PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

UNIVERSITY OF ABOU-BEKR BELKAID

– TLEMCEN –

LMD THESIS

Presented at:

FACULTY OF SCIENCE – DEPARTMENT OF COMPUTER SCIENCE

For Obtaining the Degree of:

DOCTORATE

Specialty: Software Engineering and Applications

By:

Miss MECIFI Youssera Zoukha

On the topic of:

Deep Learning-Based Automatic Diagnosis of Diseases

Publicly defended in February 14, 2026 at Tlemcen University before the jury composed of:

Dr BENAMAR Abdelkrim	Full Professor	University of Tlemcen	President
Dr MERZOUG Mohamed	Associate Professor	University of Tlemcen	Supervisor
Dr HADJILA Fethallah	Associate Professor	University of Tlemcen	Co-supervisor
Dr HADJILA Mourad	Associate Professor	University of Tlemcen	Examiner
Dr HACHEMI Mohammed Hicham	Associate Professor	University of Sciences and Technology of Oran	Examiner
Dr NEDJAR Imane	Associate Professor	Higher School of Applied Sciences Tlemcen	Examiner

Acknowledgements

Praise be to Allah, who gave me the patience, courage and health to complete this thesis despite the difficulties, discouragements encountered and for helping me all my life to get to this point and for never letting me down. I would first like to thank Dr. MERZOUG Mohammed and Dr. HADJILA Fethallah, my thesis supervisors, and Dr M'hamedi Mohamed, for their invaluable guidance and time.

I also thank the members of the jury who accepted the evaluation of our work, including Pr BENAMAR Abdelkrim, Dr. HADJILA Mourad, Dr. HACHEMI Mohammed Hicham, and Dr. NEDJAR Imane. Also, big thanks to my family for their support, attentiveness, and encouragement throughout this thesis and all my colleagues and friends that helped with a small advice, prayer, or even a sweet word.

Finally, I would like to thank all the people who participated and helped directly or indirectly in the realization of this modest work. And to all my teachers in primary, middle, and secondary school, and also my university professors who passed on their knowledge to me to become what I am today.

"GOD BLESS YOU"

I dedicate this modest work:

To my dearest parents, who raised me to be the woman I am today. Without their love, sacrifice, and unwavering support, I would not be where I am. I dedicate not only this thesis to them, but also my time and my life, hoping that I can one day repay even a part of what they have given me; though it could never truly be enough. May Almighty Allah protect and bless them always.

To my sister, and my aunts, who have loved me unconditionally since childhood and have always cheered for my success, and to my little brother and nephews who I wish nothing but success for them.

To my beloved man, who stood by me through thick and thin, helped shape me into a stronger, better version of myself, and supported me with everything he had. Thank you for never giving up on me, for believing in me, for your patience, and for making my success feel like yours too.

To my best friends, family members and neighbors who wished me luck from the bottom of their hearts and were happy to see me succeed. thank you for your kind hearts and encouragements.

Abstract

Skin cancer continues to be one of the most serious dermatological threats, with melanoma as its most aggressive form due to its high metastatic potential, and it remains a significant emotional and financial burden on patients and healthcare practitioners. That's why early diagnosis is important for improving patient outcomes and even saving their lives, but established methods of diagnosis such as visual inspection and dermoscopy are limited by human subjectivity and access to expert dermatologists. Recent advances in deep learning and artificial intelligence have revealed new fronts in medical diagnostics, especially in automatizing the processing of dermoscopic images. There, convolutional neural networks (CNNs) have matched or surpassed clinicians' diagnostic capacity in controlled environments.

This thesis explores the possibility of using deep learning techniques for improving the detection of early skin cancer with dermoscopic imaging. Specifically, it delves into how CNN architectures pre-trained on large datasets such as VGG16 and MobileNetV2 can be deployed for binary classification tasks, and how their performance can be enhanced through the use of synthetic images generated via DCGANs. It also delves into a crucial yet less explored aspect in skin cancer diagnosis: the influence of lighting modalities employed for image capture and how this influences the generalization of models. For this purpose, classification performance under various lighting conditions is measured by using Accuracy, Specificity, recall, and AUC as primary metrics.

Results indicate that dataset balancing with GAN-generated malignant images significantly enhances model resilience, and that there is a slight difference between MobileNetV2 and VGG16 performance across all test scenarios. Moreover, the work provides evidence that lighting modality influences model training and inference quality, which has crucial implications for practical applications. These results contribute to the existing body of literature on AI-driven dermatological diagnosis and offer pragmatic recommendations for the development of more generalizable and clinically pertinent deep models.

Keywords: Skin Cancer, DCGAN, VGG16, MobileNetV2, Non-polarized White Light, Cross Polarized Light, CNN, Transfer Learning, Data Augmentation.

Résumé

Le cancer de la peau demeure l'une des menaces dermatologiques les plus sérieuses, le mélanome étant sa forme la plus agressive en raison de son fort potentiel métastatique. Il constitue toujours un fardeau émotionnel et financier important pour les patients ainsi que pour les professionnels de santé. C'est pourquoi un diagnostic précoce est essentiel pour améliorer le pronostic des patients, voire leur sauver la vie. Cependant, les méthodes de diagnostic classiques, comme l'inspection visuelle et la dermoscopie, restent limitées par la subjectivité humaine et le manque d'accès à des dermatologues spécialisés. Les récentes avancées en apprentissage profond et en intelligence artificielle ont ouvert de nouvelles perspectives en diagnostic médical, notamment en automatisant le traitement des images dermoscopiques. Dans ce contexte, les réseaux de neurones convolutifs (CNN) ont démontré une capacité de diagnostic équivalente, voire supérieure, à celle des cliniciens dans des environnements contrôlés.

Cette thèse explore l'utilisation des techniques d'apprentissage profond pour améliorer la détection précoce du cancer de la peau à partir d'images dermoscopiques. Plus précisément, il examine comment des architectures CNN pré-entraînées sur de grands ensembles de données, telles que VGG16 et MobileNetV2, peuvent être adaptées à des tâches de classification binaire, et comment leur performance peut être optimisée à l'aide d'images synthétiques générées par des réseaux antagonistes génératifs convolutifs profonds (DCGANs). Le travail s'intéresse également à un aspect essentiel mais peu exploré dans le diagnostic du cancer de la peau : l'influence des modalités d'éclairage utilisées lors de l'acquisition des images, et son impact sur la capacité de généralisation des modèles. À cette fin, les performances de classification sous différentes conditions d'éclairage ont été évaluées à l'aide de métriques telles que l'exactitude, la sensitivité, la précision, le rappel et l'AUC.

Les résultats montrent que l'équilibrage des ensembles de données grâce à des images malignes synthétisées par GAN améliore considérablement la robustesse des modèles, et que MobileNetV2 surpasse légèrement VGG16 dans la majorité des scénarios de test. En outre, le travail apporte la preuve que les modalités d'éclairage influencent la qualité de l'apprentissage et de l'inférence des modèles, ce qui a des implications majeures pour leur déploiement pratique. Ces conclusions contribuent à la littérature existante sur le diagnostic dermatologique assisté par intelligence artificielle, et proposent des recommandations pragmatiques pour développer des modèles profonds plus généralisables et cliniquement pertinents.

Mots-clés : Cancer de la peau, DCGAN, VGG16, MobileNetV2, Lumière blanche non polarisée, Lumière polarisée croisée, CNN, Apprentissage par transfert, Augmentation de Données.

الملخص

يُعدّ سرطان الجلد من أخطر الأمراض الجلدية، ويُعتبر الميلانوما أكثر أنواعه عدوانية بسبب قدرته العالية على الانتشار إلى أعضاء الجسم المختلفة، مما يجعله عبئًا صحيًا وعاطفيًا وماليًا كبيرًا على المرضى والمجتمع. إن التشخيص المبكر يلعب دورًا حاسمًا في تحسين نتائج العلاج وإنقاذ الأرواح، غير أن الطرق التقليدية مثل الفحص البصري واستخدام المجهر البصري تعتمد بشكل كبير على خبرة الطبيب، مما يجعلها عرضة للأخطاء البشرية. في هذا السياق، أظهر التقدم في مجال الذكاء الاصطناعي والتعلم العميق، وخصوصًا الشبكات العصبية الالتفافية، إمكانيات واعدة في دعم تشخيص سرطان الجلد من خلال تحليل الصور الطبية بدقة وسرعة تفوق أحيانًا أداء الأطباء في بيئات خاضعة للرقابة

تهدف هذه الأطروحة إلى تحسين دقة تصنيف سرطان الجلد اعتمادًا على صور الجلد المكبرة باستخدام نماذج تعلم عميق مدعومة بتقنيات حديثة مثل التعلم بالنقل وتوليد البيانات الاصطناعية. تم استخدام نماذج مسبقة التدريب تعتمد على الشبكات العصبية الالتفافية لتصنيف الصور بشكل ثنائي (حميد أو خبيث)، كما تم الاعتماد على الشبكات التوليدية التنافسية للمساعدة في موازنة بيانات التدريب، خصوصًا بسبب عدم توازن عدد العينات بين الفئتين. بالإضافة إلى ذلك، تم دراسة تأثير ظروف الإضاءة المختلفة أثناء التقاط الصور على أداء النموذج ومدى تعميمه

أظهرت النتائج أن استخدام الصور الاصطناعية ساهم في تعزيز أداء النماذج المدربة، وأن نوع الإضاءة المستخدم في الصور يلعب دورًا ملحوظا في دقة التصنيف. كما تم تقييم النماذج باستخدام مؤشرات دقيقة مثل الدقّة، والحساسيّة، والاستدعاء، والمساحة تحت المنحنى ,وتوصّلت الأطروحة إلى أن نموذج موبايل نت يتفوق بشكل طفيف على نموذج في جي جي في مختلف سيناريوهات الاختبار. تسهم هذه الدراسة في تطوير أدوات تشخيص ذكية تساعد الأطباء وتُعزز فرص الكشف المبكر عن سرطان الجلد، مما يُحسّن فرص العلاج ويُقلل من معدلات الوفاة

الكلمات المفتاحية: سرطان الجلد، الشبكات التوليدية التنافسية العميقة، في جي جي ١٦، موبايل نت الإصدار الثاني، الضوء الأبيض غير المستقطب، الضوء المستقطب المتقاطع، الشبكات العصبية الالتفافية، التعلم بالنقل، تعزيز البيانات.

Table of Content

Table of Figures

List of Tables

General introduction

Context

Although skin cancer is not currently ranked among the top three causes of mortality worldwide according to the World Health Organization [1], it still represents a significant public health concern. Globally, Melanoma skin cancer ranks 17th in prevalence, with Europe recording the highest incidence rates and mortality figures [2]. Each year, it leads to profound emotional and economic impacts on affected families. The American Cancer Society projects that approximately 112,000 new melanoma cases will be diagnosed in the United States in 2026 [3], with an estimated 8,510 melanoma-related deaths (American Cancer Society, 2026)., highlighting the persistent burden of this disease. Skin cancer has multiple types, and among its different types, melanoma remains the most aggressive and potentially the deadliest if not detected and treated early. Although less common than other skin cancers such as basal cell carcinoma and squamous cell carcinoma, melanoma is particularly dangerous due to its ability to spread (metastasize) quickly to other parts of the body. When melanoma spreads beyond the skin and affects organs such as the lungs, liver, or brain, it becomes much more difficult to treat and can be life-threatening.

Researchers and clinicians have continuously pursued new strategies to combat this insidious disease. With the rising prevalence of skin cancer, significant efforts have been devoted to improving early diagnosis and enhancing treatment methods. Although, traditional diagnostic methods, including visual inspection and dermoscopy by skilled dermatologists, are labor-intensive and prone to human errors. In areas of lower specialist availability, misdiagnosis or delayed diagnosis leads to considerably worse outcomes. Early treatment of skin cancer always gives patients their greatest chance for full recovery, and as such any delayed diagnosis can be extremely serious and have serious repercussions on their general health. Skin cancer that has been misdiagnosed can develop for weeks or months before it is diagnosed. This gives the skin cancer time to develop and travel throughout the body of the patient. It could move to more serious stages, and where one might have been successfully treated for stage 1 cancer, suddenly they are struggling with much more aggressive form which requires more invasive cancer treatment. Worst-case in late or missed diagnosis is wrongful death, or death that could have been delayed. According to reports by Cancer Research UK, when it comes to stage 1 skin cancer, almost 100% of patients can survive their cancer for 5 years or more from the time they are diagnosed. In stage 3, this percentage drops to 70% [4].

Thanks to technology, life become much easier through-out the years, and we can notice that in many fields; whether its education, finance, sports, or our daily life; but most importantly, and the most impactful filed is the integration of technology in medical fields. In recent years

the most popular domain we hear most about is AI (Artificial intelligence), even our elders and children have heard about its impact and not only people of the filed, artificial intelligence and deep learning have emerged as powerful tools to support medical diagnosis, offering the potential to analyze large volumes of medical images rapidly and consistently. Specifically, convolutional neural networks (CNNs) have shown promise in the automatic classification of skin lesions, often performing on par with, or even exceeding, human experts in certain controlled condition [5]. That's why we hear often the terms of Compute Aided Dignosis (CAD) and Automated computer diagnosis in medicals fields. In brief, although there are some commonalities between CAD and automated computer diagnosis, there are also some discrepancies between them. Automated computer diagnosis is a computer-algorithm-based concept, whereas CAD is a concept that has emerged by equally taking into consideration both the roles of physicians and computers. Because of this fundamental difference, the requirement for computer algorithm performance has been a great difference between the two. In computer-automated diagnosis, computer performance needs to be equal to or better than that of physicians. In CAD, however, computer performance does not need to be equal to or better than that of physicians, but complementary to physicians. Actually, many CAD systems in Europe and the United States have been utilized for helping doctors in the early detection of breast cancers on mammograms.

Problem Specification

Healthcare AI has made it possible to process and analyze vast amounts of medical data much better than human beings. It enabled diagnosing sicknesses, predicting outcomes, and suggesting cures. For instance, AI programs can interpret medical scans, such as X-rays and MRIs, more accurately and faster than human radiologists and even detect diseases like cancer at more developed stages. Uses of artificial intelligence in medicine are widespread and profound. Another such breakthrough besides IBM Watson Health was Google's DeepMind Health venture that demonstrated it could identify eye disease from retinal scans with a level of accuracy equal to that of human experts. Such pioneering efforts provided insight into AI's potential to revolutionize diagnostics and precision medicine. AI in medicine's capability to rapidly review huge volumes of clinical records enables healthcare practitioners to discover markers and patterns of disease that might go undetected otherwise. The applications of healthcare and AI are wide-ranging and vast, ranging from screening radiological images for early detection to forecasting outcomes based on electronic health records. By integrating artificial intelligence into hospitals and clinics, healthcare systems can make their systems smarter, faster, and more effective in treating millions of people worldwide [6]. In skin cancer, factors like skin tone, lighting conditions, lesion size, and anatomical location can impact diagnosis. This subjectivity has led researchers to explore artificial intelligence (AI) as a means to standardize and improve early detection efforts. In this context, dermoscopy has emerged as a key imaging technique. It offers a non-invasive magnified view of skin lesions, revealing subsurface structures that are otherwise invisible to the naked eye. The quality and modality of

these images significantly affect classification performance. Therefore, developing robust AI models that generalize across different states and anatomical variances has become an active area of research.

