared to a standard life expectancy.

- Years Lived with Disability (YLD): This measures the years lived with a health condition or its consequences, adjusted by a severity weight.

Formula:

$$DALYs = YLL + YLD$$

2.7 Conclusion

In this chapter, we introduced the topic of skin diseases by explaining their different types, how they are diagnosed, and how skin tone can affect the diagnosis process. We also shared important facts and statistics about skin cancer to show how serious and widespread these diseases can be.

CHAPTER III

OVERVIEW OF ARTIFICIAL INTELLIGENCE AND NEURAL NETWORKS.

3.1 Introduction

In this chapter, we will talk about Artificial intelligence(AI) and Machine learning (ML) and have improved significantly in medical image analysis over the past few years, particularly in classifying dermatological conditions. Convolutional Neural Networks (CNNs) have been one class of neural network with superb performance in skin condition classification by extracting features from images systematically. Deep Learning is one of the branches of ML that enables automatic and intricate pattern recognition and therefore lifts diagnostic accuracy . Transfer Learning, with the help of pretrained models such as ResNet50 and Xception enables users to process data from huge databases with minimal labeled data. This study employs the benchmark datasets HAM10000 to classify seven skin diseases using advanced deep learning techniques aimed at enhancing diagnostic accuracy. These approaches tackle the challenge of the traditional diagnoses that rely on good manual experts by leveraging advanced automated AI systems, enhancing the performance and consistency of automatic dermatology [33].

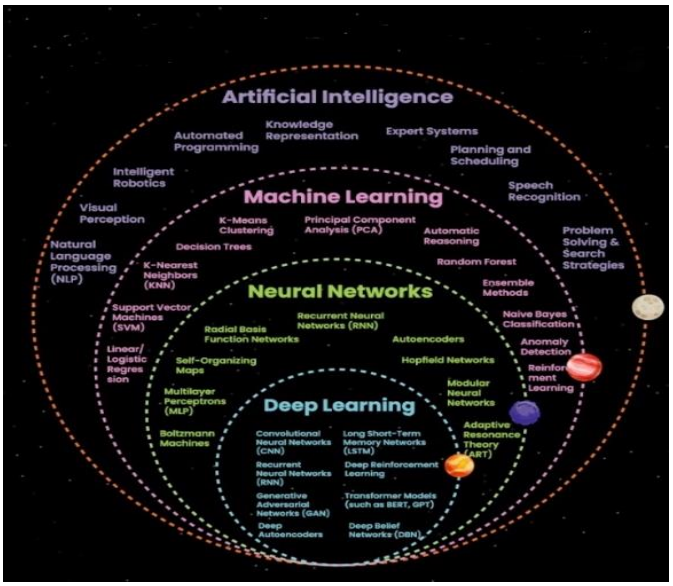

Figure 10:The relation between IA & ML &NN & CNN

3.2 Artificial intelligence

Artificial intelligence (AI) is the intelligence demonstrated by machines. AI in computer science is the study of "intelligent agents," products that perceive their environment and act to maximize their chances of success at some task. Generally, however, the term "artificial intelligence" will be used when a machine mimics the cognitive processes typical of the human mind, "learning" or "problem-solving.".

AI functionalities at present comprise comprehending human language, high level strategic play in games such as chess and Go, autonomous automobiles, smart routing in content delivery networks, war gaming, and deciphering intricate information. The main goals of AI research are reasoning, knowledge, planning, learning, natural language processing (communication), perception, and the capability to move and manipulate objects. Acquiring general intelligence is some of the long term goals in the field. Statistical methods, computational intelligence, and classical symbolic AI are some of the approaches to AI. Search algorithms, mathematical optimization, logic, and probability and economics based methods are some of the tools used in AI. AI is a field that draws from an extensive list of disciplines including computer science, mathematics, psychology, linguistics, philosophy, neuroscience, and artificial psychology, among others[34].

3.3 Deep learning

Deep learning, also known as deep structured learning, hierarchical learning, or deep machine learning, is a field of artificial neural networks and algorithms that use multiple layers of nonlinear processing units for feature extraction and transformation. These algorithms can be supervised or unsupervised and can be used for pattern analysis and classification. DL is part of a broader machine learning field that focuses on learning representations of data, such as faces, facial expressions, and more abstract representations like edges and regions. Research in this area aims to create better representations and models for learning from large-scale unlabeled data, inspired by neuroscience and neural coding. Deep learning architectures, such as deep neural networks, convolutional networks, deep belief networks, and recurrent neural networks, have been applied to fields like computer vision, automatic speech recognition, natural language processing, audio recognition, and bioinformatics, producing state of the art results on various tasks [34].

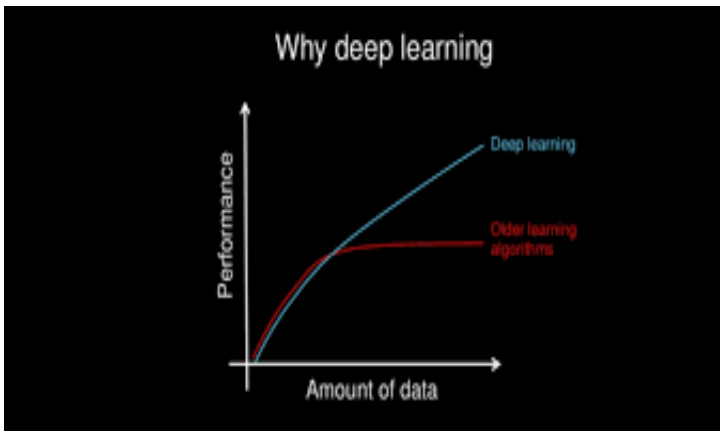

Figure 11:Other Algorithms VS deep learning

3.4 Transfer Learning

Transfer learning is a machine learning technique where a model trained for a given function is reused to enhance accuracy on a different, related function. It's particularly useful with deep learning because training complex neural networks is made possible even with limited labeled data through pre-trained models. Architectures like InceptionV3 or ResNet50, for instance, initially trained on large datasets such as ImageNet, can be reused for classifying medical images, for example, images for skin lesions [35]. It not only reduces

computational cost and accelerates training but also maintains a high accuracy rate, making transfer learning a valuable tool for areas such as dermatology and more.

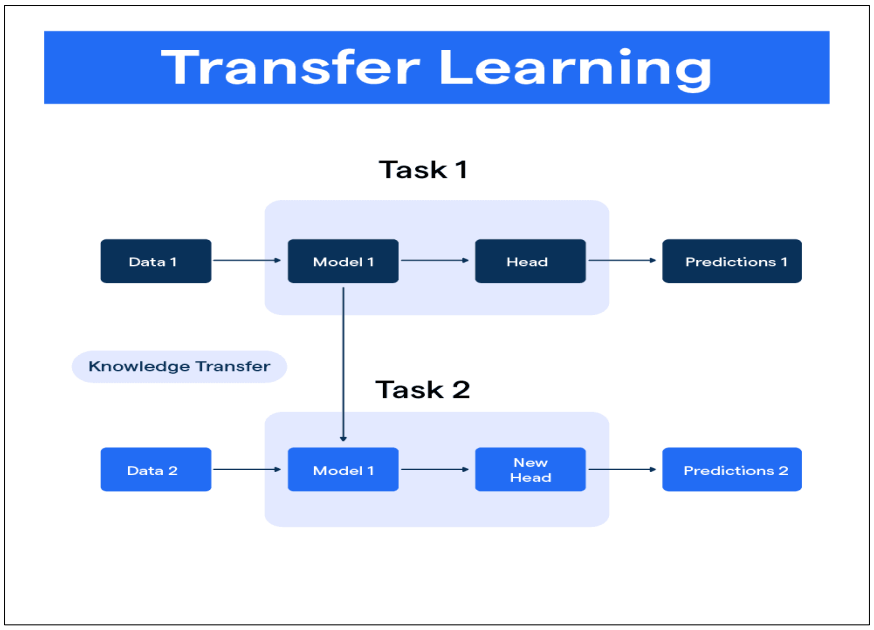

Figure 12:Technique of Transfer Learning

3.5 Brain's Neural Network

The brain can efficiently encode, store, and retrieve information thanks to neural networks that are responsible for memory. These networks are made up of interconnected neurons that form associative and auto associative networks. By using recurrent synaptic connections that get stronger with learning, key areas such as the hippocampus allow partial inputs to trigger the recollection of whole memories through pattern completion. Weighted connections between nodes in this biological process adapt during training to identify and recall patterns, much like in artificial neural networks. Both systems use weight updates in artificial networks and synaptic plasticity in the brain to adjust connection strengths in order to dynamically retain information. Therefore, for tasks like classification and recall, machine learning models that imitate associative memory and pattern recognition processes are inspired by the neuronal memory networks found in the brain [49].

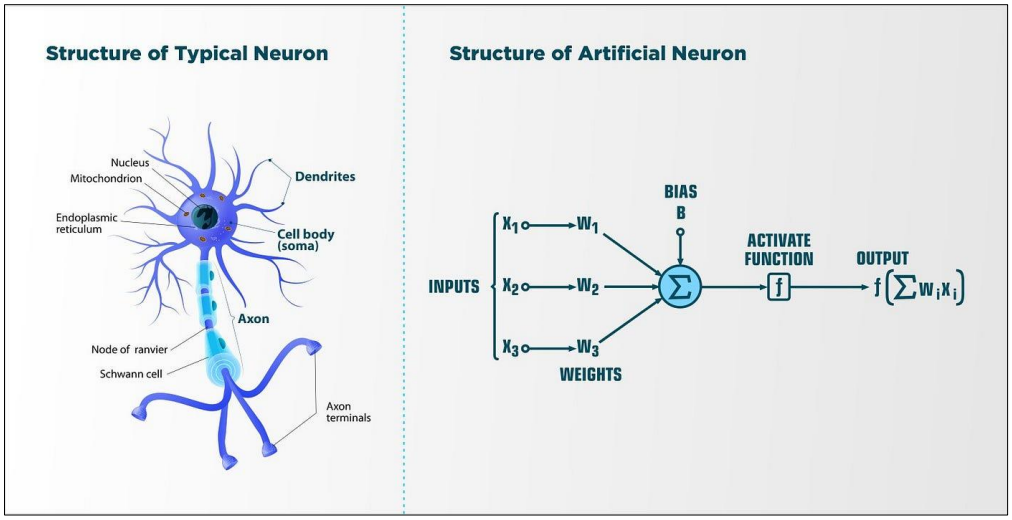

Figure 13:Structure of Typical neuron and Artificial neuron

3.6 Convolutional neural networks (CNNs)

Convolutional neural networks, often known as (CNNs or ConvNets), are feed forward artificial neural networks used in machine learning where the connectivity pattern between their neurons is modelled after the structure of the visual cortex in animals. The receptive field is a constrained area of space where individual cortical neurons react to inputs. The visual field is tiled by the partial overlap of the receptive fields of various neurons. A convolution process can be used to mathematically model how a single neuron responds to inputs within its receptive region. Convolutional networks, which are versions of multilayer perceptron made to require less preprocessing, were inspired by biological processes. They are widely used in natural language processing, recommender systems, and image and video recognition [34].

Convolutional neural networks (CNNs), for instance, use several layers of receptive fields in image recognition applications. These are groupings of tiny neurons that process specific areas of the input image. To create a higher-resolution representation of the original image, the outputs of these collections are then tiled so that their input regions overlap this process is done for each layer of this type. CNNs can withstand input picture translation thanks to tiling [34].