The advent of deep convolutional neural networks (CNNs) has revolutionized image-based classification tasks, including in medical imaging. Architectures such as AlexNet, VGG16, ResNet, and MobileNet have been successfully repurposed to classify skin lesions, sometimes matching or exceeding the diagnostic performance of experienced dermatologists [3][4].One of the most impactful contributions came from Esteva et al. [5], who demonstrated that a CNN trained on over 129,000 clinical images could achieve dermatologist-level classification accuracy for melanoma and other skin conditions [5]. Since then, the field has rapidly progressed, with methods incorporating transfer learning, data augmentation, and generative adversarial networks (GANs) to overcome several challenges.

However, the effectiveness of these AI-based systems heavily depends on the quality, diversity, and balance of the datasets used for training. Medical datasets often suffer from class imbalance, where benign cases significantly outnumber malignant ones. This imbalance can bias models and reduce their diagnostic reliability, especially for detecting rare or aggressive cancers.

Therefore, improving the quality of datasets and the training strategies of deep learning models remains a crucial challenge. Techniques such as transfer learning and data augmentation; particularly using Generative Adversarial Networks (GANs); are now being explored to overcome these limitations and enhance diagnostic accuracy in AI-driven skin cancer detection. In addition, variations in image acquisition conditions, such as lighting, further puts a wide area for questioning and wondering whether any changes in this artifact can affect the model's performance, while this question seems very obvious and the answer might also seem simple, it is widely under explored in the literature.

These limitations underscore the urgent need for more refined data preprocessing, augmentation, and model training strategies tailored to the particular challenges of skin lesion analysis in dermoscopy.

Contributions

The overarching goal of this thesis is to enhance the accuracy and robustness of skin cancer diagnosis through the application of deep learning techniques on dermoscopic images plus studying the effects of lighting modality while capturing the images, on the overall training performance. Given the high stakes involved in early detection, improving classification performance can lead to significant clinical benefits by assisting dermatologists in early intervention. To achieve this objective, we sought to address a range of specific challenges encountered in the field:

- The extensive class imbalance in medical image datasets with benign images far exceeding malignant images.
- The limited sample size and heterogeneity of existing annotated data sets.
- The effect of imaging conditions (lighting modalities) on model performance.
- The need to establish whether transfer learning structures (like VGG16 and MobileNetV2) would be able to generalize well to skin lesion classification problems.
- The necessity of quantitative evaluation based on comprehensive metrics like accuracy, precision, recall, and AUC.

Accordingly, the work that has been conducted in this thesis is founded on the following main research questions:

- Are synthetic data from Deep Convolutional GANs able to improve CNN model classification accuracy on unbalanced skin lesion datasets?
- How do the fine-tuned pre-trained models (e.g., VGG16 and MobileNetV2) compare with each other when applied for dermoscopic image binary classification?
- What effect have image acquisition modalities (e.g., lighting conditions) on learning and generalization of classification models?
- Which model architecture yields the optimal trade-off between accuracy and generalization for skin cancer detection?

To address these questions, two main research axes were developed:

1. **Synthetic Data Generation and Transfer Learning**: In this part, DCGAN-generated malignant skin lesion samples were integrated with real dermoscopic images to mitigate class imbalance. Transfer learning models (VGG16 and MobileNetV2), enhanced with tailored modifications, were then trained and evaluated.
    - This work was concretized in a paper published in the Cybernetics and Information Technologies journal [7].
2. **Evaluation of Lighting Modality on Model Performance**: A comparative study was conducted using ResNet50, EfficientNet-B0, MobileNetV2, and DenseNet121 across dermoscopic images captured under different lighting conditions (white light, cross-polarized, and mixed). The aim was to quantify the effect of acquisition modality on model robustness.
    - The results were disseminated in a submitted paper which is still under review.

By systematically exploring these two axes, the thesis contributes to the literature on AI-based skin cancer detection by highlighting (i) the role of synthetic data augmentation using DCGAN in addressing dataset imbalance, and (ii) the importance of acquisition conditions in ensuring reliable model generalization. The outcomes also emphasize reproducibility, rigorous performance evaluation, and clinical applicability.

MANUSCRIPT ORGANIZATION

The present thesis manuscript is organized into four chapters, each addressing a critical component of the research undertaken in the domain of skin cancer classification using deep learning. The first chapter, entitled *Artificial intelligence, Machine Learning, and Deep Learning*, presents the fundamental concepts of artificial intelligence and machine learning, to lay the foundation for understanding these technologies. This chapter then introduces convolutional neural networks (CNNs) and their essential role in computer vision, a subdiscipline of AI dedicated to analyzing and interpreting visual data. Emphasis is placed on transfer learning techniques, which enable the reuse of pre-trained CNN models to accomplish specific tasks, thus reducing the need for massive data sets to train effective models. Finally, a review of key pre-trained models, such as VGG16 and MobileNetV2, is presented, highlighting their particular relevance to medical image classification, where accuracy and reliability are crucial.
The second chapter, *Background and general information on skin cancer*, contains an in-depth description of skin cancer types to contextualize the importance of early detection and diagnosis. This chapter then focuses on malignant melanoma, examining its clinical and pathological features that distinguish it from other skin cancer types. Subsequently, a prevention and diagnostic methods are presented along with detailed presentation of imaging techniques used for data acquisition plus, a number of popular datasets of skin cancer were listed. Finally, the use of artificial intelligence for the diagnosis of skin cancer is discussed, highlighting the significant advances and persistent challenges in the application of these technologies to improve the diagnosis and treatment of this form of cancer.

The third chapter, *Early skin cancer detection using transfer learning model and DCGAN approach for data balancing*, focuses on the dataset curation process and the handling of class imbalance using Deep Convolutional Generative Adversarial Networks (DCGANs). It describes the procedures followed to refine the image dataset and details the design and training of generative models aimed at augmenting the minority class with high-quality synthetic images, along with the details of implementing and adapting two pre-trained architectures, VGG16 and MobileNetV2, for binary classification of skin cancer. This section describes model modifications, training settings, and optimization strategies used to enhance generalization and stability, particularly under data-limited conditions.

Chapter four, entitled *Effect of Dermoscopic Lighting Modality on Skin Cancer Classification: A Deep Learning-based Approach*, presents a thorough assessment of the classification models across a series of metrics including Accuracy, Recall, Specificity, and Area Under the Curve (AUC) to evaluate the effect of lighting modality changes of the input images meant for training. Comparative analyses, confusion matrices, and visual plots are provided to support interpretation of the results and to demonstrate the models' diagnostic potential.

Finally, we present a *General Conclusion*, that summarizes the contributions of the thesis, reflects on the limitations encountered, and proposes future research directions such as multi-modal data fusion, advanced hybrid architectures, and potential clinical validation studies.

1. Chapter 1. Background and general information on skin cancer

1.1. Introduction

According to the American Cancer Society's Cancer Facts & Figures 2026 report, an estimated 2,114,850 new cancer cases and 626,140 cancer deaths are expected to occur in the United States in 2026. Despite a continued decline in overall cancer mortality (attributable to reductions in smoking, earlier detection, and improved treatments) significant gaps persist across cancer types and population groups. The cancer mortality rate has dropped substantially since its peak in 1991, contributing to millions of deaths averted, yet incidence continues to rise for many common cancers [8].

According to statistics from the World Health Organization, cancer is nowadays considered one of the leading causes of human death [9], as statistical models predict that cancer incidence and mortality rates will escalate to approximately 26 million new cases and 17 million deaths per annum by 2030 [10]. Among all types of cancer, skin cancer is the most common form in of cancer worldwide [11], it is not necessarily fatal, however, early diagnosis plays a crucial role in saving lives. To understand the early detection and diagnosis of skin cancer, it is important to examine the human skin and the different types of skin cancers.

The most dangerous skin cancer is melanoma, which is more likely to spread and become deadly if not found early. Basal cell carcinoma (BCC) and squamous cell carcinoma (SCC) are more prevalent, but the severity and lethality do not compare to that of melanoma, in more details:

- Melanoma is the deadliest form of skin cancer because it is most likely to spread to other areas of the body.
- Basal Cell Carcinoma (BCC) is quite common [12], but it normally grows slowly and almost never spreads, so it is less deadly than melanoma.
- Squamous cell skin cancer (SCC) is more likely to spread than BCC, but it is also usually easier to treat than melanoma.
- Merkel Cell Carcinoma is a rare, but aggressive form of skin cancer and can be fatal if it spreads.

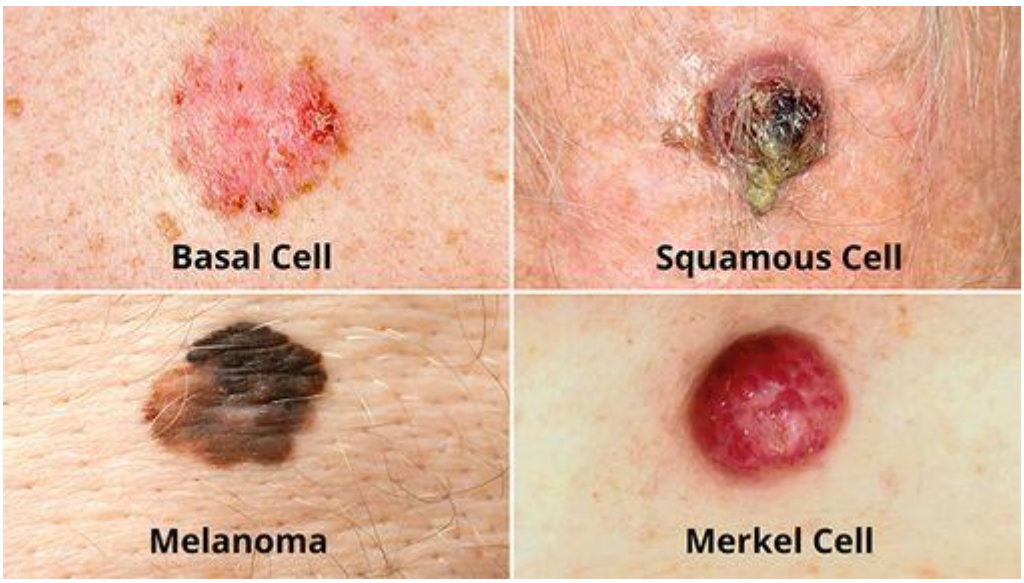


Figure 1-1: Some examples of dangerous skin cancers

These feared types of cancer represent the most serious forms of skin cancer, possible risk factors include exposure to sunlight or UV rays, viral infections, arsenic exposure, immunosuppression, and certain genetic conditions. In the early days of melanoma treatment, surgery was the only available method; over time, sentinel lymph node biopsy, wide excision, immune checkpoint inhibitors, targeted therapy drugs, and chemotherapy have been used to treat deadly skin cancers.

In modern medicine, relying solely on traditional methods or approaches to cure or mitigate the severity of cancer is considered inadequate. Thankfully, due to technological advances in the medical field, physicians and medical experts now use various medical equipment to quickly diagnose skin cancer and prevent it from progressing.

This chapter is divided into three parts: the first describes skin lesions, the second explains the different types of skin cancer, and the third part focuses on diagnostic techniques, computer-aided diagnosis systems, and artificial intelligence.

1.2. Presentation of Cutaneous Lesions

A cutaneous lesion refers to an area of skin exhibiting abnormal growth that differs in appearance from surrounding tissue, particularly in color, shape, size, or texture. Such lesions are highly prevalent and often result from localized skin damage (e.g., sunburns, contact dermatitis). However, some may indicate systemic conditions, including infections, diabetes, or autoimmune/genetic disorders. While most cutaneous lesions are benign, others may be malignant or precancerous, with potential to progress into skin cancer.

Lesions are classified as follow [13]:

1. **Primary Lesions**: Arise on previously healthy skin due to specific etiologies. Examples include freckles, nevi (moles), and blisters, which may be congenital or acquired.
2. **Secondary Lesions**: Develop from the progression or manipulation of primary lesions (e.g., scratching, friction, treatment). Common types include scabs, ulcers, erosions, and scars.

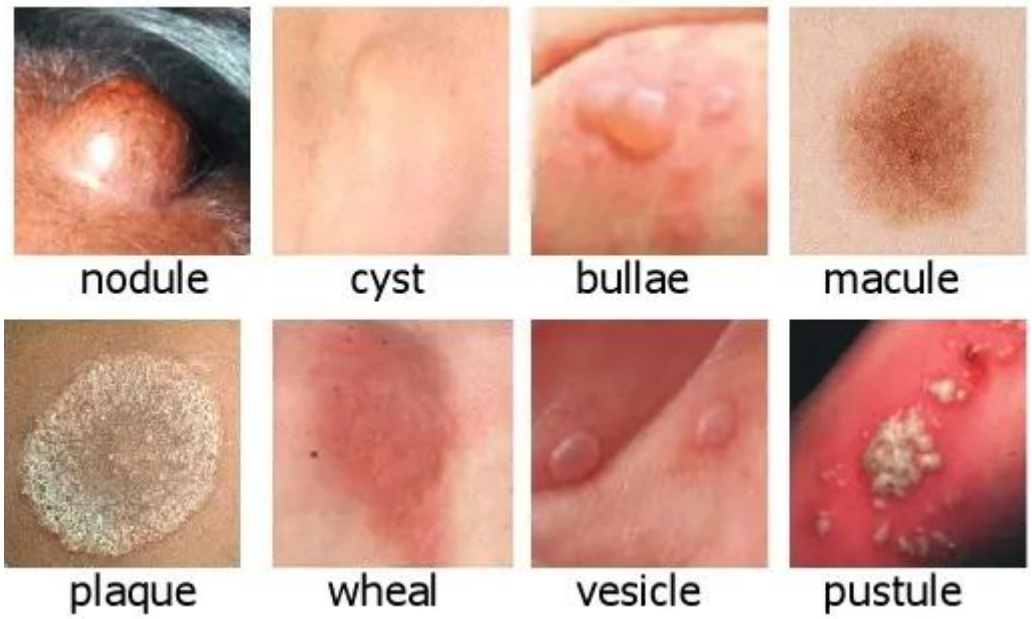


Figure 1-2: Primary lesion Examples.