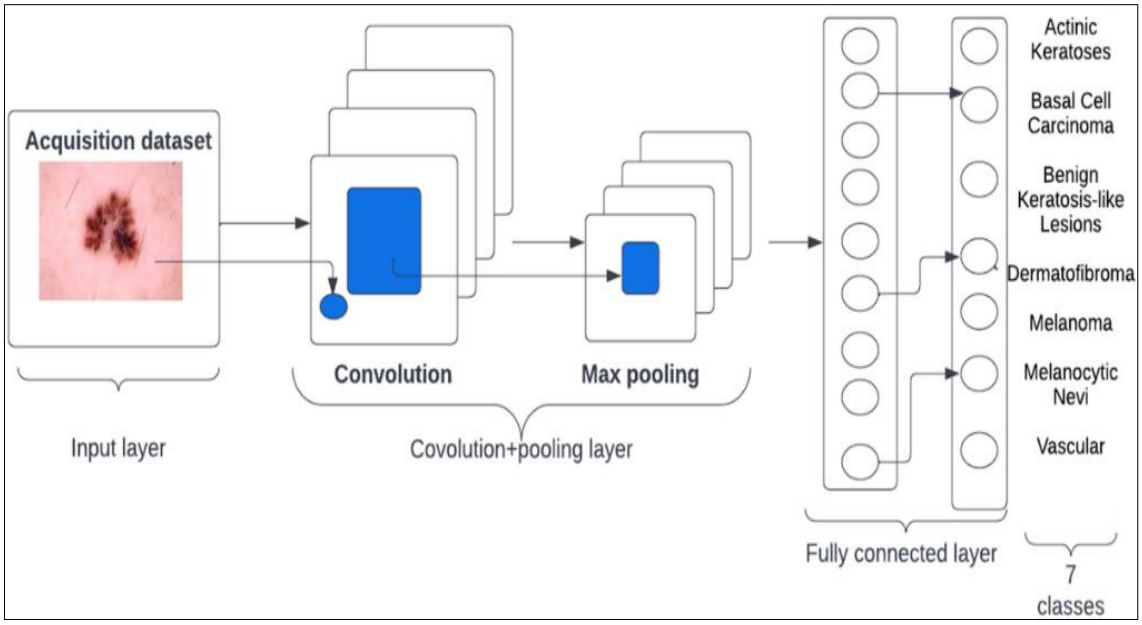

Figure 14:Architecture of CNN

3.7 Blocks of CNN architecture

CNNs comprise three main kinds of layers that are convolutional, pooling, and fully connected layers. The process begins at the convolutional layer, and filters pass through the input image to detect the features like edges or texture by performing convolution operations. These feature maps are passed to activation functions like ReLU in order to provide non linearity. The pooling layers then down sample the spatial dimensions of feature maps, keeping important information while reducing computational complexity. Finally, the fully connected layers convert the learned features to make predictions or classifications. Hierarchical features are learned automatically by CNNs, starting with simple features in shallow layers and proceeding towards intricate patterns in deeper layers. This makes them very effective for tasks like image recognition, object detection, and medical image processing, outdoing the traditional neural networks since they can automate the feature extraction process and handle huge data sets without any hassle [45].

3.7.1 Convolutional layer

Convolutional layers are the chief constituents of CNN architectures. CONV layers alter The input image is transformed with a filter or kernel. The dot product in the region is calculated by the layer. The input layer neurons and the filters create the feature map. Figure 15 demonstrates this [46].

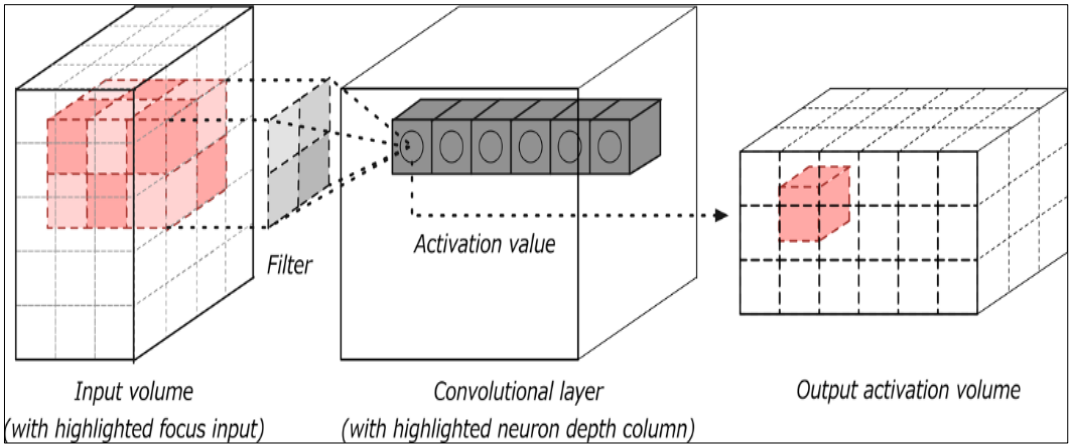

Figure 15:Convolution layer with input and output volume

The output generated after the convolution has the same dimensions usually as the input as seen in figure 16.

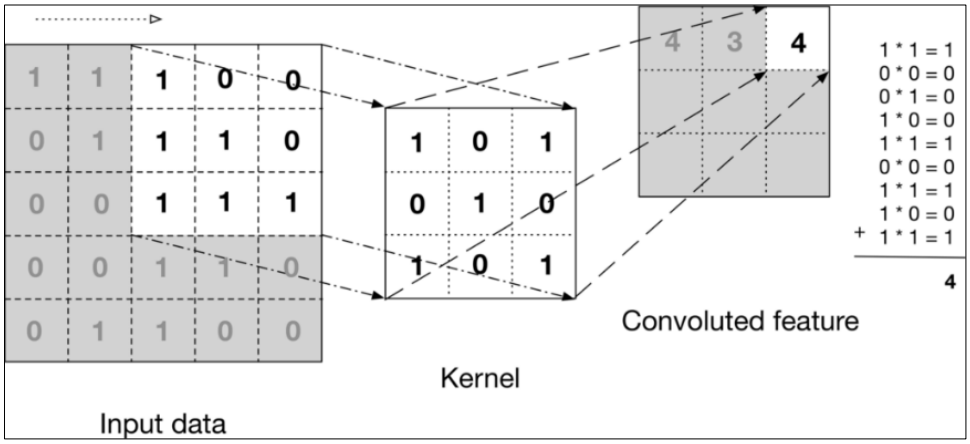

Figure 16:The convolution operation

As illustrated in figure 16, the kernel or filter is smaller than the input size that it slides over. It utilizes one or more. The input data's stride value is used to form a convolved feature. It is also referred to as a feature detector. The activation map, or feature map (figure 17), for any one filter is summed together along Depth serves to produce the 3D result. How the feature detector operates makes it clear. The activation value. It means that each filter will learn to seek out some particular feature. The filter relocates it generates an activation map with two directions in the given filter [46].

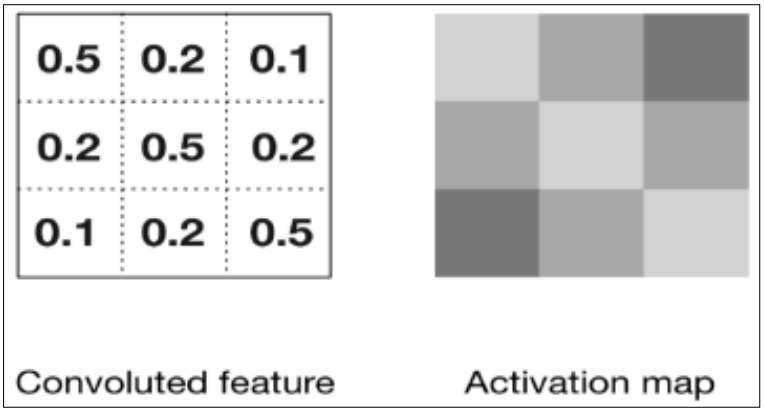

Figure 17:Convolution and activation maps

Stacked activation maps form the output volume. Activation values in this volume align with one another. The nerve cells' signals occupy an area of input space that is small.

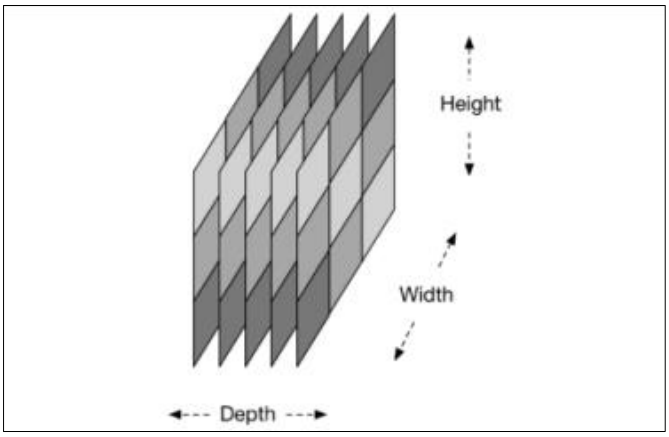

Figure 18:Activation volume output of convolutional layer

- **Padding and Stride**

  Padding and stride are key hyperparameters in convolutional neural networks (CNNs) that control spatial output dimensions. Padding is a process of adding zeros around the input borders to preserve spatial resolution and allow filters to meaningfully process edge-located pixels. For example, "same" padding ensures input output dimension equality, calculated as:

$$p = \frac{f-1}{2}$$ (where f is the filter size) [39]

Stride defines the number of pixels that the filter moves during convolution: larger strides result in reduced output size and reduced computational requirements[39]. The output dimensions are computed using

$$Output\ size = \left[\frac{n + 2p - f}{s} + 1\right]$$

n =input size, p =padding, f =filter size, and s =stride

3.7.2 Pooling Layer

Pooling layers are subsequent from consecutive convolutional layers. Pooling layer in CNNs combines the output of adjacent neurons within the same map/filter to reduce the spatial size of the feature maps. Pooling layers reduce the representation of the data and prevent overfitting from happening. All pooling compress the space they cover. Max pooling reduces the size of the inputs spatially by using the max() function. For example with a 2 x 2 filter, the max() uses the highest of the four values. The result is a condensed feature map [46]. There are two main types of pooling:

- **Max pooling:** The filter chooses the pixel with the highest value to send to the output array as it passes over the input. In addition, this approach is typically employed more often than ordinary pooling.
- **Average pooling:** The filter determines the average value in the receptive field to send to the output array as it passes over the input.

Pooling layers give CNNs a lot of benefits, but they also lose a lot of information. They lessen the chance of overfitting, increase efficiency, and simplify things.

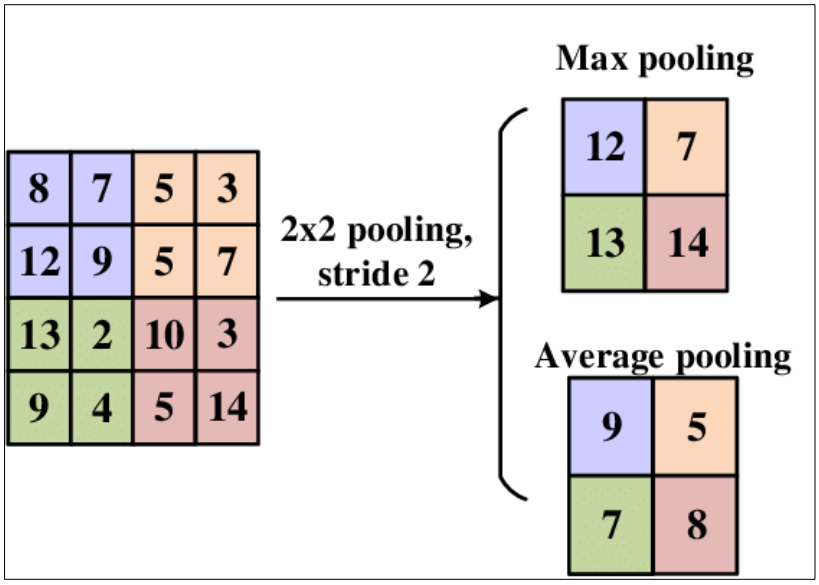

Figure 19:Pooling Layer

3.7.3 Fully Connected Layer

This layer is responsible for calculating the probabilities for the output classes for the given input data. The output produced is a vector of N values, each of which represents the probability for one of the N output classes [46].

3.8 Concepts and techniques

3.8.1 Forward Propagation

In forward propagation, or the forward pass, input data is provided to the neural network, and calculations are carried out layer by layer until the output is obtained. Every neuron in the network takes input signals, applies some activation function to them, and sends the output to neurons in the next layer. This goes on until the final output is calculated. Forward propagation is accountable for predicting using the present parameters (biases and weights) of the network.

3.8.2 Weights

Weights are quantitative parameters used to specify the strength of connection between neurons within a neural network. They are trained using optimization methods such as gradient descent in order to minimize prediction errors. Weights are initialized randomly but learn to give more importance to similar input features, allowing the network to

generalize patterns. Re-calibration of weights is an easy, but a lengthy process. The only nodes where we know the error rate are the output nodes. Re-calibration of weights on the linkage between hidden node and output node is a function of this error rate on output nodes [47].

3.8.3 Activation Function: ReLU

Neuron output activation is enabled by the use of the Rectified Linear Unit (ReLU) nonlinearity. The input x, $f(x) = \max(0, x)$. CNNs with ReLUs train in lesser time than their equivalents with Hyperbolic tangent function. Also, the CIFAR-10 data achieves a training error of 25% in significantly less time. iterations. One of the reasons for using ReLUs is that they do not demand normalized inputs. Inputs with positive values show learning happening in those neurons. Thus, local response normalization helps to reduce error rate. ReLU does not saturate and thus does not have the vanishing gradient problem, unlike the sigmoid or tanh function. But it can lead to dying neurons if inputs are always negative [46].

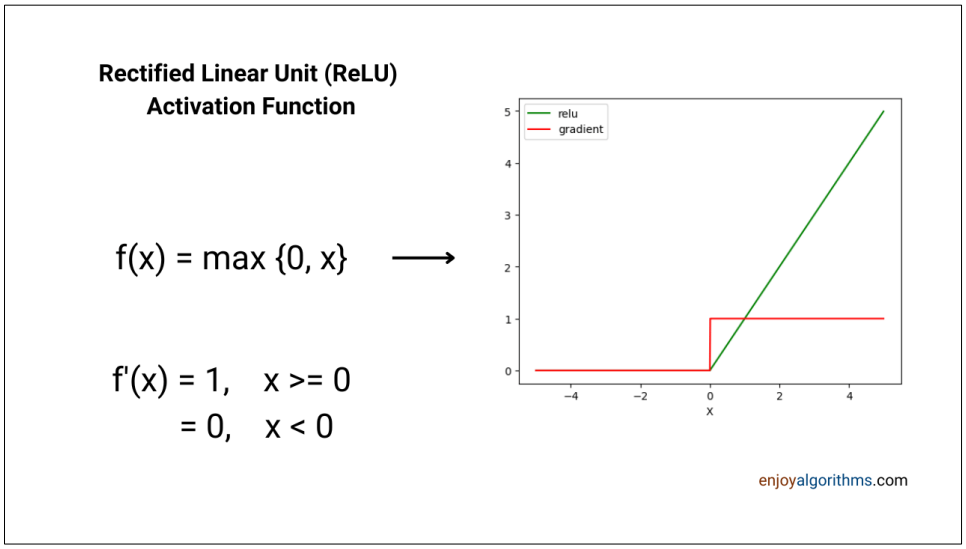

Figure 20:Activation ReLU Function

3.8.4 Loss function

After obtaining the output, the second part is to assess the model's performance by comparing the model's predictions against the true target values of the training data. This is accomplished with a loss function, also referred to as a cost function or objective function, that measures the disparity between predicted output and the ground truth.

Examples of common loss functions include mean squared error (MSE), cross-entropy loss, and hinge loss. At training time, the goal is to minimize this loss function so that the model's accuracy will be better [1].

$$loss\ function = (yi - \hat{y})^2$$

$$Cost\ function = \frac{1}{2n}\sum_{i=1}^{n}(yi - \hat{y})^2 = \frac{1}{2n}\sum_{i=1}^{n}(yi - b0 - b1xi)^2$$

3.8.5 Gradient descent

Gradient descent is a popular optimization technique that reduces loss by cyclically modifying the network's learnable parameters, such as kernels and weights. The direction of the function's sharpest climb is indicated by the gradient of the loss function. Thus, using a user-specified step size established by the learning rate hyperparameter, each learnable parameter is adjusted in the opposite direction of the gradient (see Figure 2.13). The gradient can be expressed mathematically as a partial derivative of the loss with respect to each learnable parameter. The expression for an individual parameter change is as follows:

$$\omega := \omega - \alpha * \frac{\partial L}{\partial \omega}$$

where w stands for each learnable parameter, stands for a learning rate, and L stands for a loss function [1].

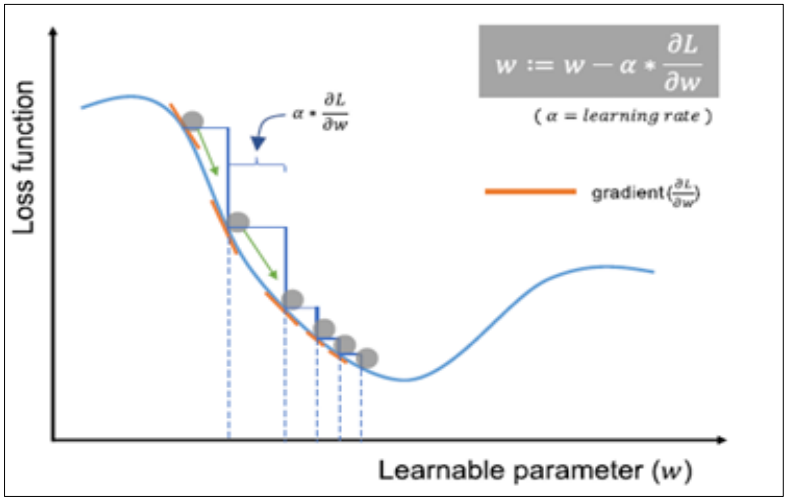

Figure 21:Gradient descent is an optimization algorithm that iteratively updates the learnable parameters so as to minimize the loss, which measures the distance between an output prediction and a ground truth label. The gradient of the loss function.

3.8.6 Vanishing gradient

This occurs when gradients become extremely small during backpropagation, preventing weight updates in deep networks. Solutions include the use of activation functions like ReLU or normalization techniques to stabilize training.

3.8.7 Optimizer

Optimizers such as stochastic gradient descent (SGD) and Adam update weights and learning rates to reduce loss. These algorithms utilize gradients from the loss function to update parameters, with more sophisticated methods incorporating momentum or adaptive learning rates to prevent local minima and enhance convergence.

3.8.8 Backward propagation

Once the loss function has been established, the network will modify its internal parameters, the weights and biases, to reduce the loss and enhance performance. Backward propagation, or backpropagation, is the modification of these parameters by computing the gradient of the loss function with regard to each parameter. This is done by backpropagating the error backwards through the network layer by layer, employing methods like the chain rule of calculus. The gradients are then utilized to update the parameters through optimization methods like gradient descent or variants. By continuously updating the parameters based on gradients calculated, the network iteratively learns to predict more accurately and reduce errors [47].

Collectively, these three processes forward propagation, calculation of the loss function, and backward propagation are fundamental processes of artificial neural network training. They all allow the network to learn from data, adjust parameters, and enhance its performance over time, ultimately enabling it to solve difficult tasks and make precise predictions.

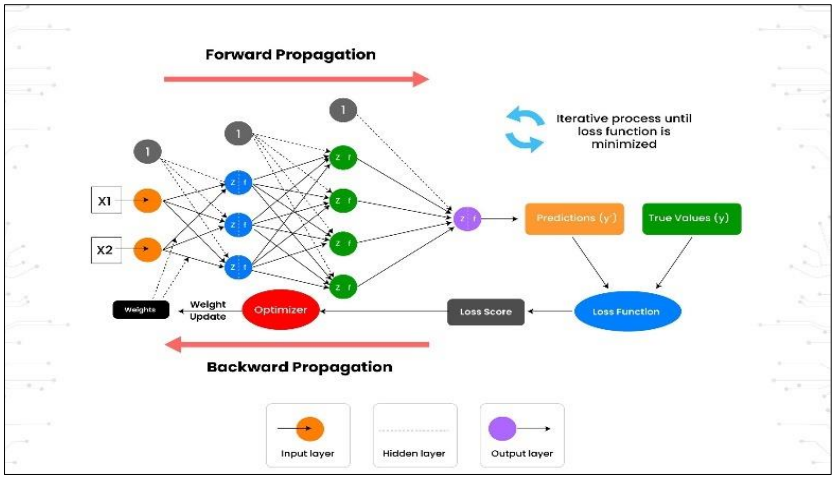

Figure 22:Forward and Backward propagation

3.8.9 Batch Size

Batch size controls the number of training examples that are updated between weight updates. Small batches reduce memory use but increase noise, whereas large batches stabilize the updates at the cost of computation.

3.8.10 Epochs

An epoch is a once-through of the whole training set by the network. The model typically requires more than one epoch to converge, with early stopping being used to avoid overfitting.

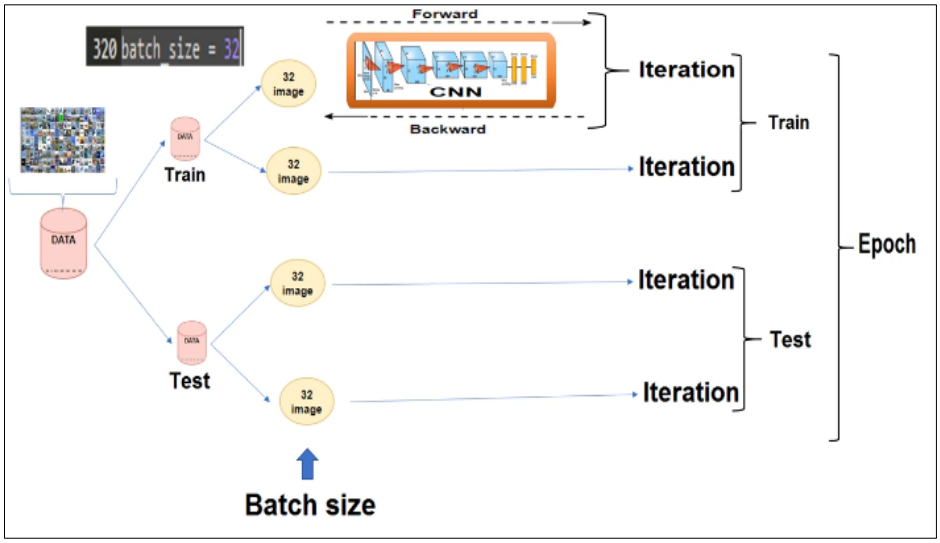

Figure 23:Relation between batch size and epochs

3.8.11 Split valid technique

In machine learning, split validation is a method where the dataset is randomly split into two parts: one for training the model and the other for testing how well it works. The training set is used to teach the model what to do, and the test set is used to see how well it does on data it hasn't seen yet. This method checks to see if the model works much better on training data than on test data, which helps find overfitting.

3.8.12 Learning rate

The learning rate (α) is a hyperparameter that regulates the step size during weight updates. A high rate leads to oscillation or divergence, while a low rate slows down convergence. Ideal values trade off speed and accuracy, usually found using methods such as learning rate schedules or adaptive methods (Adam).

3.8.13 SoftMax

For multi-class classification problems, the last layer of neural networks uses the SoftMax activation function. By exponentiating each score and then normalizing them by dividing by the sum of all exponentials, it converts raw output scores (logits) into probabilities. This guarantees that the output values can be interpreted as probabilities since they are between 0 and 1 and add up to 1. Normalize by the sum of these exponentials:

$$Pi = \frac{e^{Zi}}{\sum_{j=1}^{N} e^{Zj}} .$$

3.9 Different Models of CNN algorithms

3.9.1 Xception

François Chollet unveiled Xception, a deep convolutional neural network architecture, in 2016. It expands and develops upon the Inception model by substituting depthwise separable convolutions for its modules. Inception modules are interpreted by this design as a transitional step between standard convolutions and depthwise separable convolutions. The latter are composed of a pointwise 1x1 convolution (combining features across channels) and a depthwise convolution (applying a single filter per input channel to capture spatial features). Compared to InceptionV3, Xception performs better on large-scale picture classification tasks by making better use of model parameters without increasing their number by completely separating spatial and cross-channel correlations. Stacks of depthwise separable convolutional layers with batch normalization, ReLU activations, and residual connections make up the architecture's three flows entry, middle, and exit. This framework makes Xception accurate and efficient by lowering computing complexity while preserving powerful feature extraction capabilities[40].

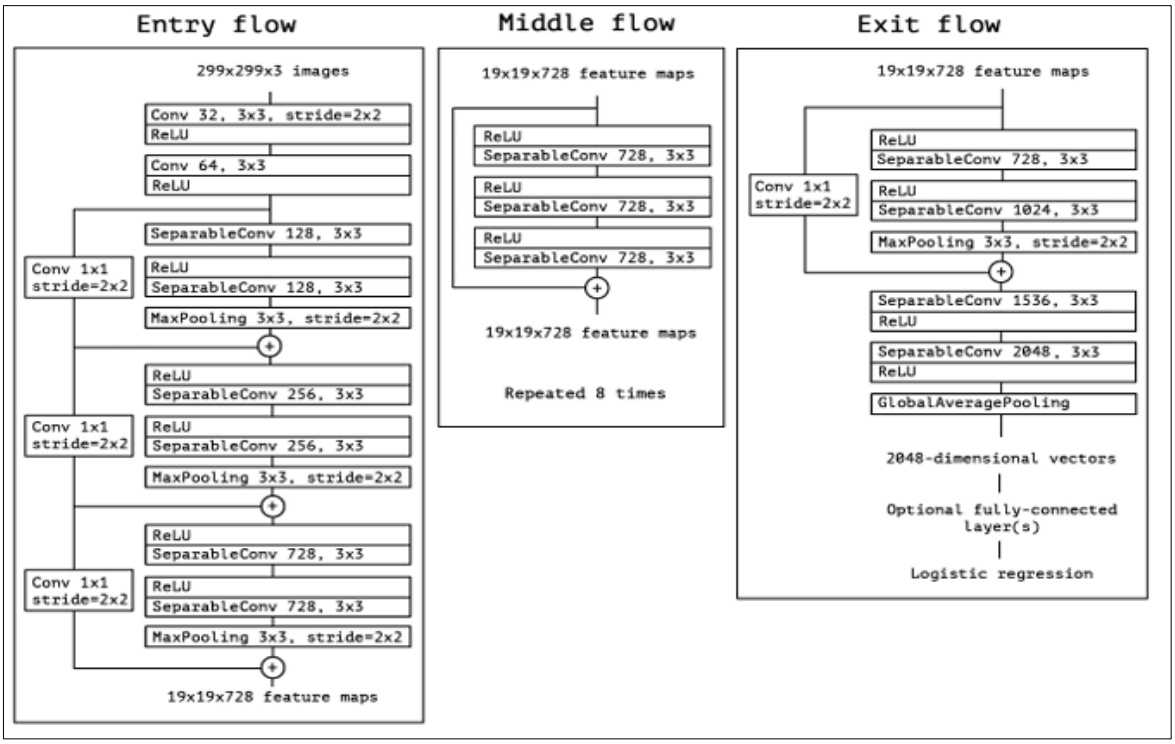

Figure 24:The Xception architecture: the data first goes through the entry flow, then through the middle flow which is repeated eight times, and finally through the exit flow.