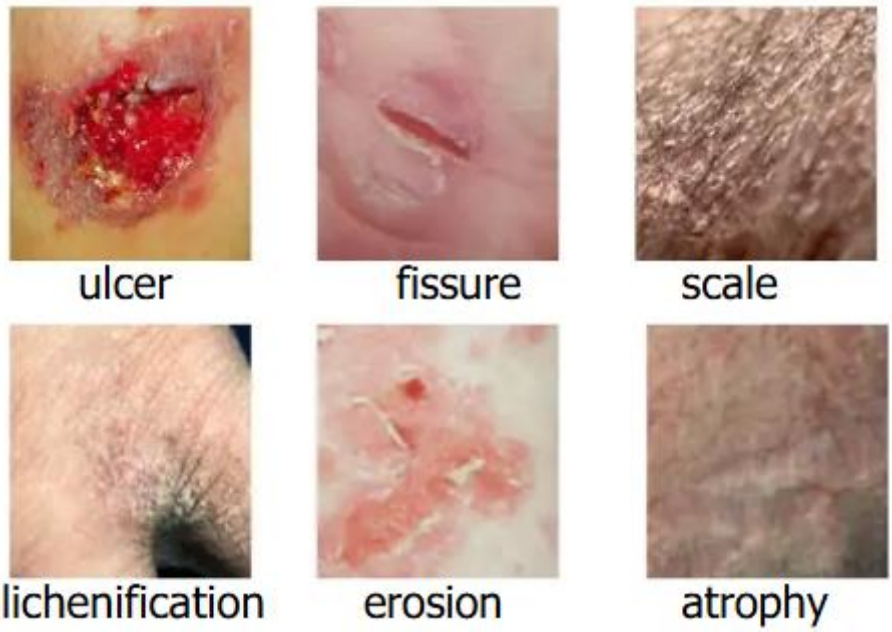


Figure 1-3: Secondary lesion Examples.

Cutaneous lesions vary in morphology (size, shape, texture) and distribution (isolated, clustered, localized, or generalized). Key risk factors include:

- **Congenital nevi**: Present in 1–2% of newborns, with ~5% risk of malignant transformation to melanoma.
- **Atypical/dysplastic nevi**: Associated with higher melanoma risk. The associated moles are typically large, irregularly shaped, and range in color from pink to dark brown, representing atypical or dysplastic nevi.

A benign skin lesion is a non-cancerous skin abnormality, growth, or tumor that can appear anywhere on the body, these lesions manifest differently depending on their cause and tissue of origin. Common benign lesions include most melanocytic nevi (commonly called moles), seborrheic keratoses, skin tags, cherry angiomas, and lipomas, most are harmless and require no treatment unless causing discomfort or itching [14]. Unlike malignant lesions, benign ones are typically symmetrical, well-circumscribed, uniform in appearance, and stable or slowly evolving. However, when differentiation from malignant lesions is challenging, a biopsy or surgical excision may be performed to rule out malignancy.

1.3. Skin Cancer

Skin cancers are primarily classified by their cell of origin: carcinoma (non-melanoma) and melanoma [15].

1.3.1. Non-Melanoma Skin Cancer (NMSC)

The term *non-melanoma skin cancer* is commonly used to refer to basal cell carcinomas (BCC) and squamous cell carcinomas (SCC). It therefore identifies a population of cancer sufferers by what they do not have, cutting an area out of the scope of study and downplays the importance of this disease.

Basal cell cancer (BCC) is a kind of skin cancer, which is also known as a rodent ulcer, and it's the least serious form of skin cancer. It is also the most common malignant neoplasm in humans [16]. This cancer grows from the basal cells at the base of the epidermis (the outer layer of skin). Unusual growth in this layer is the characteristic of a BCC.

BBC usually appears in areas of chronic sun exposure such as the face, ears, neck, scalp, and shoulders, but can also develop on the back or legs. BCC typically grows slowly and very rarely spreads, although it can cause significant local tissue destruction and disfigurement if neglected or improperly treated; it is generally not life-threatening especially when prognosis is done in the right moment [17].

In addition, non-melanoma skin cancer also includes other malignant skin tumors such as cutaneous lymphoma, Merkel cell carcinoma, Paget's disease, angiosarcomas, and malignant histiocytomas [18]. Squamous cell carcinoma (SCC) is the second most common form of skin

cancer, accounting for approximately 20% (1 in 5) of diagnosed skin cancers [19]. This cancer begins in cells called keratinocytes, located in the upper part of the epidermis. SCC most often develops in sun-exposed areas, but it can also develop in scars, previously burned skin areas, and long-term ulcerated skin regions. In rare cases, SCC can develop in genital areas. Moreover, chronic skin inflammation or medical conditions that suppress the immune system over time can promote SCC development.

SCC is more aggressive than BCC because it grows faster, has less defined margins, and a greater c potential. While 96–97% of SCCs are localized, the same percentage of remaining cases (3–4%) may spread to other parts of the body, often with fatal outcomes [20]. Due to the very large number of cases recorded each year, non-melanoma skin cancers represent a major public health issue.

1.3.2. Risk Factors for Non-Melanoma Skin Cancer

Multiple interconnected risk factors are associated with the development of non-melanoma skin cancers. Both genetic and environmental causes are involved in the multistep carcinogenesis process, including tumor initiation, pre-malignant progression, and malignant transformation of normal skin cells into cancerous lesions. Among the risk factors, we mention [21]:

- **Ultraviolet radiation, latency, and dose effects**: Sunlight is the most significant external cause of non-melanoma skin cancers. Ultraviolet rays B (UVB), A (UVA), and C (UVC) can cause DNA damage in normal skin cells. While most studies show that UVB is the main part of the solar spectrum responsible for the development of these cancers, UVA is also a risk factor. UVC is absorbed by the ozone layer and does not reach the Earth's surface.
- **Skin Type**: According to a meta-analysis, individuals with skin phototypes I or II have a 70% higher risk of developing basal cell carcinoma (BCC) compared to those with skin phototypes III or IV [22]. The following table lists the skin phototypes and their features:

Table 1-1: Overview of Skin Phototypes and Their Typical Features and Sun Response. Based on: [23]

Skin Phototype	Typical Features	Tanning Ability
Type I	Tends to have freckles, red or fair hair, and blue or green eyes.	Often burns, rarely tans.
Type II	Tends to have light hair, and blue or brown eyes	Usually burns, sometimes tans.
Type III	Tends to have brown hair and eyes.	Sometimes burns, usually tans.

Type IV	Tends to have dark brown eyes and hair.	Rarely burns, often tans.
Type V	Naturally black-brown skin. Often has dark brown eyes and hair.	
Type VI	Naturally black-brown skin. Usually has black-brown eyes and hair.	

- **Chemicals and tobacco**: Exposure to ionizing radiation—whether iatrogenic (e.g., acne or scalp ringworm treatment), professional (e.g., uranium mining), or accidental (e.g., atomic bomb explosion [24])—increases the risk of BCC.

  It has been reported that smoking was linked to a higher risk of squamous cell carcinoma (SCC) but appeared to lower the risk of basal cell carcinoma and melanoma. However, the exact biological reasons behind these contrasting effects remain unclear [25].

- **Elderly individuals**: There is a clear relationship between immune system function, aging, and the incidence of non-melanoma skin cancers. In the study of Sinikumpu et al. [26], they found that 25.5% of older adults had skin cancer or its precursors, with higher prevalence in men. A history of previous skin cancer significantly increased the risk of developing new cancers, and outdoor work was identified as a key risk factor. Age and socioeconomic status also influenced skin cancer occurrence, emphasizing the importance of regular skin examinations in this population.

1.3.3. Diagnosis of Non-Melanoma Skin Cancers

Several factors influence the progression of a disease, the simplest way to assess the effect of health in a particular population is to estimate morbidity (number of disease cases) and associated mortality. There are more complex health indicators, such as discomfort, disability, and dissatisfaction (quality of life).

1.3.4. Factors Related to Non-Melanoma Skin Cancers

These factors include the type of cancer coverage and the tumor's risk profile. For SCC, higher mortality is observed in patients with tumors located on the lips and ears, as well as in anogenital areas. Other high-risk precursors include:

- Large tumor size (greater than 2 cm)
- Poor to moderate differentiation
- Deep invasion

In the case of BCC, tumor aggressiveness also depends on location. Tumors of the head and neck (particularly those on the ear and eyelid) are more frequently metastatic.

❖ ***Treatment Factors***

In the management of patients with non-melanoma skin cancer, treatment factors must be considered. These can be influenced by:

- ❖ The treating clinician's performance
- ❖ The patient's acceptance of the treatment plan

These tumors are considered less severe than others due to their slow growth and low metastasis rates. However, clinicians observe substantial rates of treatment refusal for surgical interventions among patients, which is often associated with advanced age, frailty, and socioeconomic disadvantages, including racial minority status and lower income [27]; this can result in increased morbidity and, in the worst cases, mortality .

❖ ***Patient-Related Factors***

Among the patient-related factors influencing non-melanoma skin cancer diagnosis are general demographic characteristics, such as age, sex, and race. Mortality linked to this type of cancer is higher with: Older age, Male sex, and White populations [28].

1.3.5. Malignant Melanoma Skin Cancer

Malignant melanoma is considered the dangerous type of skin cancer that develops uncontrollably from pigment-producing cells called "melanocytes" and constitutes a potentially fatal disease. Diagnosing malignant melanoma at a preliminary stage is difficult to establish because most shared characteristics are similar to the clinical features of an atypical nevus.

Although malignant melanoma can appear anywhere in the body, including internal organs, this section focuses solely on the experiences of those who developed skin melanoma, often referred to as cutaneous malignant melanoma. There are four fundamental types of melanomas distinguished by their severity level and location on the body.

1.3.6. Types of Cutaneous Malignant Melanoma

Malignant melanoma, the most aggressive form of skin cancer, manifests in four distinct clinical subtypes—each with unique characteristics that influence diagnosis, treatment, and prognosis. While all melanomas arise from the uncontrolled growth of pigment-producing melanocytes, their presentation varies significantly.

A. **Superficial Spreading Melanoma**

This melanoma type is the most frequent, with approximately 70% belonging to this category, as its name suggests, it grows on the surface of the top layer for months or even years before penetrating deeper into the skin. This melanoma appears as a flat or slightly raised lesion, often with irregular borders and color variations. Lesions most

commonly occur on the torso, legs, and upper back in both sexes and are more frequent in middle-aged individuals.
The first sign of a new superficial melanoma is the darkening of part of an existing mole or the appearance of a new mole on previously unaffected normal skin. This melanoma type only becomes dangerous when it begins growing vertically into deeper skin layers.

B. **Nodular Melanoma**
Nodular melanoma is the most aggressive type of melanoma and represents about 25% (approximately 1 in 4) of all diagnosed melanomas. It typically develops more rapidly in depth and thickness than other melanoma types and may not have a clearly visible radial growth phase. Instead of developing from a pre-existing mole, it can form in a location where there was no previous lesion. This melanoma usually appears as a blue-black bump-shaped nodule, although 5% of lesions are pink or red and 5% are amelanotic (without pigmentation), making diagnosis more difficult. Nodular melanoma is more frequent in men than women, it appears most commonly in people over 60 years old (though it can develop at any age), and is most often found on the torso or head and neck.

C. **Lentigo Maligna Melanoma**
Lentigo maligna melanoma typically appears on sun-damaged skin in middle-aged and elderly individuals, particularly on the face. In its early stages, this melanoma type may be confused with a benign age spot or sunspot. Since the cancer can easily be mistaken for other conditions, it may go unnoticed for years, which can be particularly dangerous. It is often confused with lentigo simplex, a benign brownish spot that can develop in elderly individuals after years of sun exposure.
Lentigo maligna melanoma represents about 10% of diagnosed melanomas (approximately 1 in 10).

D. **Acral-Lentiginous Melanoma**
Acral-lentiginous melanoma is the most common melanoma in Black and Asian individuals, accounting for 50% of melanomas in people with these skin types. "Acral" comes from the Greek word "akron," meaning extremity, and the melanoma typically appears on the palms, soles of the feet, or under nails. It is much more frequent on feet than hands. Lesions are usually beige, brown, or black in color, with color variations and irregular borders. This melanoma type is often diagnosed later than others due to the mistaken belief that melanomas only develop on sun-exposed areas and that Black and Asian individuals are less likely to be affected. Confusing early signs of acral-lentiginous melanoma with bruises or injuries to palms, soles, or nail beds can also delay diagnosis.

Figure 1.4 represented examples of histological subtypes of malignant melanoma that demonstrates distinct characteristics:

1. **Superficial spreading melanoma** initially appears as a pigmented macule with irregular borders (Fig. 1A), progressing to show atypical melanocyte proliferation in the papillary dermis during invasion (Fig. 1B).
2. **Nodular melanoma** presents as a brown-black exophytic tumor (Fig. 1C) dominated by vertical growth, featuring pigmented epithelioid or spindle-shaped atypical melanocytes infiltrating the reticular dermis (Fig. 1D).
3. **Lentigo maligna melanoma** manifests as large, irregularly bordered macules on sun-damaged skin (Fig. 1E), histologically showing lentiginous proliferation of atypical spindle melanocytes at the dermo-epidermal junction with papillary dermis invasion; characteristic actinic damage and dermal elastosis (*) are evident in surrounding tissue (Fig. 1F).
4. **Acral melanoma** may present as an amelanotic nodule on extremities (Fig. 1G), with non-pigmented spindle melanocytes proliferating throughout the dermis (Fig. 1H).

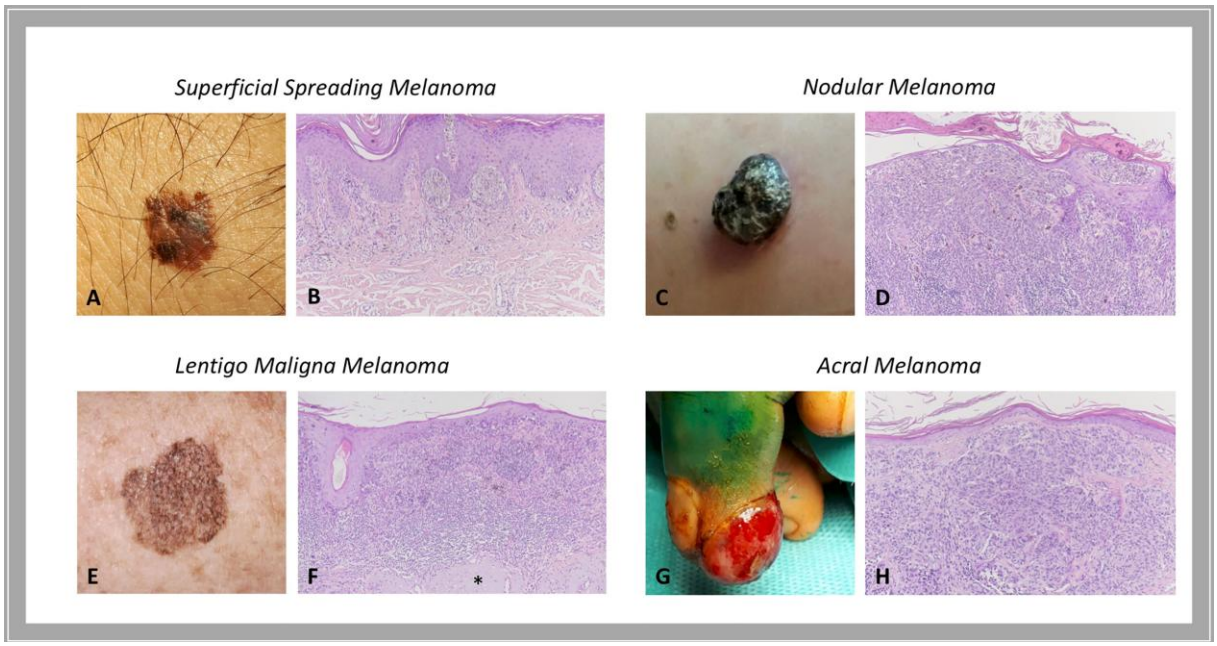


Figure 1-4: Examples of Cutaneous Malignant Melanoma.