Table 2:Xception Architecture

Stage	Layer Type	Output Size	Repetition	Details
Input	Image	299×299×3	-	-
Entry Flow	Conv 3×3, stride 2	149×149×32	1	Convolution
	Conv 3×3	147×147×64	1	Convolution
	Separable Conv Block	73×73×128	1	2 Separable Conv + MaxPool
	Separable Conv Block	37×37×256	1	2 Separable Conv + MaxPool
	Separable Conv Block	19×19×728	1	2 Separable Conv + MaxPool
Middle Flow	Separable Conv Block	19×19×728	8	3 Separable Conv each block
Exit Flow	Separable Conv Block	10×10×728	1	2 Separable Conv + MaxPool
	Separable Conv Block	10×10×1024	1	2 Separable Conv
	Global Average Pooling	1×1×1024	1	-
	Fully Connected	1×1×1000	1	SoftMax

3.9.2 ResNet-50

ResNet50, proposed by Microsoft Research in 2015, represented a breakthrough in deep learning by addressing the vanishing gradient problem prevalent in deep neural networks. This is achieved through the use of residual blocks with skip connections, allowing for stable training of all 50 layers. Each residual block incorporates convolutional filters of size 1x1, 3x3, and 1x1 in a "bottleneck" design, efficiently reducing computational requirements without sacrificing the feature hierarchies. The 1x1 filters reduce and increase the dimensions of channels, while 3x3 filters play a key role in capturing spatial patterns, increasing the network's ability to refine features in its layers. By adding the input to the block's output directly through skip connections, ResNet50 eliminates the possibility of performance degradation for deeper layers, achieving state of the art accuracy on the ImageNet dataset and becoming a cornerstone for medical image analysis, object detection, and other applications [44].

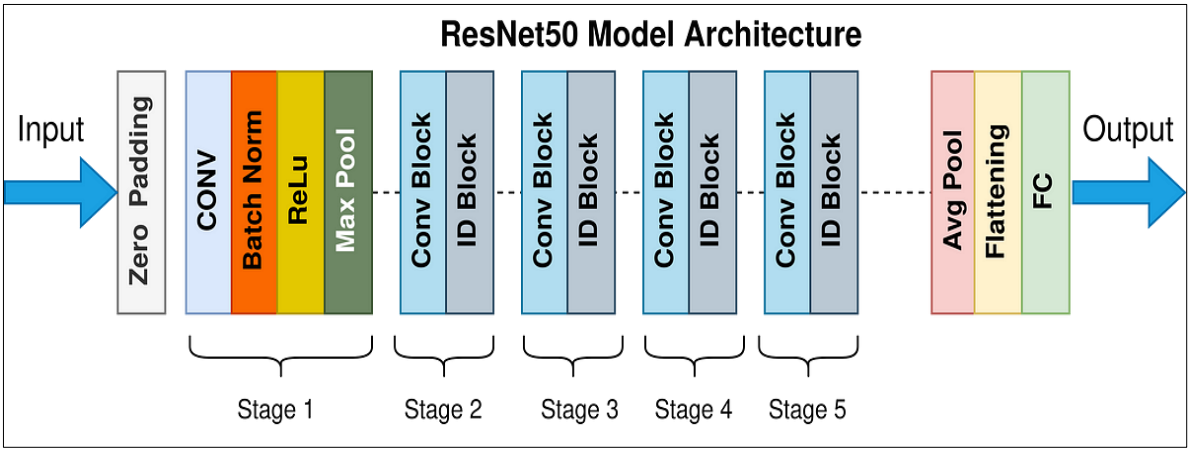

Figure 25:ResNet50 Model Architecture

Table 3:ResNet50 Architecture

Stage	Layer Type	Output Size	Repetition	Details
Input	Image	224×224×3	-	-
Stage 1	Conv 7×7, stride 2	112×112×64	1	Convolution
Stage 2	MaxPool 3×3, stride 2	56×56×64	1	Max Pooling
Stage 3	Conv Block	56×56×256	3	Conv[1×1, 64] → Conv[3×3, 64] → Conv[1×1, 256]
Stage 4	Conv Block	28×28×512	4	Conv[1×1, 128] → Conv[3×3, 128] → Conv[1×1, 512]
Stage 5	Conv Block	14×14×1024	6	Conv[1×1, 256] → Conv[3×3, 256] → Conv[1×1, 1024]
Stage 6	Conv Block	7×7×2048	3	Conv[1×1, 512] → Conv[3×3, 512] → Conv[1×1, 2048]
Stage 7	Average Pooling	1×1×2048	1	-
Stage 8	Fully Connected	1×1×1000	1	SoftMax

3.9.3 EfficientNetV2-B1

Google researchers released EfficientNetV2-B1 in 2021. It is a part of the EfficientNetV2 family and is meant to improve training speed and parameter efficiency compared to its predecessors. EfficientNetV2-B1 adds Fused-MBConv layers, which replace traditional MBConv blocks in early stages by combining 3x3 depthwise separable convolutions with squeeze-and-excitation modules. These layers reduce computational overhead while keeping accuracy. They do this by balancing depth, width, and resolution. By combining spatial and channel wise feature extraction, these fused layers make processes easier and faster. The model also uses adaptive regularization and progressive

learning to keep accuracy stable. It starts training with lower picture resolutions and slowly raises them. EfficientNetV2-B1 is designed to be as light as possible and has 8 stages of optimized MBConv blocks. It works well for tasks like gesture recognition and is therefore perfect for edge devices [43].

Table 4:EfficientNetV2-B1 Architecture

Stage	Layer Type	Output Size	Repetition	Details
Input	Image	240×240×3	-	-
Stage 1	Conv 3×3, stride 2	120×120×32	1	Convolution
Stage 2	Fused-MBConv	120×120×16	2	Conv[3×3, 16], Expansion: 1, SE: 0.25
Stage 3	Fused-MBConv	60×60×24	4	Conv[3×3, 24], Expansion: 4, SE: 0.25
Stage 4	Fused-MBConv	30×30×48	4	Conv[3×3, 48], Expansion: 4, SE: 0.25
Stage 5	MBConv	15×15×96	6	Conv[3×3, 96], Expansion: 4, SE: 0.25
Stage 6	MBConv	15×15×112	9	Conv[3×3, 112], Expansion: 6, SE: 0.25
Stage 7	MBConv	8×8×192	15	Conv[3×3, 192], Expansion: 6, SE: 0.25
Stage 8	Conv1×1 + Pooling	8×8×1280	1	[1×1 conv], Global Average Pooling
Stage 9	Fully Connected	1×1×1000	1	Dense Layer + SoftMax

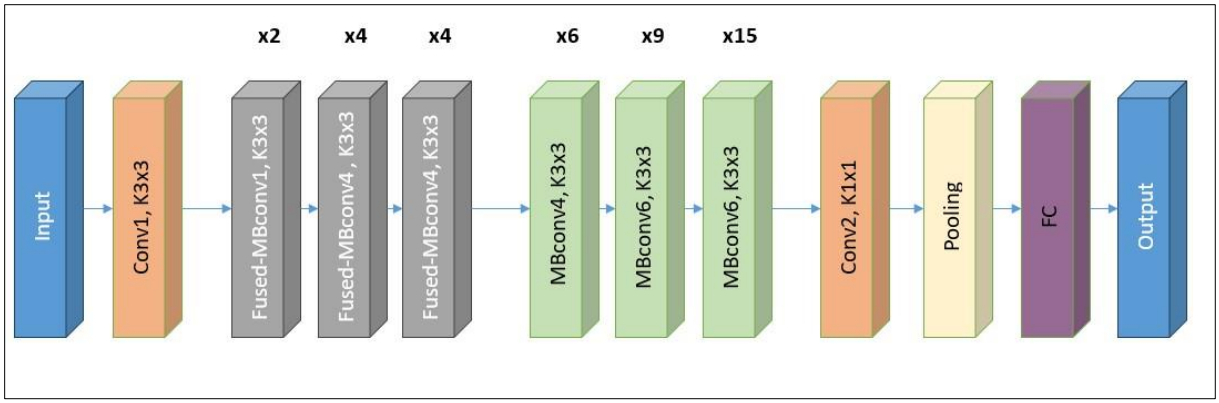

Figure 26:EfficientNetV2-B1 Architecture

3.10 Conclusion

The preceding sections have made an effort to present an examination of the principles of deep learning algorithms at a level beyond the simple basics. We started with a comparative analysis between "deep learning" and "machine learning", . That was followed by a detailed explanation of convolutional neural networks (CNN). In addition, we reviewed popular CNN architectures with a focus on four models Xception, ResNet50, EfficientNetV2B1 and outlined their general architectures. The purpose of this section was to give the reader the necessary background information so that he or she can appreciate the experimental results given later in the paper.

CHAPTER IV

Models Evaluation, Results and Discussion.

4.1 Introduction

With the rapid progress in artificial intelligence (AI) and deep learning, we now have powerful tools that can assist in medical diagnosis especially in detecting and classifying complex conditions like skin diseases. Building on the concepts discussed in the previous chapters, this chapter focuses on the design, implementation, and evaluation of a deep learning-based system aimed at improving early detection of skin conditions, which can ultimately help save lives.

We begin by describing the development environment, including the hardware and software setup used for training and testing our models. After that, we present the experimental results obtained from applying and comparing different deep learning algorithms for skin disease classification.

Next, we analyze the performance of each model using key metrics such as accuracy, precision, recall, and F1-score, to understand which one performs best in identifying various skin conditions. Finally, we discuss the outcomes in detail, reflecting on their practical value and real-world impact in the field of medical diagnosis.

4.2 Work Environment

4.2.1 Hardware Environment

In this section we present the hardware environment used for the application, the characteristics are: a DELL Laptop:

- **Processor:** Intel(R) Core(TM) i7-8650U CPU @ 1.90GHz, 2.11 GHz.
- **RAM:** 16.00 GB.
- **Hard Drive:** 256 GB SSD.
- **System Type:** 64-bit operating system, x64-based processor.
- **OS:** Windows 10 Pro.

4.2.2 Software Environment

For the development and training of our skin disease classification system, we used Google Colab of which provide free, cloud-based environment equipped with high-performance GPUs. Specifically, we utilized the P100 GPU, which significantly accelerated the training process of our deep learning models.

The programming was done in Python, a popular language for machine learning and data science. Several essential libraries were used in this project, including:

- **OpenCV (cv2):** For image processing tasks.
- **NumPy and Pandas:** For data handling and manipulation.
- **Scikit-learn:** For preprocessing and evaluation metrics.
- **TensorFlow:** As the main deep learning framework.
- **Matplotlib:** For data visualization.

Colab environment come pre-installed with most of these libraries, and additional packages were easily installed using built-in commands (!pip install).

4.3 Dataset collection

Neural network learning for computer-based diagnosis of pigmented skin lesions is slowed down by the small size and diversity of the currently available datasets of dermatoscopic images.

- **Dataset Ham10000**

HAM10000 ("Human Against Machine with 10000 training images") dataset. This dataset collected dermatoscopic images from different populations and they archived by different modalities. Because of this diversity we had to employ different acquisition and purification processes and developed semi-automatic workflows with especially trained neural networks. The last dataset contains 10015 dermatoscopic images which are

provided as a training set for educational machine learning and are made available freely using the ISIC archive. This is an appropriate benchmark dataset for machine learning use and comparison with human experts. Cases include a representative subset of all important diagnostic categories in the context of pigmented lesions. More than 50% of the lesions have also been confirmed by pathology, and ground truth for the rest was either follow up, consensus expert, or confirmed by in vivo confocal microscopy [48].

We took this dataset from the Kaggle website and it includes 7 classes, which are Mel, Bkl, Akiec, Df, Nv, Bcc and Vasc . This dataset has been processed, it contains 38569 images with size (224 x 224) pixels.

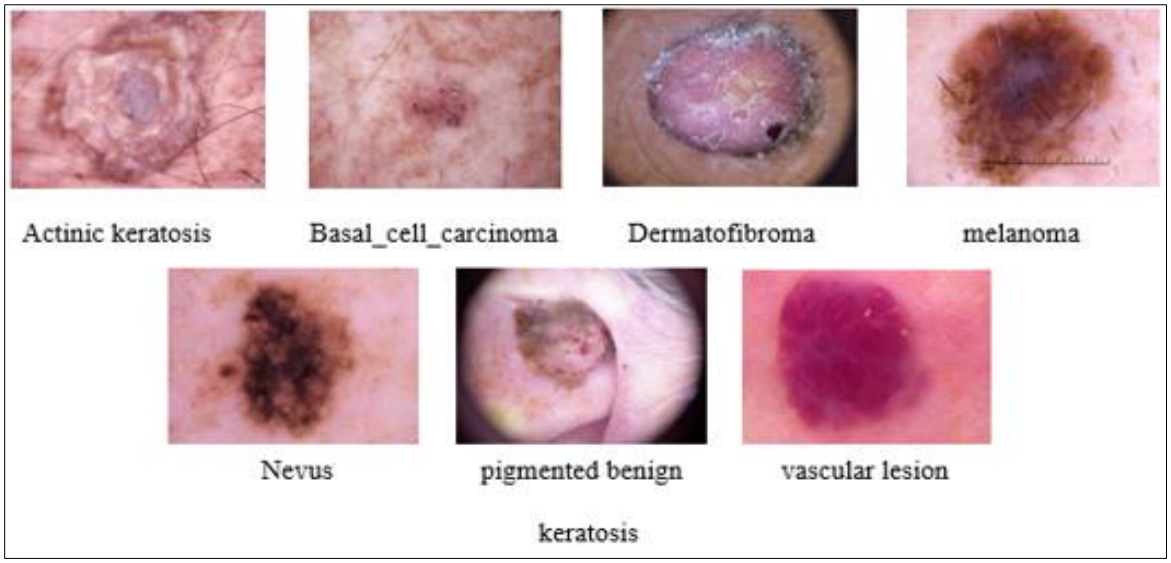

Figure 27:Types of Dataset Used

4.4 Evaluation Parameters

We then subjected the models to their test with the held over test set once they had been trained. Subsequently, the confusion matrix was used in the calculation of performance measures. For purposes of signaling both the intended and actual classifications, portions of the confusion matrix are used. Two classes arise as a result of classification: correct and wrong. To calculate the prediction model, we conducted a study of the following four basic case studies [36]:

- The term "true positive" refers to the percentage of true positives that are identified with high degree of precision (TP).

- Incorrect forecasts are referred to as having a false negative (FN). It finds situations that are malevolent, despite the fact that the model incorrectly anticipated them to be normal.
- A false positive, sometimes known as an (FP), is an incorrectly positive prediction made when the observed assault is, in fact, normal.
- The percentage of false positives that correctly identify attacks is the value that is being measured by the true negative (TN).

4.4.1 Accuracy

Model accuracy is used to refer to the ability of the model to predict the truth or predict a new sample correctly. As a measure, the higher the accuracy, the more the model is able to predict the truth or predict a new sample. From this, a higher model accuracy depends on choosing and using the appropriate tools to train and implement the model based on the available data, technical and cultural requirements, and other determinants [42].

$$Accuracy = \frac{Tp + Tn}{Tp + Tn + Fp + Fn}$$

4.4.2 Precision

Precision AI measure is the measure which identifies the proportion of items that are appropriately categorized as falling under the classification of the said category out of the total items that are categorized as falling under the classification of the said category, whenever new data is reported to a trained machine learning system that is trained to categorize items under different categories. Accuracy is used to evaluate the precision and quality of the machine learning model to identify the correct category [42].

$$precision = \frac{Tp}{Tp + Fp}$$

4.4.3 Recall

Sensitivity measure in AI is a measure of the ability of the artificial intelligence model to correctly identify certain cases, and it mentions the rate of the proportion of actual positive cases that are detected (True Positives) to the total number of correct positive cases

found in the initial data. If the desired value is all positive cases being correctly detected, then the model ought to have a sensitivity value close to 100% [42].

$$Recall = \frac{Tp}{Tp + Tn}$$

4.4.4 F1-score

The F1-score is a popular way to rate machine learning models, which finds the harmonic mean of precision and recall and adds them together into a single number. It gives a fair evaluation of a model's accuracy and is especially helpful when datasets aren't balanced and one class is much less common than others. Precision is the percentage of correctly predicted positive cases out of all expected positive cases. Recall is the percentage of correctly identified positive cases by the model [42].

$$F1 - score = \frac{2 * Precision * Recall}{Precision + Recall}$$

which makes sure that you only get a high score if both accuracy and recall are high. This makes the F1 number very useful in situations where both false positives and false negatives cost a lot. It is a better way to measure performance in these situations than just accuracy.

4.4.5 Confusion matrix

A multi-class confusion matrix is a square table used to test how well classification models with three or more classes work. Each row of the matrix shows the predicted classes, and each column shows the actual classes. The diagonal elements show how many instances were correctly classified for each class, and the off-diagonal elements show misclassifications, which show where the model gets one class mixed up with another. This matrix gives you a lot of information about the model's strengths and weaknesses by showing errors by class, so you can figure out metrics like accuracy, precision, and recall for each class. It is an important tool for understanding and improving multi class classification models beyond overall accuracy [42].

Figure 28:Confusion Matrix multiple classes.

4.5 Data Augmentation Technique

Data augmentation is a powerful strategy used to improve a model's ability to generalize by creating modified versions of existing images. This process enhances the diversity of the training dataset without the need to collect more data, making it especially valuable when working with limited medical image samples.

Various augmentation methods were applied in this study:

- **Rescaling:** Normalizes pixel values, making images easier for neural networks to process and improving convergence speed.
- **Rotation:** Rotates the image around its center. This technique helps the model become invariant to orientation changes, useful in object classification and detection tasks.
- **Shear:** Applies a distortion to the image that simulates changes in perspective. This helps the model adapt to shifts in viewpoint.
- **Zoom:** Scales images in or out. As a geometric transformation, zooming helps the model recognize objects regardless of their size or distance within the frame.

These transformations were applied dynamically during training using TensorFlow, allowing the model to see a new variant of the image at each epoch strengthening its learning without increasing the dataset size on disk.

4.6 Parameter Settings

To ensure consistent training and fair model comparisons, we followed a methodical approach to parameter tuning. In each experiment, we tried to modify one parameter or more at a time to study the changes.

The following settings were explored:

4.6.1 Networks Used:

Xception (Extreme Inception), ResNet50, EfficientNetV2

4.6.2 Batch Size:

Determines how many samples are processed before the model's weights are updated. And at this study we tested different values 32, 64.

Larger batch sizes increase training time and memory usage, sometimes at the cost of accuracy.

4.6.3 Epochs:

Represents how many times the model sees the entire training set. The Values used are: 20, 38, 40 and 50.

4.6.4 Validation Split:

Controls the percentage of data used for validation during training. We tried various splitting ratios: (80% for training & 10% for testing & 10% for validation), (95% for training & 2.5% for testing & 2.5% for validation), (90% for training & 5% for testing & 5% for validation), (85% for training & 15% for testing) and (80% for training & 20% for testing).

Proper splitting ensures that the model is evaluated on unseen data, helping avoid overfitting.

4.6.5 Fine-tuning

Fine-tuning involves unfreezing some layers of the pre-trained model and retraining them on our dataset. We applied it after the initial training to improve accuracy and adapt the model more effectively to our skin disease classification task.

By adjusting and observing these parameters, we aimed to identify the optimal configuration for achieving high accuracy and reliable classification across all tested CNN models.

4.7 Results and Analysis

During model training, we explored various code implementations and adjusted parameters to identify the best-performing setup for each CNN architecture. All experiments were conducted using the HAM10000 dataset as the sole data source. Final results were evaluated across two different train-test-validation split strategies for each model, and the training process is illustrated in the flowchart:

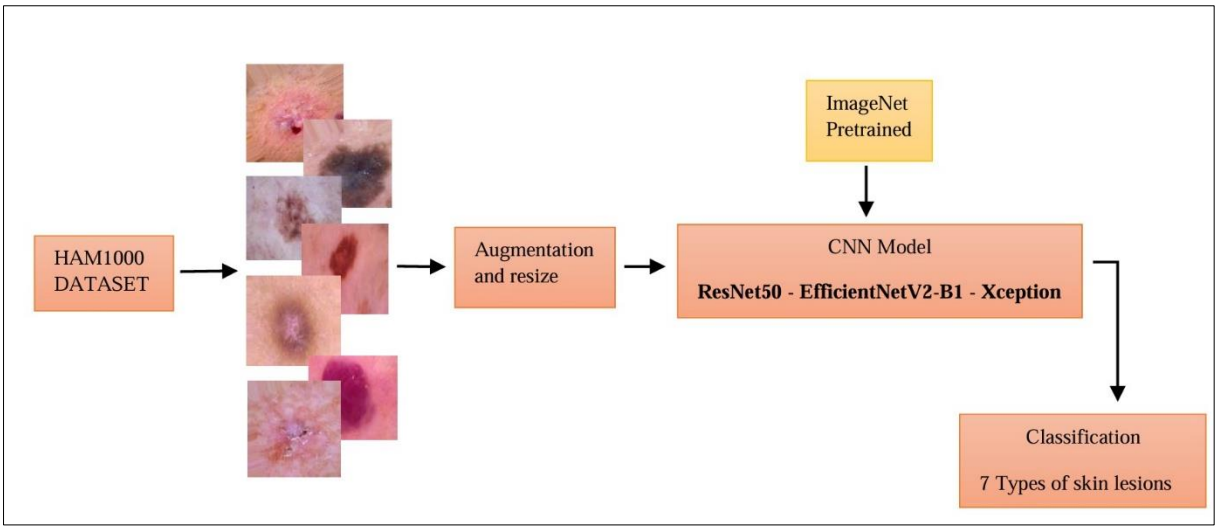

Figure 29:Training Process

4.7.1 ResNet-50 Performance

ResNet50 is widely recognized for its strong feature extraction capabilities, particularly in medical image analysis. We selected it for its proven efficiency in disease classification tasks and applied it to our dataset to evaluate its diagnostic performance in skin cancer detection.

In our experiments, we utilized the ResNet-50 architecture with the parameter (include_top=False). This approach loads the pre-trained base of ResNet50, excluding its final classification layers. By doing so, we retained the powerful feature extraction capabilities of ResNet-50 while allowing the model to train the last layers tailored to our specific dataset. This strategy is a common practice in transfer learning, enabling the model to adapt to new tasks by leveraging learned features from large datasets like ImageNet [50].

The model delivered high accuracy, confirming its reliability in identifying skin-related conditions. Below is a summary of the training setups and outcomes for ResNet50 :

Table 5:ResNet-50 setups and results

Model	Data Splitting	Epochs	Batch Size	Accuracy
ResNet-50	95% Train / 2.5% Test / 2.5% Val	50	32	93%
ResNet-50	90% Train / 5% Test / 5% Val	50	32	90%

The results show that accuracy improved significantly when using 95% of the data for training. With only 90% training data, the model lacked sufficient exposure to diverse patterns, slightly limiting its performance. On the other hand, the 95% split allowed deeper feature learning, leveraging ResNet's capacity to learn hierarchical representations due to its deep architecture of up to 152 layers. This led to better generalization and improved classification accuracy.