1.3.7. Prevention Strategies

Cancer prevention has been defined as the set of measures that limit the progression of the disease at any stage of its development. Preventive measures can be implemented at any stage of the health process. Melanoma prevention operates at three levels:

- **Primary prevention:** UV avoidance and sunscreen use in sun exposed areas.
- **Secondary prevention:** Early detection via skin self-exams and professional screening.

- **Tertiary prevention:** Rehabilitation post-treatment.

Both primary and secondary prevention are important when it comes to malignant melanoma. Secondary prevention remains particularly crucial, as early detection dramatically improves prognosis [29], It measures include attempts to detect melanoma early in asymptomatic individuals by encouraging them to adopt practices such as skin self-examination and annual professional checkups.

1.3.8. Diagnostic Methods

The goal of effective clinical detection of primary melanoma remains paramount to the successful management of this cancer.

Visualization of skin lesions using one of the imaging techniques is not sufficient to distinguish malignant from benign melanoma. There is a need for reproducible diagnostic techniques that can be used by clinicians to understand the types of skin cancer. There are four commonly accepted reproducible methods for diagnosing skin cancers, particularly melanoma. These are the ABCD-E rule, the 3-point checklist, the 7-point checklist, the Menzies method, and pattern analysis.

- **The ABCDE Rule:**

This method was introduced in 1994 by Nachbar et al. [30]. ABCDE stands for Asymmetry, Border, Color, Diameter, and Elevation. These are five dermoscopic image criteria for the semi-quantitative assessment of skin lesions. Melanomas are usually asymmetrical, with irregular borders, and measure more than 6 mm in diameter, and they usually have mixed colors and change in size, color, shape, and bleed. These criteria (except E) have their possible scores based on the appearance of the skin lesion (see Table 2.2). These scores are multiplied by the associated weighting factors to obtain a Total Dermoscopic Score (TDS).

Table 1-2: Values utilized in multivariate analysis for determination of total dermoscopy score. [31]

Features	Description	Points	Weight Factor	TDS Score
A=A**symmetry**	The lesion is bisected by two lines that are placed 90 degrees to each other. Lesions that are symmetric in both axes are given zero points, one axial asymmetry one point and biaxial asymmetry two points. So, the points range from 0 to 2	0-2	1.3	0 – 2.6

B=**Border Sharpness**	First the lesion is divided into eight pie-shaped pieces. Next, the number of segments that have an abrupt perimeter cut off are counted. Thus, the points range from 0 to 8.	0-8	0.1	0 – 0.8
C=Color	Number of following colors present: light brown, dark brown, black, red, white, blue-grey. Thus, the points range from 1 to 6.	1-6	0.5	0.5 – 3.0
D=Dermoscopic structures	Number of following five structures: dots, globules, structureless (homogenous) areas, Network and branched streaks.	1-5	0.5	0.5-2.5

Table 1-3: Interpretation of total dermoscopic score (TDS). [31]

CATEGORY	SCORE RANGE
Benign	<4.75
Suspicious	4.75 - 5.45
Malignant	>5.45

- ## The 3-Point Checklist

It was presented in 2004 by Soyer et al.[32] and this method looks for three criteria: (1) asymmetry, (2) atypical pigment network, and (3) blue-white structures. The presence of any one of these criteria indicates the possibility of melanoma.

- ## The 7-Point Checklist

This method uses the scoring technique of the ABCDE rule. It was presented in 1998 by Argenziano et al. [33] and they explained the three major evaluation criteria and four minor criteria. The major criteria are the presence of one or more atypical pigmented tissues, one or more blue-gray areas, and one or more atypical vascular patterns, each scored as (2). For example, if all these criteria are present in a lesion, it is scored as 6. There are also minor criteria. These include the presence of radial streaming, irregular diffuse pigmentation, irregular dots and globules, and a regression pattern in a skin

lesion, each scored as (1). To establish a melanoma diagnosis, a minimum total score of (3) is required.

○ The Menzies Method

This method is based on a checklist of eleven features. In this method, the absence or presence of a total of eleven features is examined. It distinguishes benign lesions from malignant lesions using two sets of negative features and nine sets of positive features. The negative set includes only two features, namely symmetry and a single color, while the positive set includes nine features: the presence of a blue-white veil, multiple brown dots, pseudopods, radial streaming, scar-like depigmentation, peripheral black dots, multiple colors, multiple blue/gray dots, and a broad pigment network. The presence of at least one feature from the positive feature list and the absence of both features from the negative feature list are required to diagnose a lesion as malignant melanoma [34].

○ Pattern Analysis

Pattern or shape analysis is another method used to diagnose melanoma lesions and differentiate benign lesions from malignant melanomas. The pattern analysis method is used to identify specific patterns of skin lesions, which can be global or local. Some global patterns include reticular, globular, cobblestone, homogeneous, starburst, parallel, multicomponent, and nonspecific, which refer to benign lesions. Local patterns include pigment network, dots/globules/moles, streaks, blue-white veil, regression structures, hypopigmentation, blotches, and vascular structures, which also refer to benign lesions [35].

1.4. Skin cancer imaging techniques

When diagnosed at an early or minimally advanced stage, skin cancer can be successfully treated in 90% of cases, compared to just 50% when detected at an advanced stage. The development of non-invasive, high-resolution imaging techniques has significantly improved diagnostic accuracy for skin cancers and cutaneous lesions. Regarding melanoma specifically, diagnostic imprecision remains the primary cause of overtreatment (due to false-positive diagnoses) or inadequate treatment (due to false-negative diagnoses).

False-positive diagnoses substantially increase treatment costs by prompting unnecessary excision of benign lesions for biopsy and pathological examination. However, high-resolution imaging techniques show strong potential for improving diagnostic specificity, thereby reducing unnecessary procedures. The most commonly used imaging techniques for skin cancer diagnosis currently include [36]:

1.4.1. Reflectance Confocal Microscopy

This non-invasive imaging method uses a focused laser to visualize cellular details of the skin in real time. By detecting differences in refractive indices between cellular structures (cells, melanin, hemoglobin, etc.), it can differentiate light reflected by various skin components. However, it remains the most expensive skin imaging technique available.

1.4.2. Optical Coherence Tomography

This technique captures microscopic structures (several micrometers) in real time and can distinguish healthy from cancerous tissue. Its limitations include:

- Inability to visualize subcellular elements and membranes
- Limited detection of early-stage tumors
- Inability to definitively diagnose melanoma without histological confirmation

1.4.3. Ultrasonography

As one of the most common non-invasive techniques, ultrasound offers versatility, painless application, and minimal risk. While ultrasonic waves can penetrate deep skin layers to assess tumors, its low resolution cannot differentiate lesions at a histomorphological level. It also cannot detect tumors in their earliest stages.

1.4.4. Dermoscopy

Also known as epiluminescence microscopy, this practical, real-time non-invasive method enables early detection of malignant melanoma and other pigmented lesions. It visualizes subsurface skin colors and structures, with literature showing it improves diagnostic accuracy by 5-30% depending on lesion type. Melanoma may be identified when lesions show just 2-3 dermoscopic colors (or 1-2 clinically visible colors). Deeper melanomas typically reveal more colors.

1.5. Use of artificial intelligence for the diagnosis of skin cancer

Early diagnosis of malignant skin cancer is crucial for improving cure rates. While the traditional methods mentioned earlier are reliable, they can be time-consuming and subject to inter-observer variability. Consequently, the integration of artificial intelligence, as an emergent and high-potential discipline, offered endless possibilities; more precisely, Deep learning.

Deep learning is a subfield of machine learning, that has emerged as a promising approach to automate and enhance the accuracy of dermatological diagnosis. Generally, CNNs (Convolutional Neural Networks) show great potential for widespread application in classifying skin cancers and cutaneous lesions, and that is done via training the model using large datasets as we mentioned in the last chapter. However, existing skin cancer datasets face limitations: many studies use small sample sizes, clinical image datasets remain scarce, and most methods require creating new custom datasets for their specific needs. CNNs are currently the most widely used architectures for medical image analysis, including skin cancer diagnosis. Popular models like VGGNet, ResNet, and Inception have demonstrated remarkable ability to extract discriminative features from skin lesion images.

1.5.1. Datasets

Datasets play a critical role in developing AI models for malignant lesions diagnosis. These datasets must be high-quality, diverse, and clinically representative to enable machine learning algorithms to distinguish between benign and malignant lesions.

1.5.2. Key characteristics of skin cancer datasets:

- High-quality dermoscopic images: Clear, sharp images captured under optimal lighting conditions
- Precise annotations: Each image includes accurate labels indicating malignancy status and lesion type
- Data diversity: Should include patients of different ages, genders, skin types, and ethnicities to ensure model generalizability
- Class balance: Sufficient numbers of both melanoma and benign lesions to prevent learning bias
- Metadata: Additional clinical information (patient history, risk factors, etc.) to enhance model performance

Among popular skin cancer datasets are those from ISIC (International Skin Imaging Collaboration):

ISIC 2017 Dataset [37]**:**
Part of ISIC's annual challenge focusing on skin lesion analysis and classification. Contains approximately 2,000 annotated dermatological images including melanoma, basal cell carcinoma, and benign nevi.

HAM10000 Dataset [38]**:**
("Human Against Machine with 10,000 images") was created in 2018 to facilitate neural

network training for automated pigmented skin lesion diagnosis. This publicly available dataset contains 10015 dermoscopic images from ISIC archives, including:

- 327 actinic keratoses
- 514 basal cell carcinomas
- 1,099 benign keratoses
- 115 dermatofibromas
- 1,113 melanomas
- 6,705 melanocytic nevi
- 142 vascular lesions
  (For comparison, the PH2 dataset contains only 40 melanomas and 160 melanocytic nevi)

ISIC 2019 Dataset [39]**:**
A key update in ISIC's annual dataset series containing approximately 25,000 dermoscopic images - significantly more than previous years for greater diversity. Class distribution:

- Melanoma: ~1,100-1,200 images
- BCC: ~900-1,000 images
- SCC: ~200-300 images
- Actinic Keratosis: ~400-500 images
- Benign Keratosis: ~1,000-1,200 images
- Nevi: ~3,000-4,000 images
- Vascular Lesions: ~300-400 images

ISIC 2020 Dataset [40]**:**
Represents another milestone in skin lesion analysis datasets, used for tasks like lesion classification and segmentation. Contains approximately 32,542 dermoscopic images:

- 98.2% benign lesions
- 1.8% confirmed malignant melanomas (584 images)

ISIC 2024 Dataset [41]**:**

Also known as the SLICE-3D dataset, it is a more accessible collection of over 400,000 skin lesion images gathered from seven international dermatology centers. Unlike previous datasets, these images come from standard 3D full-body scans and are similar in quality to smartphone

photos. This is important because it means we can develop AI tools that work effectively in everyday healthcare settings, not just specialized dermatology clinics.

The DERM12345 dataset [42]: The DERM12345 dataset, released in mid-2024, is a vast dermoscopic image repository. It has 12,345 annotated and curated demographically representative patient populations comprising a wide spectrum of 38 to 40 lesion subtypes within 5 large lesion superclasses. These comprise benign (e.g., melanocytic nevi, seborrheic keratoses) and malignant lesions (e.g., basal cell carcinoma, squamous cell carcinoma, melanoma).

Every picture in DERM12345 is accompanied by rich metadata, such as lesion diagnosis, lesion site, and demographic information such as patient gender and age. Such diversity aids more robust and clinically relevant classification model training, especially for fine-grained lesion analysis and multiclass classification. In addition to its diagnostic variation, the database includes images taken with uniform dermoscopic lighting and stable resolution, making it suitable for supervised learning and benchmarking. DERM12345 mitigates some of the limitations that were faced in previous databases—e.g., class imbalance and limited subtype richness—making it a critical resource to be utilized in order to create dermatology AI systems that work towards more accuracy, lesion-type equity, and greater clinical utility.

Table 1-4: Comparative table of the mentioned datasets.

Dataset	Year	Total Images	Key Classes	Annotation Type	Reference
ISIC 2017	**2017**	**2,150**	**• Melanoma (374) • BCC (514) • Benign nevi (1,262)**	**Lesion masks, diagnoses**	**ISIC 2017 Challenge** [37]
HAM10000	**2018**	**10,015**	**• Melanoma (1,113) • BCC (514) • Actinic keratosis (327) • Benign nevi (6,705)**	**Dermoscopic, clinical labels**	**Tschandl et al. (2018)** [38]
ISIC 2019	**2019**	**25,331**	**• Melanoma (1,252) • BCC (3,323) • SCC (628)**	**Segmentation masks, metadata**	**ISIC 2019 Challenge** [39]

			• Benign keratosis (2,694)		
ISIC 2020	**2020**	**33,126**	**• Malignant (584, 1.8%)** **• Benign (32,542, 98.2%)**	**Pixel-level masks, metadata**	**ISIC 2020 Challenge** [40]
ISIC 2024	**2024**	**~400,000***	**• malignant (~393)** **• Benign (~400,000)**	**3D dermoscopy, genomic links**	**ISIC 2024** [41]
DERM12345	**2024**	**12,345**	**• 5 super classes** **• 15 main classes** **• 40 subclasses**	**Image-level diagnosis (single-label classification)**	[42]

1.6. Conclusion

In this chapter, we presented an overview of skin cancer, specifically malignant lesions. Beginning with foundational medical knowledge—histological subtypes, diagnostic criteria, and imaging modalities —we established the complexity of skin cancer detection. detailing fundamental aspects and traditional diagnostic methods. The combination of traditional medical knowledge with technological innovations opens new perspectives for early diagnosis and management of skin cancer, thereby increasing the chances of successful treatment and improved outcomes for patients.

In the next chapter, we develop a skin cancer classifier based on convolutional neural networks (CNN) and Transfer learning. We leverage pre-trained models such as VGG16 and MobileNetV2 to implement it.

2. Chapter 2. Artificial intelligence, Machine Learning, and Deep Learning

2.1. Introduction

The definition of the term “intelligence” was never agreed upon in the scientific world, there for, Legg and Hutter gave a collection of 70 informal definitions for intelligence, covering a broad point of views [43]. According to these definitions, human intelligence includes a greater variety of skills than what is usually included in broad definitions. The application of such a wide range of skills is a hallmark of human intelligence since it reflects the human propensity to test and push the limits of one's own abilities.

Due to the complex nature of our daily life problems, people must constantly adapt to new and changing situations. In addition to flexibility and creative problem-solving, this adaptation necessitates more efficient and effective decision-making processes. However, human reasoning isn’t always logical and accurate, and it frequently results in biased or less-than-ideal decisions, as demonstrated by the research of Kahneman [44]. The need for intelligent systems that can support or replicate human decision-making has been proven inevitable by this disparity between human cognitive limitations and the growing demands of automated solutions. In this regard, the study of artificial intelligence (AI) has become popular as a way to automate and simulate tasks that have historically required human intelligence.

Artificial Intelligence (AI) is a subfield of computer science where it specializes in creating models aimed to perform tasks that usually requires human intelligence. This concept of AI emerged in the 1950s [45], when researchers started working on early programs intended to make machines apt of performing intelligent tasks. Interest in AI has increased continually over the last decades, inspired by constant evolving research and persistent technological advancements.

In contemporary times, AI is widely recognized as a cornerstone of the Fourth Industrial Revolution and is considered one of the defining technologies shaping the future of our societies [46]. Its impact is far-reaching, influencing numerous aspects of daily life and improving industries by providing modern, scalable, and automated solutions to complex work inconveniences.

Naturally, as any novel promising domain, a variety of advanced techniques have been developed to enhance the performance and accessibility of AI systems, among which **machine learning (ML)** and **deep learning (DL)** are the most noted. These subfields focus on enabling machines to learn from data and improve over time without being explicitly programmed for

each task [47]. The ultimate goal of AI science is to simulate, and ultimately surpass, human intelligence and cognitive capacity. While the progress achieved so far has been remarkable, several challenges persist [48], and the scientific community continues to pursue novel architectures and learning paradigms to improve AI performance and generalization [49].

In the context of deep learning, artificial neural networks (ANNs) play a central role. Among the most commonly adopted models we note convolutional neural networks (CNNs), that are widely adopted for visual recognition tasks. Unlike traditional feature extraction methods such as SIFT (Scale-Invariant Feature Transform) and LBP (Local Binary Patterns), which rely on manual engineering and domain expertise [50], CNNs are capable of automatically learning hierarchical feature representations directly from raw data or images.

This chapter lays out the conceptual foundations that need to be understood in order to understand the research work within this thesis. It initially outlines the fundamental concepts of artificial intelligence, then presents general concepts related to machine learning. It then analyzes specific problems and definitions related to machine learning issues addressed in this research work. Special emphasis is placed on deep learning techniques, with a focus on the architecture and functioning of convolutional neural networks.

2.2. History of Artificial Intelligence

Artificial intelligence has long been the aspiration and fantasy of all people in the research field, and for decades it has been changing steadily, until a breakthrough was finally made and that had hitherto appeared to be impossible. AI is rooted back to 1943, when Warren McCulloch and Walter Pitts developed the first mathematical model of a neural network.

Nine years later, in 1952, Arthur Samuel created one of the first programs that could learn to play checkers on their own without human intervention. Consequently, in 1956, the first official conference on Artificial Intelligence was organized under the guidance of John McCarthy, who is referred to commonly as the father of AI [51].

Subsequently, in 1959, Arthur Samuel used the phrase Machine Learning (ML) to denote the ability of machines to acquire knowledge through experience. However, in the 1970s, expectations regarding AI received a negative blow. The British government's "Lighthill Report" was extremely critical regarding AI research, with the result that funding and interest for it was withheld at one point of time—referred to normally as the "AI winter.".

Machine learning started to take shape again in the 1990s with the development of algorithms such as Support Vector Machines (SVMs) and the advancement of speech and image recognition programs. The most significant breakthrough occurred in 1997 when IBM's DeepBlue defeated the world chess champion Garry Kasparov [52], in a haunting piece of history in this sector.

The trend of deep learning resumed in 2006, as Geoffrey Hinton and his team resumed deep neural networks and made a pathway for the deep learning revolution. In the early 2010s, a

combination of Big Data (large datasets) and high-performance computation tools such as GPUs made incredible breakthroughs in image recognition, natural language processing, and even game playing—as exemplified by DeepMind's AlphaGo beating a world championship player at Go in 2016.

Today, AI technologies like Machine Learning and Deep Learning are integrated into various areas of society, including medicine and banking, transport and entertainment, and continue to transform the world and the way we work [53].

2.3. Subfields of Artificial Intelligence

The subfields of AI each offer unique advantages and perspectives, focusing on different aspects of its development and deployment. By specializing in these areas, researchers and practitioners can enhance the capabilities of AI systems and address complex problems more efficiently and accuratly. Key subfields include Machine Learning (ML), Deep Learning (DL), Natural Language Processing (NLP), Computer Vision, and Robotics.

- **Machine Learning (ML):**
  ML is a branch of AI and computer science that focuses on leveraging data and algorithms to enable systems to learn in a manner similar to humans—progressively improving their performance and accuracy over time without being explicitly programmed for every task.
- **Deep Learning (DL):**
  DL is a subset of machine learning that employs artificial neural networks inspired by the structure and function of the human brain. It enables systems to automatically discover and learn patterns from large volumes of data, particularly in unstructured formats such as images or audio.
- **Natural Language Processing (NLP):**
  NLP enables computers to understand, interpret, and generate human language. It encompasses tasks such as speech recognition, machine translation, text analysis, sentiment detection, and question answering.
- **Computer Vision:**
  Computer vision aims to provide machines with the ability to perceive and interpret visual information from the world. This includes object detection, image segmentation, facial recognition, and scene understanding—allowing AI systems to process and respond to visual input in real time.
- **Robotics:**
  Robotics combines AI with physical systems to create intelligent machines capable of interacting with and manipulating the physical environment. This field includes areas such as robotic perception, motion planning, control systems, and human-robot interaction.

2.4. Machine learning

Machine Learning (ML) enables computers to learn from data without being explicitly programmed. It relies on statistical models and algorithms that allow systems to identify patterns within large datasets and use this knowledge to make informed predictions or decisions. By leveraging large volumes of data, machine learning models adapt and improve over time, increasing their accuracy and performance.

ML is a dynamic and rapidly evolving field, encompassing a wide variety of methods to achieve its objectives. These methods allow systems to generalize from previous experience to new, unseen data. Today, ML is applied across a broad spectrum of domains, from healthcare and finance to autonomous systems and natural language processing.

The main categories and techniques in machine learning include [54]:

- **Supervised Learning:**
  In supervised learning, models are trained on labeled datasets—each input is paired with a known output. This category includes algorithms such as linear regression, decision trees, support vector machines (SVM), and neural networks.

- **Unsupervised Learning:**
  Unsupervised learning focuses on discovering hidden patterns or structures in data that has not been labeled. Techniques used in this category include k-means clustering, k-nearest neighbors (k-NN), principal component analysis (PCA), and autoencoders.

- **Reinforcement Learning:**

In reinforcement learning, agents learn to make sequences of decisions by interacting with an environment and receiving feedback in the form of rewards or penalties. Common algorithms in this category include Q-learning, Monte Carlo methods, Deep Q-Networks (DQN), and actor-critic policy methods.

- **Semi-Supervised Learning:**
  This hybrid approach combines elements of supervised and unsupervised learning. A small amount of labeled data is used alongside a larger set of unlabeled data to improve learning performance. Techniques in this category include Transductive SVMs and consistency-based graph methods.

2.5. Deep Learning

Deep Learning (DL) is a subset of Machine Learning (ML) (see Figure 2.1) that utilizes artificial neural networks inspired by the structure and functioning of the human brain. These neural networks are capable of learning complex patterns from large volumes of data, making them especially effective for solving high-level tasks such as autonomous driving and facial recognition.

At its core, deep learning relies on multi-layered neural architectures that automatically extract and transform features from raw input data. By modeling these networks after biological neural systems, DL algorithms are able to capture intricate data representations and progressively improve their accuracy through experience and exposure to more data.

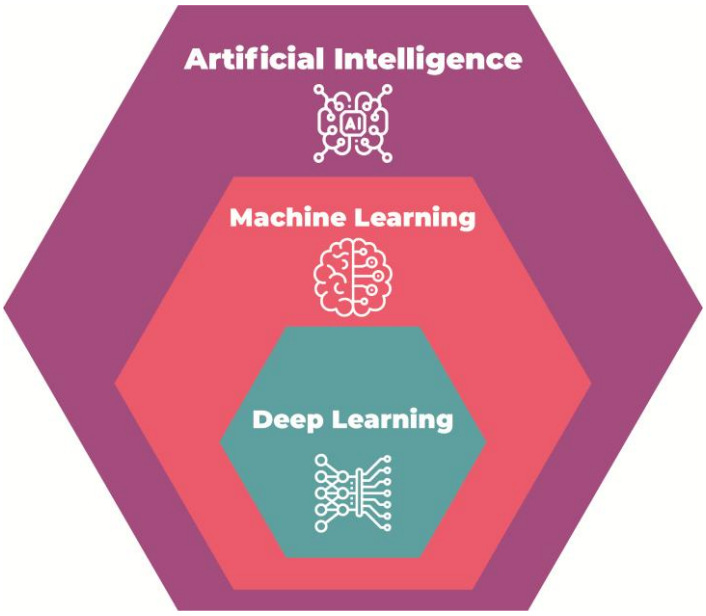


Figure 2-1: Overview: ML vs. DL vs. AI

2.5.1. Advantages of Deep Learning

Deep learning offers several advantages over traditional machine learning algorithms. Deep learning algorithms automatically learn to extract high-level features from raw data using multiple layers of artificial neural networks (ANNs). This enables deep learning models to learn from vast amounts of data and perform tasks such as image recognition and natural language processing with high accuracy. While machine learning algorithms can perform tasks such as classification and regression, they may not be as effective as deep learning algorithms when handling complex and unstructured data (see Figure 2.2) [55].

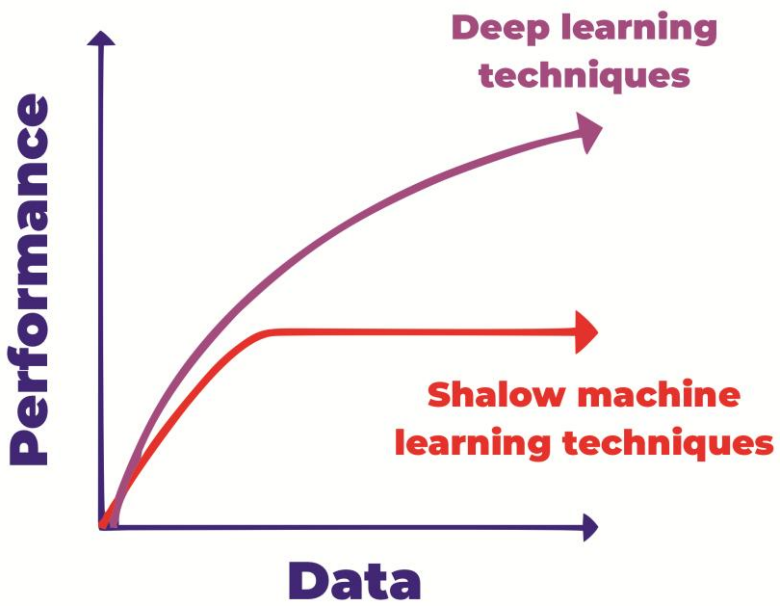


Figure 2-2: Impact of Data Volume on System Performance [56]

2.5.2. Artificial Neural Networks (ANN)

Artificial Neural Networks (ANNs) are computational models inspired by the learning processes observed in biological organisms. The human nervous system is composed of specialized cells known as neurons, which form intricate communication networks through connections involving axons and dendrites. The junctions where these axons and dendrites meet are called synapses—as illustrated in the left side of *Figure 2.3*.

Dendrites are responsible for receiving incoming signals from neighboring neurons via synaptic connections. These inputs are processed within the cell body (or soma), and if the cumulative electrical signal exceeds a defined activation threshold, the neuron generates an action potential. This electrical impulse then travels along the axon and reaches its terminal ends, where it is transmitted to other neurons in the network.

A biological neuron typically comprises three main parts:

1. The **cell body**, which processes the signal;
2. The **axon**, which carries the signal away from the cell;
3. The **dendrites**, which collect incoming signals.

When the cell body accumulates sufficient excitation (electrical charge), it transmits this excitation to the next neuron in the sequence [57]. This biological model is the foundational inspiration for the design of artificial neural networks, where neurons are represented as mathematical functions. Each artificial neuron computes a weighted sum of its inputs, applies a non-linear activation function, and passes the result to the next layer. This architecture enables the network to model highly complex relationships in data [58].

Artificial neurons are inspired by biology and mimic the functioning of the human brain. Our brain contains billions of neurons that help us interpret what we call signals. Today, artificial neurons allow us to transfer this principle into a computer by simulating a biological neuron as a software model. A real neuron can either be activated or not by an electrical signal. Similarly, an artificial neuron replicates this behavior by introducing what are known as weights. A weight is multiplied by the input value to produce a new result. In a biological neuron, these inputs accumulate and generate an electrical charge. An artificial neuron emulates this process by adding all the weighted inputs, a process known as the transfer function. The result of this summation is referred to as the net input to the neuron. When a biological neuron is sufficiently charged, it transmits an electrical impulse to the next neuron, provided a certain threshold is exceeded. This behavior is modeled in artificial neurons using an activation function, which determines whether the neuron should fire by comparing the input to a predefined threshold [59], [60]. Figure 2.3 demonstrates the similarities between the real biological neuron and the artificial neuron.