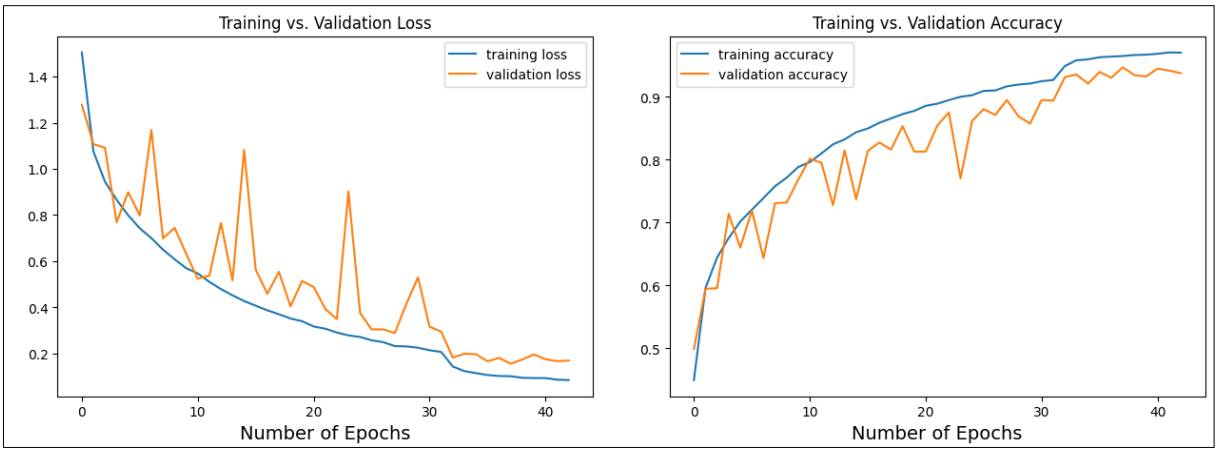

Figure 30:ResNet-50 Training loss and Accuracy curves

Classification Report:

	precision	recall	f1-score	support
akiec	0.95	0.97	0.96	131
bcc	0.97	0.96	0.97	147
bkl	0.87	0.85	0.86	148
df	0.98	1.00	0.99	111
mel	0.88	0.86	0.87	148
nv	0.88	0.90	0.89	150
vasc	1.00	1.00	1.00	133
accuracy			0.93	968
macro avg	0.93	0.93	0.93	968
weighted avg	0.93	0.93	0.93	968

Figure 31:Classification Report of ResNet-50

- **Confusion Matrix and Model Evaluation:**

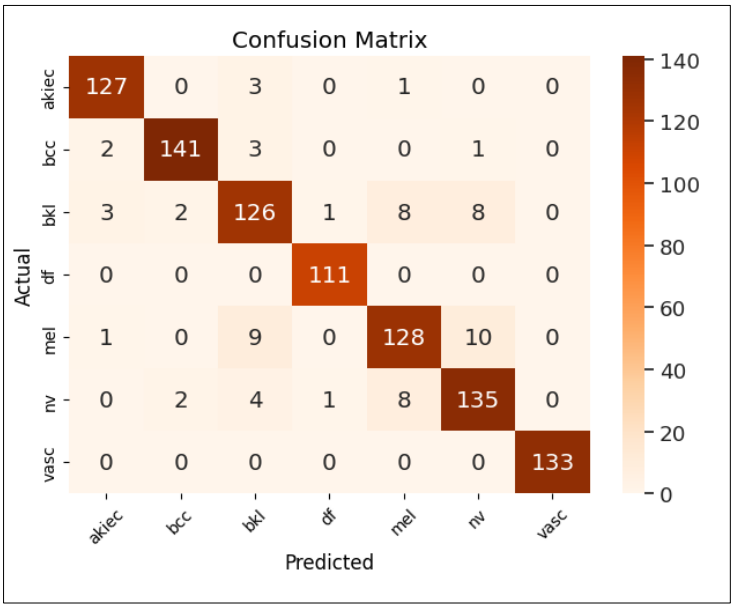

Figure 32:Confusion Matrix of ResNet-50

As shown in the attached confusion matrix figure (32), the model demonstrates strong overall performance in classifying the various types of skin lesions. Most classes are accurately identified, with high true positive rates across the diagonal. However, the model shows some difficulty in distinguishing between benign keratosis-like lesions (BKL) and melanoma (MEL), as there is a noticeable degree of misclassification between these two categories. This overlap is understandable given the visual similarities between BKL and

MEL, but overall, the model is highly effective in recognizing the majority of lesion types with minimal confusion.

4.7.2 EfficientNetV2-B1 Performance

EfficientNetV2-B1 is a high-performing convolutional neural network architecture that incorporates eight stages of optimized MBConv blocks. It is specifically designed to provide a balance between efficiency and accuracy, making it a strong candidate for medical image classification tasks such as skin lesion diagnosis.

In our experiments, EfficientNetV2-B1 demonstrated outstanding accuracy, making it highly reliable for detecting skin cancer.

- **Model Development:** The convolutional neural network architecture used for our model implementation was EfficientNetV2-B1, implemented using Keras and TensorFlow. The deep learning task was executed in the Colab notebook environment, leveraging GPU acceleration for optimized training performance. All input images from the HAM10000 dataset were resized to 240 × 240 pixels, and preprocessing was handled using the built-in (preprocess_input) function specific to EfficientNetV2, which scales pixel values in accordance with the ImageNet training distribution. The EfficientNetV2-B1 architecture used in this work retained its pretrained weights from ImageNet, benefiting from transfer learning. We fine-tuned the entire model to adapt to the domain-specific features of dermoscopic skin lesion images. The architecture consists of 8 optimized MBConv and Fused-MBConv stages, which efficiently balance depth, width, and resolution for improved performance with reduced computational cost. Our model employed data augmentation techniques including random flipping, rotation, zooming, and contrast adjustment to enhance generalization and mitigate overfitting. A cross-entropy loss function was applied, and the final classifier layer used a SoftMax activation function to output class probabilities across seven categories of skin lesions. The training was conducted using the Adam optimizer, with an initial learning rate of 1e-5, and training lasted for 40 epochs with a batch size of 32. Regularization strategies such as dropout layers (0.3 and 0.25 rates) and early stopping were incorporated to further improve generalization. Additionally, a ReduceLROnPlateau callback was employed to dynamically lower the learning rate when validation performance plateaued.

The table below summarizes the training configurations and results:

Table 6:EfficientNetV2-B1 configurations and results

Model	Data Splitting	Epochs	Batch Size	Accuracy
EfficientNetV2-B1	95% Train / 2.5% Test / 2.5% Val	40	32	96%
EfficientNetV2-B1	90% Train / 5% Test / 5% Val	40	32	93%

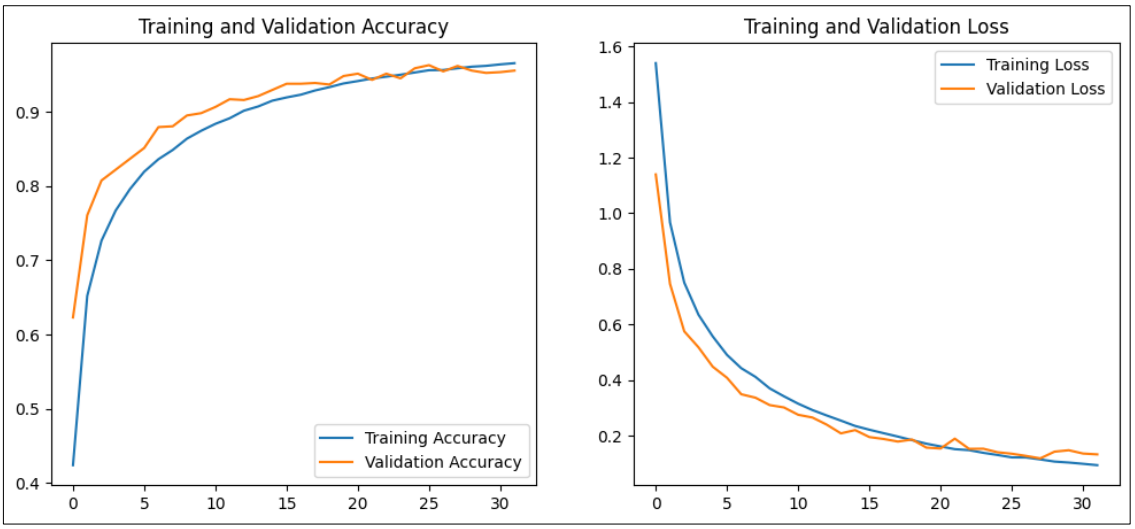

Figure 33:EfficientNetV2B1 Training curves

	precision	recall	f1-score	support
akiec	0.99	0.97	0.98	152
bcc	0.99	0.98	0.99	150
bkl	0.89	0.94	0.92	143
df	1.00	0.99	1.00	120
mel	0.93	0.89	0.91	141
nv	0.91	0.93	0.92	143
vasc	1.00	1.00	1.00	120
accuracy			0.96	969
macro avg	0.96	0.96	0.96	969
weighted avg	0.96	0.96	0.96	969

Figure 34:Efficient Net V2B1 classification report

- **Confusion Matrix and Model Evaluation:**

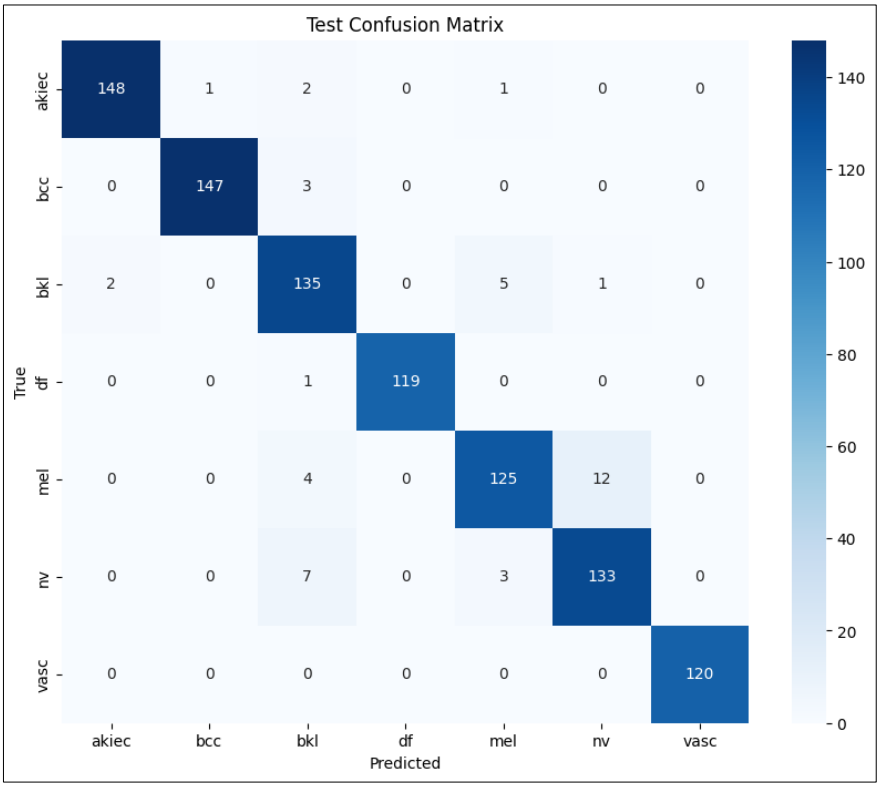

Figure 35:Confusion Matrix of the trained Efficient-Net V2B1 model

As shown in the confusion matrix figure (35), the EfficientNetV2-B1 model demonstrates high precision in classifying skin lesion images from the HAM10000 dataset, achieving a classification accuracy of 96%. The model excels in distinguishing most classes such as “bcc”, “df”, “nv”, and “vasc”, while showing only slight confusion between “bkl” and “melanoma”, which is reasonable given their visual similarity. This high performance underscores the model's ability to generalize well across varied lesion types.

To achieve such accuracy, the model was trained with a low learning rate of 1e-5, which is essential when fine-tuning deep architectures like EfficientNetV2-B1. This approach allows the model to update weights gradually and avoid catastrophic forgetting of pre-trained features from ImageNet.

The architecture used in this experiment is both semi lightweight and deep, enabling it to extract hierarchical and discriminative features. Below is a summary of the model architecture:

Table 7:Trained EfficientNetV2-B1 architecture

Layer (type)	Output Shape	Param #
input_layer_2 (Input Layer)	(None, 240, 240, 3)	0
efficientnetv2-b1 (Functional)	(None, 8, 8, 1280)	6,931,124
global_average_pooling2d (GlobalAveragePooling2D)	(None, 1280)	0
dropout (Dropout)	(None, 1280)	0
dense (Dense)	(None, 128)	163,968
dropout_1 (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 7)	903

4.7.3 Xception Performance

The convolutional neural network architecture implemented for this study was Xception, a high-performing deep learning model that extends the Inception architecture by employing depthwise separable convolutions. The name "Xception" stands for "Extreme Inception", and it demonstrates superior performance on large scale datasets like ImageNet, while maintaining fewer parameters and higher training speed than InceptionV3. For these reasons, Xception was selected as the backbone model for our most accurate skin lesion classification task. The Xception architecture is divided into three functional parts: entry flow, middle flow, and exit flow. The entry flow performs early-stage convolution and down sampling, the middle flow repeated eight times extracts deeper hierarchical features, and the exit flow performs final feature refinement before classification. All layers of the original Xception architecture were retained. Transfer learning was applied using pretrained weights from ImageNet to leverage general visual features and adapt them to the medical imaging domain.