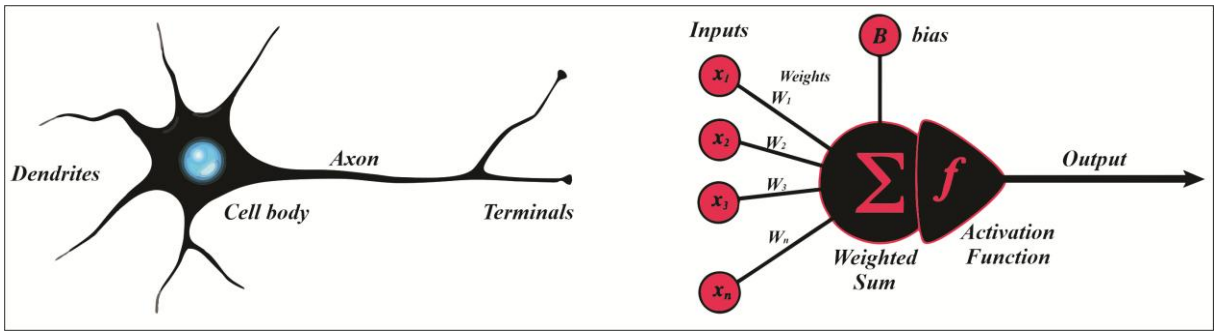


Figure 2-3: Differences between biologic and artificial neuron

2.5.3. Components of Artificial Neural Networks

Artificial Neural Networks (ANNs) are composed of several elements that work together to process information and make decisions. These components include [61]:

1. **Neurons (Units):** Also known as *perceptron*, these are the fundamental units of a neural network. Each neuron receives input signals, processes them, and sends output signals to other neurons in the network.
2. **Neuron Layers:** A layer is a group of interconnected artificial neurons. Layers are stacked on top of each other to form the network's architecture, allowing it to learn complex relationships between inputs and desired outputs.
    - The **input layer** receives the raw data,
    - The **hidden layers**, situated between the input and output layers, perform most of the computation,
    - The **output layer** produces the final predictions or outputs of the network (see Figure 2.5).
3. **Weights and Biases:** As illustrated in Figure 2.4, *weights* represent the strength of connections between neurons. They are initialized randomly and adjusted during training to minimize prediction error. *Biases*, on the other hand, are constant values

added to the weighted sum of inputs. They allow the model to better fit the training data by shifting the activation threshold.

4. **Activation Functions:** An activation function is a mathematical function applied to a neuron's weighted input to determine its output. Activation functions introduce non-linearity into the network, enabling it to learn complex relationships between inputs and outputs.
   Common functions include ReLU (Rectified Linear Unit), Sigmoid, Tanh (Hyperbolic Tangent), and Softmax.
5. **Loss or Cost Function:** This function measures the difference between the model's predictions and the actual values. It is used to quantify prediction error and guide the optimization process during training. Common loss functions include the Mean Squared Error (MSE) and Cross-Entropy.
6. **Optimization Process:** This refers to the algorithms used to adjust weights and biases in order to minimize the loss function. Popular optimization techniques include Stochastic Gradient Descent (SGD), Adam (Adaptive Moment Estimation), and RMSProp (Root Mean Square Propagation).
7. **Data:** The learning process relies on two types of data:
   - **Training data**, which is used to train the neural network and must be representative of the problem,
   - **Test data**, a separate dataset used to evaluate the model's final performance on unseen data.

2.5.4. Types of Artificial Neural Networks

In this section, we aim to delve into some different types of artificial neural networks to better understand their impact on modern technology. In general, each type of artificial neural network has its own strengths and specific use cases, so here is a non-exhaustive list of neural network types:

- **Feedforward Neural Networks (FFNNs)** [62]**:** Also known as deep feedforward networks, these are among the simplest types of artificial neural networks (ANNs). In FFNNs, data flows through several nodes until it reaches an output node. These networks have no cycles or loops, meaning that information cannot flow backward to influence previous steps. Such networks may struggle to accurately predict outputs for a given input, especially in complex tasks. This is where backpropagation comes in: it helps the network adjust its weights and biases. Backpropagation computes the error between the predicted and actual output, then propagates this error backward through the network, layer by layer, determining how much each weight and bias contributed to the total error.
- **Multilayer Perceptron (MLP)** [63]**:** An MLP is a specific type of FFNN where learning typically takes place through the backpropagation algorithm. It contains at least three layers and can classify non-linearly separable data (i.e., data that cannot be divided by a straight line). In an MLP, the network is fully connected, meaning every node (artificial neuron) in one layer is connected to all the nodes in the next layer.
- **Recurrent Neural Networks (RNNs)** [64]**:** These networks are similar to FFNNs but retain the output of a specific layer and feed it back as input. This allows the network to take previous information into account and potentially predict multiple outputs from a given input. For instance, if the first layer receives the output and sends it back as input,

subsequent layers will begin a recurrent neural process where each node retains the previous step. Thus, the system "remembers" previous errors and learns from them to improve future predictions. In short, RNNs can learn step-by-step to predict future outcomes.

- **Long Short-Term Memory Networks (LSTMs)** [65]**:** LSTMs are a subset of RNNs capable of learning long-term dependencies. Like RNNs, LSTMs are sequential networks, meaning they can retain information from previous steps and use it for current input processing. However, basic RNNs struggle to remember long-term dependencies and cannot deliver accurate results for inputs dependent on very old information. This is where LSTMs shine: they can choose to retain useful information for future use and discard irrelevant information.
- **Convolutional Neural Networks (CNNs)** [66]**:** These are a variant of MLPs with multiple layers, often fully connected. CNNs are primarily used to extract features from images, such as facial features. They identify characteristics based on the spatial relationship between pixels.
- **Graph Neural Networks (GNNs)** [67]**:** GNNs are designed to handle data organized as graphs. A graph is a structure made up of nodes (entities) and edges (relationships). GNNs analyze and exploit the relationships between graph entities to perform tasks like prediction or classification. Unlike CNNs, which are designed for grid-like image data, or RNNs for sequential data, GNNs process more flexible structures. They work by propagating messages between neighboring nodes in the graph. At each propagation step, a node aggregates information from its neighbors and updates its own representation. This process allows the network to learn complex relationships among entities in the graph.

2.6. Convolutional Neural Network (CNN)

The convolutional neural network operates differently from other types of networks because it processes data in a spatial manner. Instead of connecting each neuron to every neuron in the next layer, it uses local and shared connections, which allows it to capture local patterns in the data. Convolutional Neural Networks (CNNs) are used in various computer vision tasks, such as image classification, similarity grouping, and object recognition in scenes [68].

To understand computer vision, let's first consider human vision. Human vision is the ability of the human eye and brain to perceive and recognize objects. Computer vision is the process of giving a machine a similar, or even superior, ability to perceive and identify objects in the real world. The most commonly applied artificial neural network (ANN) in computer vision is the convolutional neural network (CNN).

There are two main reasons why CNNs are used in computer vision problems:

1. With traditional ANNs, solving computer vision problems is already difficult even for small-sized images.

2. CNNs require relatively minimal image preprocessing compared to other image classification algorithms, meaning that CNNs can learn the filters themselves.

Figure 2.4 shows an example of a CNN designed for cats and dogs' classification.

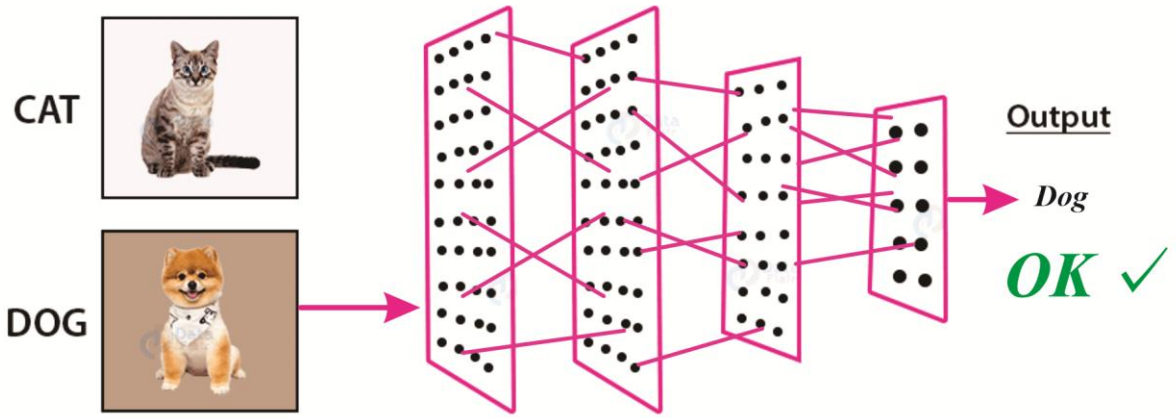


Figure 2-4: Example of a CNN classifier for cats and dog

2.6.1. CNN Architecture

Convolutional Neural Networks (CNNs) are a specialized type of artificial neural networks designed specifically for the analysis and processing of visual data. They replicate the biological visual process through the use of localized receptive fields, where every neuron responds to a specific region of an input image. CNNs are made up of many layers that consist of convolutional layers to extract features from learnable filters, pooling layers to reduce dimensionality and retain valuable information, and fully connected layers to interpret the extracted features for classification or prediction.

More specifically, the CNN architecture is composed of **Input layer, output layer, and multiple hidden layers**, and the hidden layers consist of other types of layers such as [58]:

- Convolutional Layers
- ReLU Layers (*Rectified Linear Unit*) as the activation function,
- Normalization Layers,
- Pooling Layers, and
- Fully Connected (FC) Layers.

This architecture enables CNNs to automatically learn hierarchical feature representations from low-level patterns like edges and textures to high-level structures like objects or facial features. Their capacity to extract features automatically and their high level of precision makes them extremely useful for numerous applications, including medical image analysis, object detection, face recognition, and image classification [69]. Figure 2.5 showcases the graphical representation of the CNN layers and architecture.

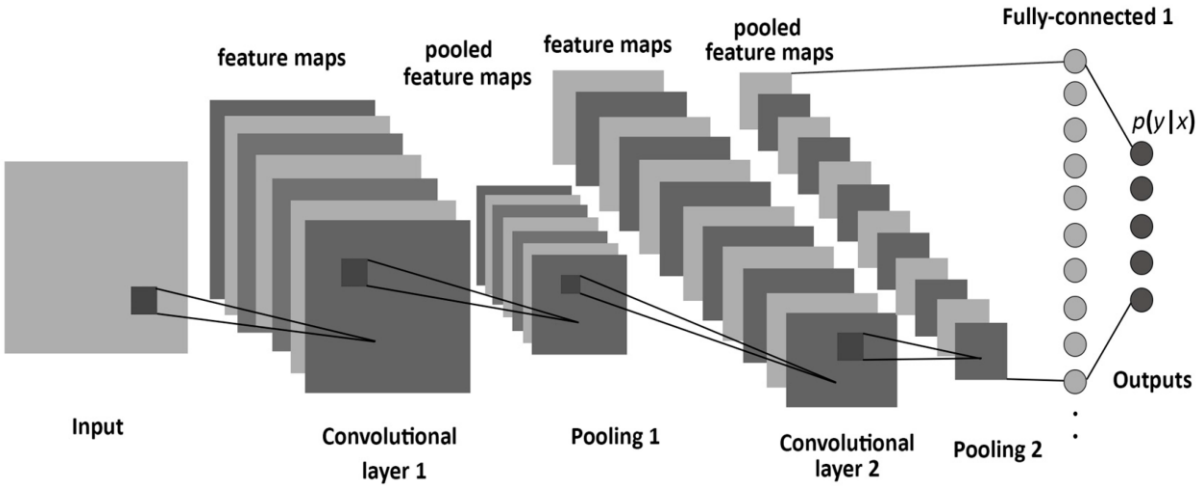


Figure 2-5: Architecture of a CNN. [70]

CNN architecture has basically four components, which are as follows:

2.6.1.1. Input Image

At the input of a Convolutional Neural Network (CNN), the raw image is fed into the network for processing. This image can be of any type: a human, an animal, a medical image (X-ray, CT scan, MRI, etc.), or a general object.

2.6.1.2. Convolutional Layer

The convolutional layer is the starting point of image processing. It detects patterns and local features in the input data. It consists of filters (small matrices like 3×3 or 5×5) that slide over the image and compute dot products with the local pixel values. The result is stored in a *feature map*. Multiple filters create multiple feature maps

2.6.1.3. Pooling Layer

The pooling layer (also called down-sampling or subsampling) reduces the spatial size of the feature maps while preserving important features. This helps ignore irrelevant data. Common types:

- **Max Pooling**: selects the maximum value in each region.
- **Average Pooling**: computes the average value.
- **Global Pooling**: computes a single value from the entire feature map. Max pooling is the most common. A 2×2 matrix window slides across the feature map, selecting the highest value in each region.

2.6.1.4. Flattening Layer

Flattening transforms the 3D data from the convolution and pooling layers into a 1D vector to feed into fully connected (dense) layers. For example, a 3×3 matrix becomes a vector of length 9.
Figure 2.6 demonstrates how a computer visualizes the letter X using the mentioned steps.

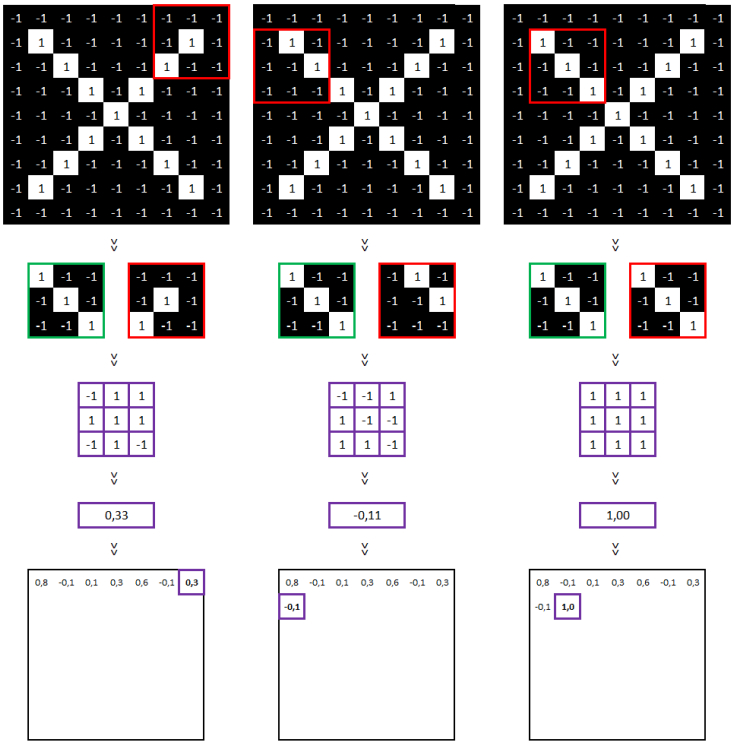


Figure 2-6: How a computer processes the letter "X".

2.6.2. Example of a CNN implementation

This section shows a complete example of building a CNN for an image classification task using the Tensorflow-Keras module.

Table 2.1 shows the sequences of a typical convolutional neural network.