This model was implemented using Keras within TensorFlow in a Google Colab environment, utilizing a Tesla P100-PCIE GPU for accelerated training. Input dermoscopic images were resized to 224 × 224 pixels and normalized to standardize pixel intensity for efficient learning. Preprocessing was handled using the standard (preprocess_input) method for Xception. We trained and evaluated the model using two different dataset split strategies: 90% training, 5% validation, and 5% test for the initial experiment. 95% training, 2.5% validation, and 2.5% test for the refined experiment.

The training process was conducted using the Adamax optimizer, with a learning rate of 0.0001 and batch size of 64. Training was run for 15 and 20 epochs in the respective experiments. To mitigate overfitting, we employed early stopping, halting training when the validation accuracy plateaued.

Table 8:Xception Model setup and results

Model	Data Splitting	Epochs	Batch Size	Accuracy
Xception	90% Train / 05% Test / 05% Val	20	64	96%
Xception	95% Train / 2.5% Test / 2.5% Val	20	64	98%

The model achieved impressive results with an overall accuracy of 96% and 98% in the two configurations. Notably, the confusion matrix (Figure 38) shows excellent class-level performance across all categories. The final architecture included a Flatten layer followed by Dropout, a Dense layer with 128 units (ReLU activation), another Dropout layer, and a final Dense layer with 7 output neurons (SoftMax activation) corresponding to the seven skin lesion classes. The model comprised a total of 21,124,655 trainable parameters.

Table 9:Trained Xception architecture

Layer (type)	Output Shape	Param #
Xception (Functional)	(None, 2048)	20,861,480
Flatten (Flatten)	(None,2048)	0
Dropout (Dropout)	(None,2048)	
dense (Dense)	(None, 128)	262,272
Dropout_1 (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 7)	903

While the original Xception model is designed with a default input size of (299×299×3), we used (224×224×3) in this study to ensure consistency across all models and reduce computational cost. This decision was made intentionally for two reasons:

- ✓ **Efficiency:** Smaller input dimensions lead to faster training and lower GPU memory usage, which is practical for environments like Google Colab.
- ✓ **Comparability:** Using the same input size enables a fair comparison of performance because other models in the study (such as EfficientNet and ResNet) were trained with 224×224 images.

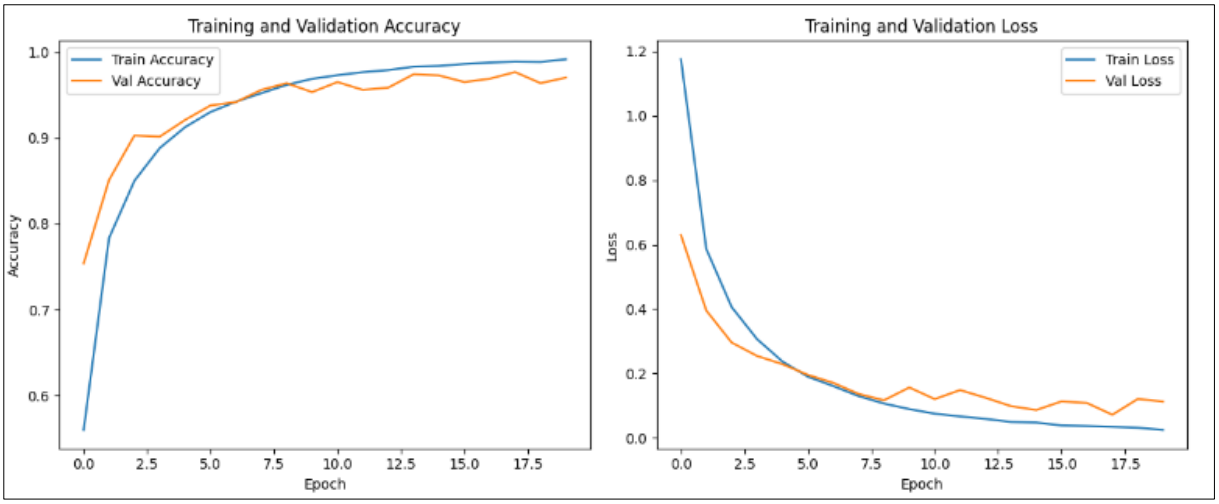

Figure 36:Accuracy and Loss curves (Xception model)

Classification Report:	precision	recall	f1-score	support
akiec	0.99	1.00	1.00	110
bcc	1.00	1.00	1.00	110
bkl	0.97	0.94	0.95	110
df	1.00	1.00	1.00	111
mel	0.94	0.94	0.94	110
nv	0.93	0.96	0.95	110
vasc	1.00	1.00	1.00	111
accuracy			0.98	772
macro avg	0.98	0.98	0.98	772
weighted avg	0.98	0.98	0.98	772

Figure 37:Classification report of Xception Model

- **Confusion Matrix and Model Evaluation:**

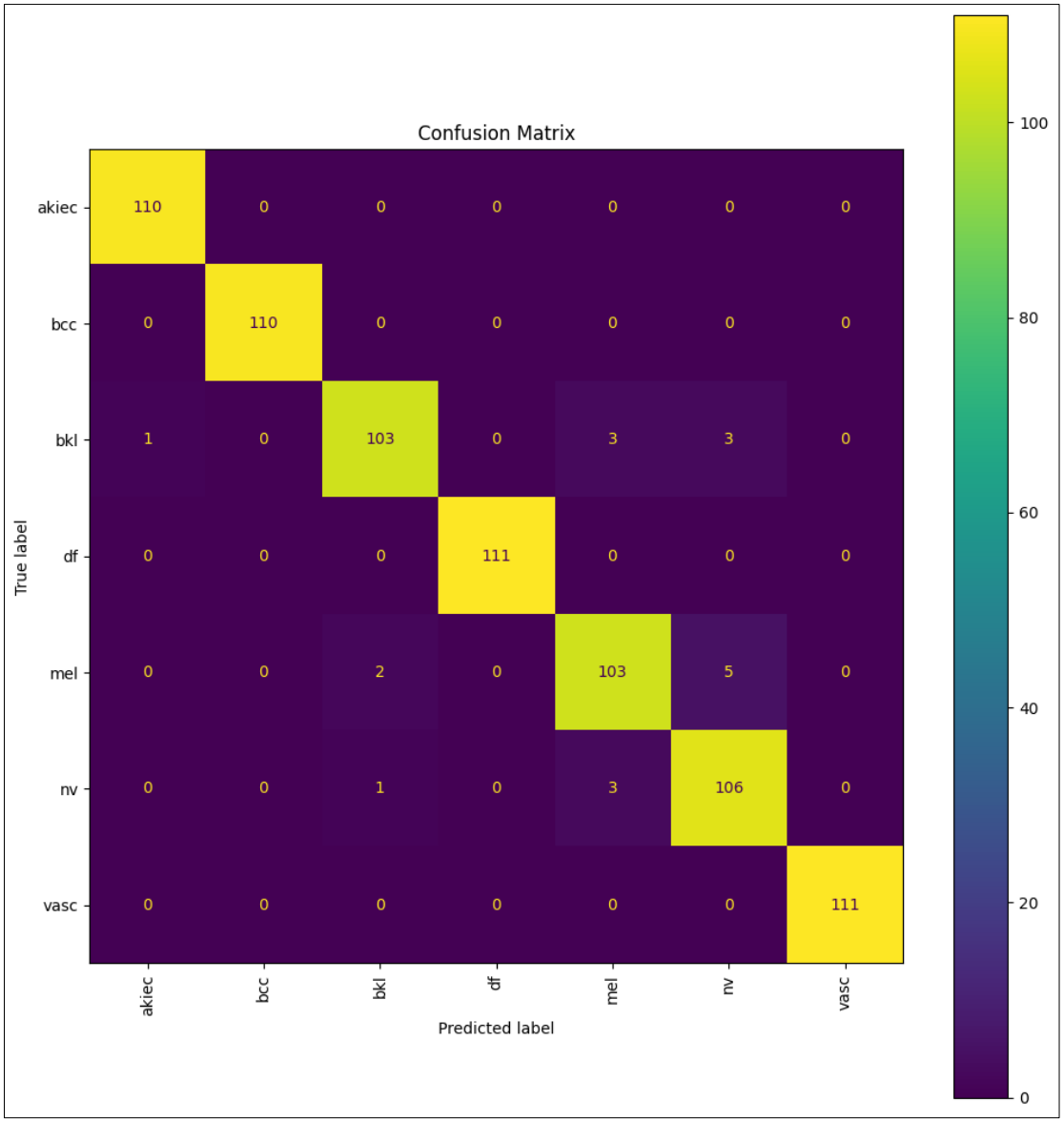

Figure 38:Confusion matrix of the Xception model trained on HAM10000

As demonstrated in the final confusion matrix figure (38), the Xception model excelled in classifying all skin lesion types with very few misclassifications. In particular, the model effectively addressed prior weaknesses seen in other architectures such as ResNet and EfficientNet, especially in distinguishing between BKL (benign keratosis-like lesions) and MEL (melanoma) two challenging classes that often exhibit similar features. This makes Xception the most reliable and high-performing architecture in our study.

- **Potential Applications of the Xception Algorithm in Clinical Practice**

With its many useful uses in clinical practice, the Xception algorithm shows great promise in classification of skin cancer using medical imaging. Improving the precision and effectiveness of cancer diagnosis is one of its most potential applications, especially in situations where early-stage or tiny lesions can be missed by human observers. This can be particularly helpful in environments with limited resources or in places where dermatological specialists are not easily accessible.

Additionally, Xception can help dermatologists and radiologists make decisions by offering a second viewpoint that could increase inter-observer consistency and lower diagnostic mistakes. It may be crucial to early diagnosis and tests because of its strong classification ability, particularly in differentiating between superficially similar lesion types like BKL and MEL (as shown by the confusion matrix results). Its ability to support individual treatment planning is another crucial use. The algorithm may help physicians choose the best treatment options for each patient's diagnosis by precisely identifying specific types of skin cancer. This could enhance treatment results and cut down on unnecessary procedures.

Despite these promising advantages, the integration of AI models like Xception into clinical workflows must be approached with caution. There exists a risk of overreliance on algorithmic outputs, which could diminish human clinical reasoning over time. Furthermore, model performance also needs to be carefully evaluated across a range of patient populations. It is also essential to establish clear protocols for how AI-generated predictions should be interpreted and used in conjunction with clinical judgment.

- **Limitations of the Xception Algorithm**

Although the Xception model has been good at classifying skin cancers, there are certain limitations in the present work that may be addressed in future research. One of the limitations is the requirement of having a lot of labeled data for the model to learn.

For rare skin diseases or underrepresented classes, getting enough high-quality data might be hard, which can detract from the generalizability and resilience of the model. Future work may examine more sophisticated data augmentation strategies or synthetic data generation to counter the effects of small datasets. Another promising direction is the development of model interpretability tools to provide transparent insights into how

the Xception algorithm makes its predictions. This is particularly important in clinical settings, where trust and accountability in AI-driven decisions are critical.

While Xception has proven to be a powerful model in our study, continued research is essential to enhance its accessibility, interpretability, and clinical integration.

4.8 Conclusion

The results of this study demonstrated the effectiveness of various deep learning models in classifying skin lesions using the HAM10000 dataset. ResNet50 leveraged its deep architecture to extract powerful hierarchical features, EfficientNetV2-B1 offered an excellent balance between accuracy and efficiency, while Xception emerged as the most accurate and reliable model, particularly in distinguishing visually similar classes such as BKL and MEL.

Performance was further enhanced through transfer learning, careful tuning of training parameters, and appropriate preprocessing strategies. Dataset splitting and well-structured experiments contributed to improved generalization and reduced overfitting. Based on classification accuracy and confusion matrix analysis, Xception can be considered the most dependable architecture for skin disease classification in this experimental setup.

CHAPTER V
General Conclusion

Skin cancer is considered one of the most serious and widespread health concerns worldwide, with a significant impact on patients' quality of life and survival. The timely and accurate diagnosis of skin cancer is essential for effective treatment and improved outcomes. Deep learning models have shown considerable promise in assisting dermatologists with skin cancer diagnosis in recent years. In this study, we developed and evaluated three state-of-the-art convolutional neural networks Xception, EfficientNetV2-B1, and ResNet-50 for skin lesion classification using the HAM10000 dataset. Our results demonstrated that Xception achieved the highest accuracy (98%), outperforming other models in distinguishing challenging classes like benign keratosis-like lesions (BKL) and melanoma (MEL), while EfficientNetV2-B1 and ResNet-50 also delivered robust performance with 96% and 93% accuracy, respectively. The success of these models, particularly Xception's balance of accuracy and computational efficiency, underscores their potential as reliable decision-support tools for dermatologists, enabling earlier detection and improved patient outcomes in skin cancer diagnosis. In conclusion, this study highlights the potential of deep learning, particularly Xception, as a decision-support tool for dermatologists.