Table 2-1: Components of a Convolutional Neural Network (CNN) and their code

Component	Role	Python Code Example
Input	Receives the raw image and prepares it for processing.	input_shape = (28, 28, 1)
Convolution	Applies filters to detect patterns (edges, textures).	Conv2D(32, (3, 3), activation='relu', input_shape=(28, 28, 1))
ReLU Activation	Introduces non-linearity and activates only positive outputs.	Activation('relu')
Normalization	Stabilizes and speeds up training by normalizing activations.	BatchNormalization()
Pooling (Max)	Reduces dimensions while retaining essential features.	MaxPooling2D((2, 2))
Second Convolution	Extracts more complex features.	Conv2D(64, (3, 3))
Second ReLU	Adds non-linearity after the second convolution.	Activation('relu')
Second Normalization	Normalizes activations again.	BatchNormalization()
Second Pooling	Further reduces the size of the feature maps.	MaxPooling2D((2, 2))
Flattening	Flattens data into a 1D vector for the dense layers.	Flatten()
Dense Layer	Combines extracted information to make a prediction.	Dense(128, activation='relu')
Output	Produces final results (e.g., probabilities for each class).	Dense(10, activation='softmax')

Figure 2.7 shows the code corresponding to the example given by Table 2.1:

```
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv2D, MaxPooling2D, ReLU,
    BatchNormalization, Flatten, Dense
# Definir le modele CNN
model = Sequential([
    # Couche d'entr e et couche convolutive
    Conv2D(32, (3, 3), input_shape=(28, 28, 1)),  # Couche
        convolutive
    ReLU(),  # Activation ReLU
    BatchNormalization(),  # Normalisation des activations
    MaxPooling2D((2, 2)),  # Couche de pooling
    # Deuxi me bloc de couches
    Conv2D(64, (3, 3)),  # Couche convolutive
    ReLU(),  # Activation ReLU
    BatchNormalization(),  # Normalisation des activations
    MaxPooling2D((2, 2)),  # Couche de pooling
    # Aplatir les donn es avant les couches FC layers
    Flatten(),

# Une seule couche entierement connectee
    Dense(128, activation='relu'),  # Couche dense
# Couche de sortie
    # 10 classes pour une t che de classification
```

Figure 2-7: Python code for a CNN architecture

2.6.3. Transfer learning and pre-trained models

In this section, we present the key concepts of transfer learning and pre-trained networks, also known as pre-trained models.

Suppose we want to use two approaches to determine whether a photo shows a dog or a cat. The first approach involves building a deep learning model from scratch and then feeding the new images into the network. Another option is to use a pre-trained deep learning neural network model that has already been trained on images of cats and dogs, instead of creating a neural network from scratch. Using a pre-trained model saves computational time and resources. A pre-trained network may also have some unexpected advantages. For example, nearly all photos of dogs and cats include other objects such as trees, the sky, and furniture. We could even use this pre-trained model to identify objects like trees, sky, and furniture.

2.6.3.1 Definition of a Pre-trained Model

Humans learn through experience; we apply knowledge acquired in one situation to similar situations we encounter in the future. Suppose you already know how to code in Python, and now you want to learn JavaScript. While the syntax and some rules differ, many core

programming concepts such as loops, conditionals, variables, and functions are similar. The logic and problem-solving skills you've developed in Python will help you quickly adapt to JavaScript. This is the essence of transfer learning—using knowledge from one domain to accelerate learning in another related domain.

This is precisely what transfer learning is. By definition, transfer learning is a concept in machine learning in which knowledge gained from performing one task is stored and reused when learning a different but related task [71], [72].

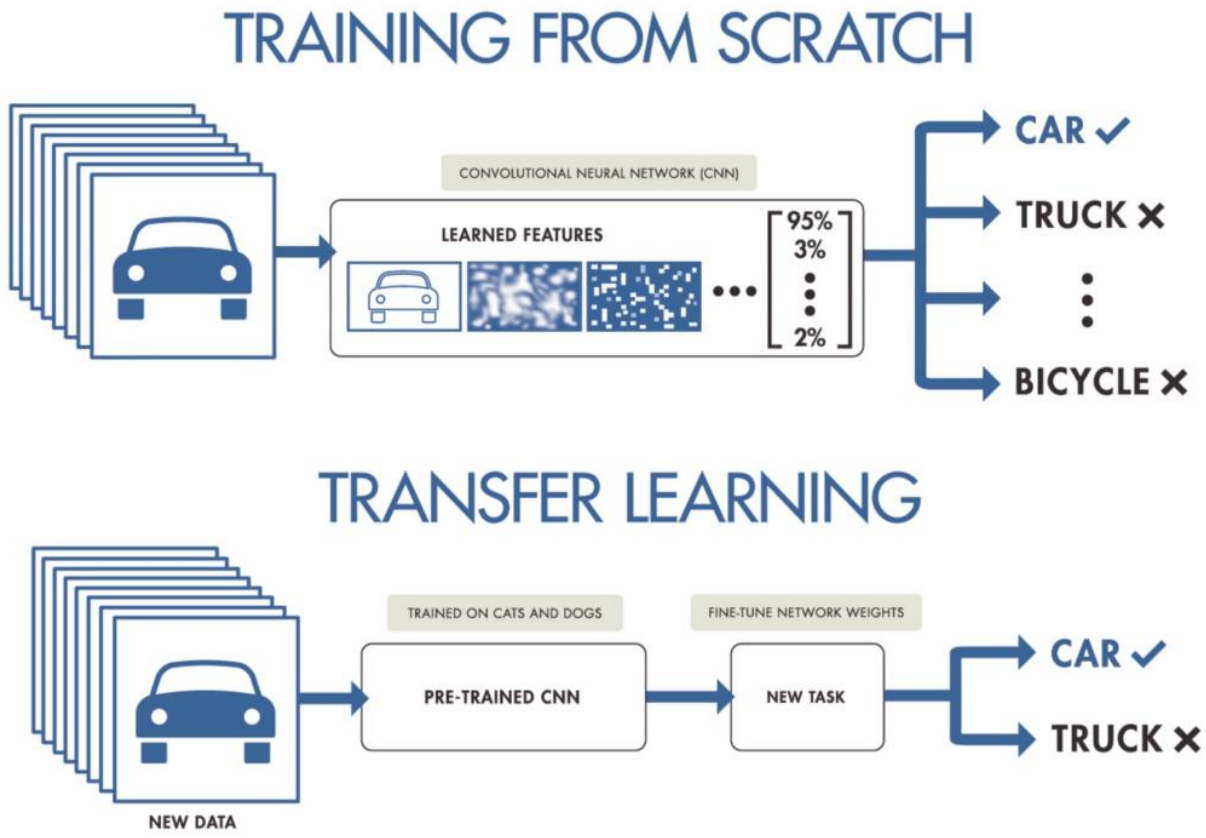


Figure 2-8: Differences between Transfer Learning and Training from Scratch

A pre-trained network is a saved network that has already been trained on a very large dataset, most often on image classification problems. To work with a pre-trained network, it is essential to understand the concepts of **feature extraction** and **fine-tuning**.

2.6.3.2 Feature Extraction

To understand feature extraction, we must revisit the architecture of a Convolutional Neural Network (CNN) (see Figure 2.5). The complete architecture of a CNN consists of the following components: convolutional layers, pooling and flattening layers, and an Artificial Neural Network (ANN).

Now, if we divide this architecture into two parts, the first part contains the convolutional and pooling/flattening layers (excluding the ANN), while the second part consists solely of the ANN. In this case, the first part is referred to as the **convolutional base**, and the second part is called the **classifier** (see Figure 2.9).

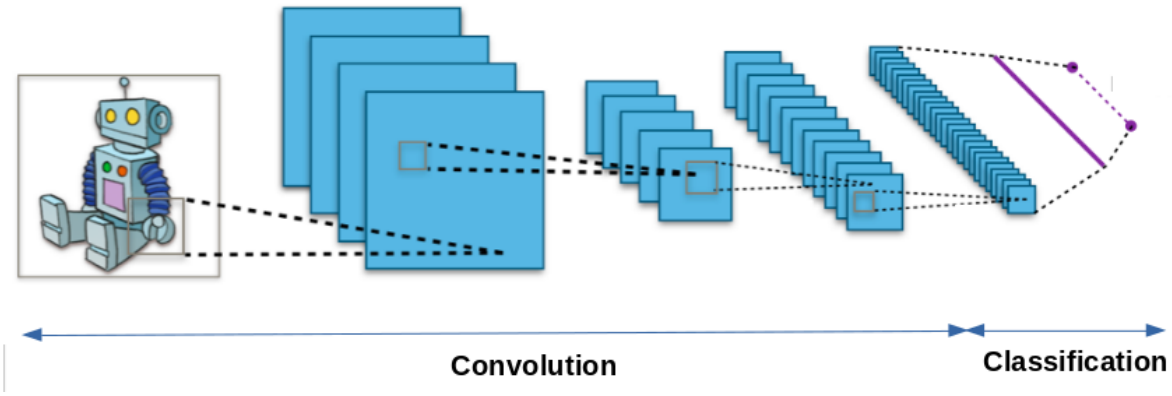


Figure 2-9: CNN split architecture – convolutional base and classifier

In the feature extraction process, the convolutional base is retained, while the classifier is modified [73]. This approach allows the reuse of the learned features from the convolutional layers while enabling the customization of the classifier for specific tasks. For instance, a classifier can be trained to distinguish between cats and dogs, bicycles and cars, or even medical radiographic images to classify tumors, infections, and other conditions.

The learning performed by the classifier is generally specific to the categories for which the model was trained. Therefore, it is recommended to transfer only the convolutional base and exclude the original classifier.

One of the most important concepts in working with pre-trained models is the freezing of certain layer**s**. This technique involves preventing the modification of parameters in specific layers of a pre-trained model during its adaptation to a new task.

Freezing essentially means stopping the weight updates of the lower layers within the convolutional base. These layers have already learned abstract representations from large amounts of training data. By freezing them, we preserve these useful features and prevent them from being lost during fine-tuning for a specific task [74]. Additionally, freezing the lower layers reduces the risk of overfitting [73], a phenomenon in which the model becomes too closely adapted to the training data, thereby impairing its ability to generalize to new, unseen data.

Finally, freezing layers helps conserve computational resources and training time, as it decreases the number of parameters that need to be updated during training. The process of freezing certain layers while training others is referred to as fine-tuning the network.

2.6.3.3 Fine-Tuning

We can freeze some of the initial layers in the network in order to preserve the information stored in them, this leads to a process called **fine-tuning**. As illustrated in Figure 2.10, fine-tuning involves adjusting a neural network so that it better fits the specific task at hand.

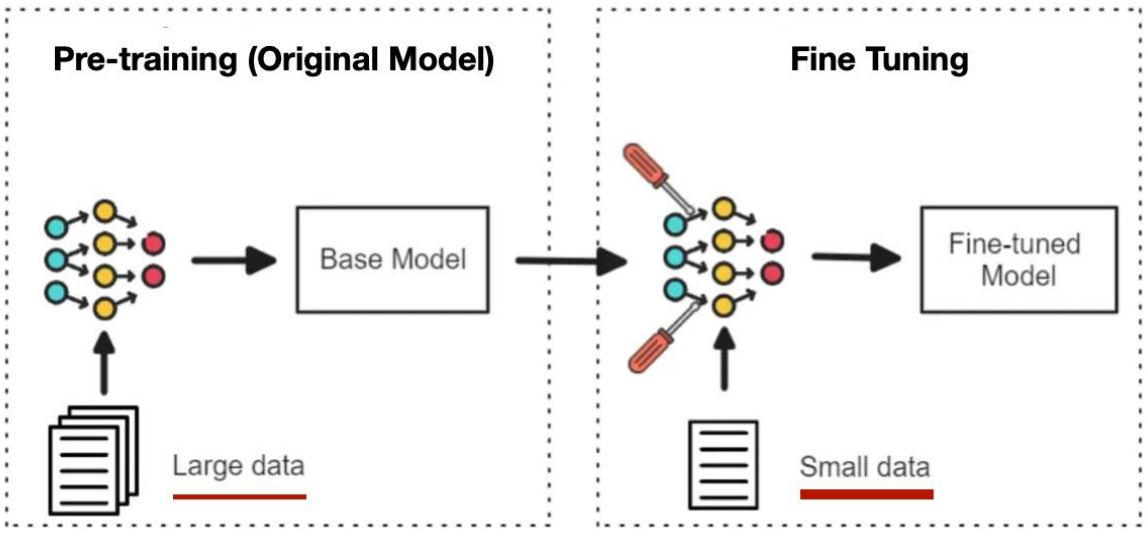


Figure 2-10: Fine Tuning Process.

Fine-tuning a pre-trained network is a powerful technique in deep learning that leverages models which have already learned general representations from large-scale datasets. Suppose we have a pre-trained network capable of identifying animals. If we want to identify specific animals such as dogs and cats, we can slightly modify some of the layers to teach the model what dogs and cats look like. This is similar to reusing the entire pre-trained network and adding a new classifier trained specifically on dog and cat images. The same concept applies when we use a pre-built network and append a custom classifier to it, which will be trained on our specific dataset.

Here are the three main steps involved in fine-tuning [75]:

- **Add a classifier (ANN)** to a pre-trained model by adding new fully connected (FC) layers that will be trained on the specific task.
- **Freeze the convolutional base** (convolutional and pooling layers), which prevents their weights from being updated during training.
- **Compile and train** the added layers and classifier jointly.

For more advanced fine-tuning, one may **unfreeze the top layers** of the base model and retrain them using a lower learning rate. This step is optional and depends on the similarity between the new dataset and the original one.

2.6.3.4 Popular Pre-trained Models

In real-world professional settings, it is rare to build a convolutional neural network (CNN) entirely from scratch. Instead, we usually rely on pre-trained models. These models serve as convolutional bases to which custom classifiers (ANNs) can be attached. Among the most commonly used pre-trained models are presented in the following table:

Table 2-2: Pre-trained CNN models

Model	Architecture	Applications	Examples
VGGNet	CNN with simple convolutional layers	Image classification, object recognition	VGG16 , VGG19 [76]
ResNet	Deep CNN with residual connections	Image classification, object detection, image segmentation	ResNet50, ResNet101 [77]
Inception	CNN with Inception modules	Image classification, object detection in complex scenes	InceptionV3 [78], GoogleNet [79]
EfficientNet	CNN with compound scaling coefficients	Image classification, object detection, image segmentation	EfficientNetB0, EfficientNetB7 [80]
MobileNet	Lightweight CNN optimized for mobile devices	Real-time object detection on mobile and embedded devices	MobileNet [81], MobileNetV2 [82]
Xception	Inception-inspired CNN with depthwise separable convolutions	Image classification, semantic segmentation, object recognition	[83]

The table showed some of the most known pretrained models, but for our coming research here are the models that drew our attention or that we found very helpful for our case. Among CNN architectures, ResNet50, EfficientNet-B0, MobileNetV2, DenseNet121, and VGG16 have emerged as popular choices due to their balance between performance and computational efficiency.

ResNet50 (Residual Network with 50 layers) addresses the problem of vanishing gradients in deep networks through residual connections, or "skip connections," which allow the network to learn identity mappings alongside residual functions. These connections enable the training of very deep architectures without degradation in accuracy, making ResNet50 particularly effective for complex image classification tasks, including skin lesion recognition. Its modular

design, consisting of bottleneck residual blocks, allows the network to capture both local and global contextual features, improving generalization on diverse datasets.