Suggestions for future work Includes:

Future research on the Xception algorithm could explore its application to other cancer types beyond the seven skin diseases studied, particularly those with lower incidence rates or in different medical imaging domains. Integrating multimodal clinical data such as patient history, lifestyle factors, and genetic information may further enhance diagnostic accuracy and support more personalized treatment strategies. Additionally, improving the interpretability of the model could help clinicians better understand and trust AI-based decisions. Investigating the algorithm's deployment in real-world clinical settings, across diverse populations and imaging conditions, is also essential to ensure generalizability. Finally, optimizing Xception for low-resource environments through model compression or edge computing could expand its accessibility and practical utility in global healthcare.

References

[1] Rikiya Yamashita, Mizuho Nishio, Richard Kinh Gian Do, and Kaori Togashi. Convolutional neural networks: an overview and application in radiology. Insights into imaging, 9:611–629, 2018.

[2] G. S. Narayan and M. R. Bendre, "Skin Cancer Identification using CNN: A Lightweight Model with Fine-Grained Classification," 2024 MIT Art, Design and Technology School of Computing International Conference (MITADTSoCiCon), Pune, India, 2024, pp. 1-6, doi: 10.1109/MITADTSoCiCon60330.2024.10575526

[3] Malik SG, Jamil SS, Aziz A, Ullah S, Ullah I, Abohashrh M. High-Precision Skin Disease Diagnosis through Deep Learning on Dermoscopic Images. Bioengineering (Basel). 2024 Aug 27;11(9):867. doi: 10.3390/bioengineering11090867. PMID: 39329609; PMCID: PMC11440112

[4] Timothy McCalmont, MD. Director, UCSF Dermatopathology Service, Professor of Clinical Pathology and Dermatology, Departments of Pathology and Dermatology, Nov 01, 2019. “Melanocytic Nevi” from https://emedicine.medscape.com/article/1058445-overview?form=fpf

[5] Cleveland Clinic “Melanoma”. from https://my.clevelandclinic.org/health/diseases. Last reviewed on 06/21/2021.

[6] Scott R, Oakley A. Benign Keratosis: A Useful Term? Dermatol Pract Concept. 2023 Apr 1;13(2):e2023115. doi: 10.5826/dpc.1302a115. Epub ahead of print. PMID: 37196289; PMCID: PMC10188172

[7] Scurtu LG, Petrica M, Scurtu F, Simionescu AA, Popescu MI and Simionescu O (2024) Basal cell carcinoma—a clinical indicator of immunosuppression. Front. Med. 11:1381492. doi: 10.3389/fmed.2024.1381492

[8] Reinehr CPH, Bakos RM. Actinic keratoses: review of clinical, dermoscopic, and therapeutic aspects. An Bras Dermatol. 2019 Nov-Dec;94(6):637-657. doi: 10.1016/j.abd.2019.10.004. Epub 2019 Nov 6. PMID: 31789244; PMCID: PMC6939186.

[9] Argenziano G, Zalaudek I, Corona R, et al. Vascular Structures in Skin Tumors: A Dermoscopy Study. Arch Dermatol. 2004;140(12):1485–1489. doi:10.1001/archderm.140.12.1485.

[10] Yao MX, Wang YT, Zhou NH. Multiple Eruptive Dermatofibroma: A Case Report. Clin Cosmet Investig Dermatol. 2024;17:457-461 https://doi.org/10.2147/CCID.S424707.

[11] World Health Organization. (2023). Global Report on Skin Diseases. https://www.who.int

[12] Li Z, Fang Y, Chen H, Zhang T, Yin X, Man J, Yang X, Lu M. Spatiotemporal trends of the global burden of melanoma in 204 countries and territories from 1990 to 2019: Results from the 2019 global burden of disease study. Neoplasia. 2022 Jan;24(1):12-21. doi: 10.1016/j.neo.2021.11.013. Epub 2021 Dec 3. PMID: 34872041; PMCID: PMC8649617.

[13] A. Esteva et al., "Dermatologist-level classification of skin cancer with deep neural networks," Nature, vol. 542, no. 7639, pp. 115–118, Feb. 2017, doi: 10.1038/nature21056.

[14] Nischal U, Nischal Kc, Khopkar U. Techniques of skin biopsy and practical considerations. J Cutan Aesthet Surg. 2008 Jul;1(2):107-11. doi: 10.4103/0974-2077.44174. PMID: 20300359; PMCID: PMC2840913.

[15] https://www.mayoclinic.org/tests-procedures/skin-biopsy/about/pac-20384634

[16] Sonthalia S, Yumeen S, Kaliyadan F. Dermoscopy Overview and Extradiagnostic Applications. [Updated 2023 Aug 8]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK537131/

[17] Ayanlowo OO, Adegbulu AA, Cole-Adeife O. Cutaneous cancers in the Africans: Systematic review. Dermatological Reviews. 2022;3:369-383. doi:10.1002/der2.132

[18] Sturm RA. Human pigmentation genes and their response to solar UV radiation. Mutat Res. 1998 Nov 9;422(1):69-76. doi: 10.1016/s0027-5107(98)00176-6. PMID: 9920429.

[19] van Bodegraven B, Vernon S, Eversfield C, Board R, Craig P, Gran S, Harwood CA, Keohane S, Levell NJ, Matin RN, Proby C, Rajan N, Rous B, Ascott A, Millington GWM, Venables ZC; British Association of Dermatologists National Disease Registration Service Steering Committee. 'Get Data Out' Skin: national cancer registry incidence and survival rates for all registered skin tumour groups for 2013-2019 in England. Br J Dermatol. 2023 May 24;188(6):777-784. doi: 10.1093/bjd/ljad033. PMID: 36814132.

[20] Wunderlich K, Suppa M, Gandini S, Lipski J, White JM, Del Marmol V. Risk Factors and Innovations in Risk Assessment for Melanoma, Basal Cell Carcinoma, and Squamous Cell Carcinoma. Cancers (Basel). 2024 Feb 29;16(5):1016. doi: 10.3390/cancers16051016. PMID: 38473375; PMCID: PMC10931186.

[21] Guy GP Jr, Machlin SR, Ekwueme DU, Yabroff KR. Prevalence and costs of skin cancer

treatment in the U.S., 2002- 2006 and 2007-2011. Am J Prev Med. 2015 Feb;48(2):183-187. doi: 10.1016/j.amepre.2014.08.036. Epub 2014 Nov 10. PMID: 25442229; PMCID: PMC4603424.

[22] De Wet, J., Steyn, M., Jordaan, H., Smith, R., Claasens, S., & Visser, W. (2020). An Analysis of Biopsies for Suspected Skin Cancer at a Tertiary Care Dermatology Clinic in the Western Cape Province of South Africa. Journal of Skin Cancer, 2020. https://doi.org/10.1155/2020/9061532.

[23] Urban K, Mehrmal S, Uppal P, Giesey RL, Delost GR. The global burden of skin cancer: A longitudinal analysis from the Global Burden of Disease Study, 1990-2017. JAAD Int. 2021 Jan 4;2:98-108. doi: 10.1016/j.jdin.2020.10.013. PMID: 34409358; PMCID: PMC8362234.

[24] Gordon, L.G., Elliott, T.M., Wright, C.Y. et al. Modelling the healthcare costs of skin cancer in South Africa. BMC Health Serv Res 16, 113 (2016). https://doi.org/10.1186/s12913-016-1364-z

[25] Fontanillas, P., Alipanahi, B., Furlotte, N.A. et al. Disease risk scores for skin cancers. Nat Commun 12, 160 (2021). https://doi.org/10.1038/s41467-020-20246-5

[26] D. L. Narayanan, R. N. Saladi, and J. L. Fox, “Review: Ultraviolet radiation and skin cancer,” International Journal of Dermatology, vol. 49, no. 9, pp. 978–986, 2010, doi: 10.1111/j.1365-4632.2010.04474.x.

[27] Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., Hurst, R. T., Kendall, C. B., Gotway, M. B., & Liang, J. (2016). Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE Transactions on Medical Imaging, 35(5), 1299–1312. https://doi.org/10.1109/TMI.2016.2535302

[28] Spyridonos P, Gaitanis G, Likas A, Bassukas I. Characterizing Malignant Melanoma Clinically Resembling Seborrheic Keratosis Using Deep Knowledge Transfer. Cancers (Basel). 2021 Dec 15;13(24):6300. doi: 10.3390/cancers13246300. PMID: 34944920; PMCID: PMC8699430.

[29] J. Velasco et al., “A Smartphone-Based Skin Disease Classification Using MobileNet CNN,” Int. J. Adv. Trends Comput. Sci. Eng., pp. 2632–2637, Oct. 2019, doi: 10.30534/ijatcse/2019/116852019.

[30] Maduranga, Pasan & Nandasena, Dilshan. (2022). Mobile-Based Skin Disease Diagnosis System Using Convolutional Neural Networks (CNN). International Journal of Image, Graphics and Signal Processing. 14. 47-57. 10.5815/ijigsp.2022.03.05.

[31] Agarwal, Raghav & Godavarthi, Deepthi. (2023). Skin Disease Classification Using CNN

Algorithms. EAI Endorsed Transactions on Pervasive Health and Technology. 10.4108/eetpht.9.4039.

[32] "Disability-adjusted life year," Wikipedia. Mar. 21, 2025. Accessed: May 03, 2025. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Disability adjusted_life_ year&oldid =1281606243

[33] https://ijrpr.com/uploads/V4ISSUE5/IJRPR13532.pdf Adarsh Jadhav1, Shivani Hardade2, Vaishnavi Phadtare3, Adesh Mhetre4, Ms. Rutuja Tikait5 Skin Disease Prediction Using Deep Learning .

[34] Pariwat Ongsulee. Artificial intelligence, machine learning and deep learning. In 2017 15th international conference on ICT and knowledge engineering (ICT&KE), pages 1–6. IEEE, 2017.

[35] School of Computing and Informatics University of Louisiana at Lafayette Lafayette, LA 70503.

[36] A machine learning model for skin disease classification using convolution neural network. Viswanatha Reddy Allugunti DOI: https://doi.org/10.33545/27076636.2022.v3.i1b.53

[37] Shuchi Bhadula, Sachin Sharma, Piyush Juyal, Chitransh Kulshrestha "Machine Learning Algorithms based Skin Disease Detection" (IJITEE).

[38] Bandyopadhyay, S.; Bhaumik, A.; Poddar, S. Skin Disease Detection: Machine Learning vs Deep Learning. https://doi.org/10.20944/preprints202109.0209.v1

[39] COMPUTATIONAL INTELLIGENCE DEEP LEARNING Convolutional Neural Networks horzyk@agh.edu.pl

[40] Franc¸ois Chollet, Xception: Deep Learning with Depthwise Separable Convolutions Google, Inc. fchollet@google.com arXiv:1610.02357v3 [cs.CV] 4 Apr 2017

[41] Omiye JA, Gui H, Daneshjou R, Cai ZR and Muralidharan V (2023) Principles, applications, and future of artificial intelligence in dermatology. Front. Med. 10:1278232. doi: 10.3389/fmed.2023.1278232

[42] https://www.v7labs.com/blog/confusion-matrix-guide

[43] Networks Mingxing Tan, Quoc V. Le, EfficientNet: Rethinking Model Scaling for Convolutional Neural arXiv:1905.11946v5 [cs.LG] 11 Sep 2020

[44] Deep Residual Learning for Image Recognition Kaiming He, Xiangyu Zhang, Shaoqing Ren,

Jian Sun, Microsoft Research arXiv:1512.03385v1 [cs.CV] 10 Dec 2015.

[45] https://developer.nvidia.com/discover/convolutional-neural-network

[46] I. Aziz, "Deep learning: an overview of Convolutional Neural Network (CNN)," April 2020.

[47] Mohaiminul Islam, Guorong Chen, Shangzhu Jin. An Overview of Neural Network. American Journal of Neural Networks and Applications. Vol. 5, No. 1, 2019, pp. 7-11. doi:10.11648/j.ajnna.20190501.12

[48] Tschandl, Philipp, 2018, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions", https://doi.org/10.7910/DVN/DBW86T, Harvard Dataverse, V4, UNF:6:KCZFcBLiFE5ObWcTc2ZBOA== [fileUNF]

[49] Neural networks in the brain involved in memory and recall Edmund T. Rolls and Alessandro Treves*

[50] Olga Russakovsky* “ImageNet Large Scale Visual Recognition Challenge” arXiv:1409.0575v3 [cs.CV] 30 Jan 2015.