EfficientNet-B0 represents a family of CNNs optimized through a compound scaling method that balances network depth, width, and resolution. Unlike traditional networks that increase a single dimension arbitrarily, EfficientNet systematically scales all three dimensions using a fixed set of scaling coefficients, resulting in a model that is both parameter-efficient and computationally lightweight. EfficientNet-B0, the baseline model, has fewer parameters than ResNet50 yet achieves competitive accuracy, making it suitable for resource-constrained environments such as mobile devices or embedded diagnostic systems.

MobileNetV2 is designed for efficient mobile and embedded vision applications. It employs depthwise separable convolutions to drastically reduce the number of parameters and computational cost, while its inverted residual blocks with linear bottlenecks facilitate feature reuse and mitigate information loss. MobileNetV2 is particularly useful for real-time image processing, where inference speed and memory efficiency are critical. Despite its compact size, it retains sufficient representational power for tasks like lesion classification and preliminary diagnosis in dermatology.

DenseNet121 introduces dense connectivity patterns wherein each layer receives inputs from all preceding layers and passes its feature maps to all subsequent layers. This dense connectivity promotes feature reuse, mitigates vanishing gradients, and encourages implicit deep supervision throughout the network. DenseNet121, with 121 layers, offers high parameter efficiency compared to networks of similar depth, and its compact feature maps are beneficial in capturing subtle variations in medical images, such as differentiating between benign and malignant skin lesions.

VGG16, a classic CNN architecture, emphasizes simplicity and depth through a sequence of 3×3 convolutional layers followed by max-pooling operations. While VGG16 is computationally heavier and more memory-intensive than modern architectures, its uniform design allows straightforward visualization of learned features, making it valuable for interpretability studies in medical AI. VGG16 remains a common baseline in comparative studies and transfer learning experiments for image classification tasks.

The performance of these networks is highly dependent on the choice of optimizer, which dictates how model parameters are updated during training. Stochastic Gradient Descent (SGD) with momentum is widely used for its stability and convergence properties, whereas adaptive optimizers such as Adam or RMSProp dynamically adjust learning rates for each parameter, accelerating convergence and improving performance on datasets with heterogeneous image characteristics. The selection of an optimizer often balances training speed, generalization, and robustness against overfitting, particularly in medical image datasets with limited sample sizes.

Beyond discriminative architectures, generative models have gained prominence in medical imaging for data augmentation and synthesis. Generative Adversarial Networks (GANs),

composed of a generator and a discriminator, learn to produce realistic images by competing objectives: the generator attempts to create images indistinguishable from real ones, while the discriminator seeks to identify synthetic samples. Variants like DCGAN and StyleGAN have been employed to augment datasets with rare pathological conditions, enhancing model robustness. Additionally, Variational Autoencoders (VAEs) learn latent representations that can be sampled to generate novel images, offering a probabilistic framework for image synthesis and feature exploration.

In conclusion, deep learning in medical imaging relies on a careful selection of network architecture, optimization strategy, and, increasingly, generative modeling. Architectures like ResNet50, EfficientNet-B0, MobileNetV2, DenseNet121, and VGG16 provide diverse options, ranging from high accuracy and deep representation to lightweight and efficient inference. Coupled with effective optimizers and augmented by generative approaches, these models enable accurate, robust, and scalable solutions for challenging tasks such as skin lesion classification and early disease detection, forming the cornerstone of modern AI-driven diagnostic systems.

2.7. Data Preparation

After selecting a pre-trained model trained on a large and diverse dataset such as ImageNet [84], preparing data for transfer learning with CNNs involves several steps:

2.7.1. Data Loading and Preprocessing

Loading the dataset into the programming environment is always the first step, next the common deep learning libraries such as Keras and TensorFlow must be imported which provide convenient preprocessing functions. Preprocessing involves cleaning, organizing, and resizing the images to match the input size expected by the pre-trained model. This typically includes converting images to a uniform format and size (e.g., 224x224 pixels, RGB channels); most deep learning frameworks support common image formats such as JPEG, PNG, and BMP. Libraries such as OpenCV or Scikit-Image can be used to handle additional formats.

After that, images need to be organized into folders by class, or via the use of structured files such as CSV or Excel to indicate the class labels. This facilitates automatic discovery of the class-path relationships during loading. A metadata file typically contains information such as image names, class labels, and additional features.

Once the image information is read and loaded, preprocessing is done for use in a CNN model, either from a directory using Keras's flow_from_directory function, or from a CSV/XLS metadata file using a Pandas DataFrame and the flow_from_dataframe function provided by Keras.

2.7.2. Data Normalization

Normalization ensures that image data are on a consistent scale, which is essential for effective training of CNNs. Common normalization techniques include subtracting the mean pixel

intensity from each channel or scaling pixel values to a specific range (e.g., from 0 to 1 or from -1 to 1).

2.7.3. Data Augmentation

Data augmentation is an optional step used to artificially expand the dataset by creating modified versions of existing images. This helps the model learn features that are invariant to slight changes such as rotation, scaling, or flipping. It enhances the model's generalization performance and helps prevent overfitting [85].

Augmentation techniques include rotating, shifting, flipping, and zooming images. Applying such transformations increases dataset diversity, effectively creating more training samples from a limited dataset. For image augmentation, TensorFlow's ImageDataGenerator object is commonly used, with parameters like shear_range, zoom_range, and horizontal_flip to apply image transformations.

Data augmentation offers several benefits [86], [87]:

- **Reducing overfitting** by creating multiple variants of the same image.
- **Expanding dataset size**: A single image can produce several transformed versions. This helps balance datasets and improves model performance, especially in the presence of class imbalance.
- **Improving robustness for new predictions**: By simulating multiple views of the same object, the model becomes better at predicting unseen data.

2.7.4. Data Splitting

Data splitting is a critical step in training artificial neural networks (ANNs). It involves partitioning the dataset into separate subsets for training, validation, and testing. As shown in Figure 2.11, data should be divided into three distinct sets:

- **Training set**: The dataset used to train the model and learn hidden patterns. During each training iteration, the same training set is repeatedly fed into the network.
- **Validation set**: A portion of the training set used to evaluate model performance during training. It helps adjust hyperparameters and ensures that training is progressing effectively. Validation also helps prevent overfitting by identifying when the model becomes too specialized to the training data.
- **Test set**: A separate dataset used only after training is complete to assess the model's final performance using metrics such as accuracy and precision.

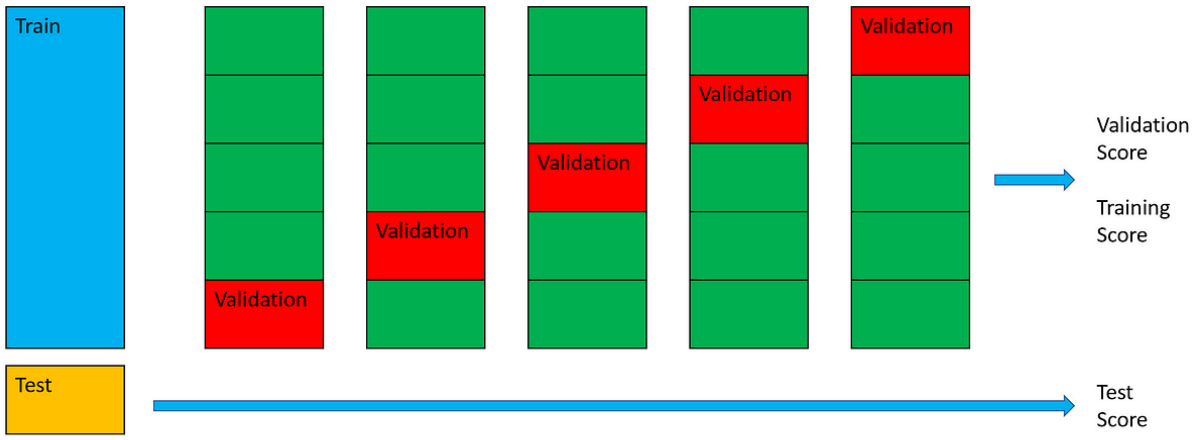


Figure 2-11: Dataset splitting into: training set, validation set, and test set.

The creation of distinct datasets allows us to fine-tune the model's real-world performance. The percentage and methodology for dataset partitioning depend on both the sample size and the model's requirements.

In machine learning, the standard practice involves splitting data into **training, validation, and test sets** using ratios such as **70:15:15%** or **60:20:20%**.

2.8. Evaluation Metrics

Effective model evaluation extends beyond technical performance, encompassing the broader context of the research problem. Common metrics for classification tasks include:

- **Accuracy**
- **Precision**
- **Recall** (also termed *Sensitivity* or *True Positive Rate, TPR*)
- **Specificity** (*True Negative Rate, TNR*)
- **False Positive Rate (FPR)**
- **False Negative Rate (FNR)**
- **F1-score**
- **Receiver Operating Characteristic (ROC) curve**
- **Area Under the Curve (AUC)**

These metrics are summarized in the following table:

Table 2-3: Classification Common Metrics

Metric	**Description**	**Equation**
Accuracy	It is the number of correct predictions divided by the total number of predictions. Percentage of correct predictions among all predictions.	$\frac{TP + TN}{Number\ of\ Samples}$
Precision	Percentage of correct positive predictions among all positive predictions. Evaluates the accuracy of the positive predictions made by a model.	$\frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}$
Recall/ Sensitivity	Percentage of true positive instances among all samples that are actually positive.	$\frac{\mathrm{TP}}{\mathrm{TP+FN}}$ or $1 - FNR$
Specificity (*True Negative Rate, TNR*)	Percentage of true negative instances among all samples that are actually negative.	$\frac{\mathrm{TN}}{\mathrm{TN+FP}}$ or $1 - FPR$
False Positive Rate (FPR)	Percentage of negative samples incorrectly classified as positive.	$\frac{\mathrm{FP}}{\mathrm{FP+TN}}$ or $1 - TPR$
False Negative Rate (FNR)	Percentage of positive samples incorrectly classified as negative.	$\frac{\mathrm{FN}}{\mathrm{FN+TP}}$ or $1 - TNR$
F1 Score	Harmonic mean of precision and recall, providing a balanced measure of both metrics.	$2 \times \frac{\mathrm{Précision} \times \mathrm{Recall}}{\mathrm{Précision} + \mathrm{Recal}}$
ROC	A curve that plots the TPR (Recall) against the FPR for different classification thresholds.	There is no single formula, but the curve is plotted with ROC = True Positive Rate (TPR) vs. False Positive Rate (FPR).
AUC	A measure of the model's ability to distinguish between classes, especially in ROC curves.	$AUC = \int_{-\infty}^{+\infty} ROC$

On the other hand, the assessment of the quality of classifications can also be based on a confusion matrix and the following four components [88]:

The confusion matrix consists of the following four instances:

- **True Positive (TP):** The number of instances that were *originally positive* and were *correctly predicted* as positive (e.g., Class 1).
- **False Positive (FP):** The number of instances that were *originally negative* but were *incorrectly predicted* as positive. This error is also known as a **Type I error**.
- **True Negative (TN):** The number of instances that were *originally negative* and were *correctly predicted* as negative (e.g., Class 0).
- **False Negative (FN):** The number of instances that were *originally positive* but were *incorrectly predicted* as negative. This error is referred to as a **Type II error**.

The objective is to **maximize TN and TP** while **minimizing FN and FP**.

The confusion matrix is an essential tool for evaluating the performance of a classification model. It provides a clear visualization of the model's errors and enables precise calculation of evaluation metrics. *Figure 2.11* presents a confusion matrix, illustrating the alignment between the model's predictions and the true labels.

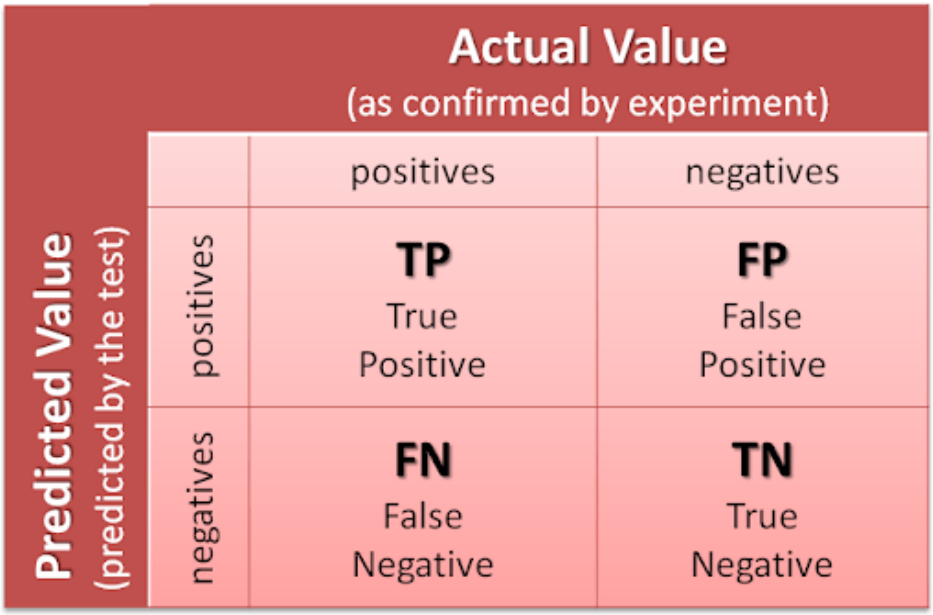


Figure 2-12: Confusion Matrix. [89]

2.9. Conclusion

In this chapter, we have explored the fundamentals and advanced techniques related to convolutional neural networks, most precisely in image classification context. We started by a presentation of the CNN architecture, while distinguishing its convolutional base, responsible for the extraction of characteristics, and the classifier, responsible for the final prediction. This

layout allowed us to transition to the concept of transfer learning, a powerful strategy that consists of exploiting gained knowledge from large data sets, with the possibility to adapt it into specific task.

We then highlighted the technique of Fine Tuning, which relies on partially freezing the layers of a pre-trained model in order to maintain general representations while adapting the classifier to a new task. This method gives the model the advantage of accelerated training and reduces the risk of overfitting, and most importantly, it optimizes the performance even with limited dataset samples.

Among the most commonly used pre-trained models, we reviewed VGGNet, ResNet, Inception, EfficientNet, MobileNet, and Xception, each with its advantages depending on the application context, computational constraints, or desired depth.

Subsequently, we detailed the essential steps of data preparation, ranging from loading and preprocessing images to their normalization, augmentation and partitioning into training, validation and test sets. These steps are crucial to ensure efficient training, robust generalization of the model, and reliable evaluation of its performance through metrics that we already saw in the same subsection such as accuracy, precision, recall, F1-score, and ROC curve.

Finally, we presented code example and demonstrated the advantages and differences of traditional deep learning versus transfer learning. Thus, this chapter establishes a solid foundation for the design of efficient computer vision models, particularly in contexts where annotated data is limited and where the use of pre-trained networks is a pragmatic and efficient solution.

In the following chapter, we focus on skin cancer as a case study, applying all the concepts and knowledge introduced earlier.

3. Chapter 3. Early skin cancer detection using transfer learning model and DCGAN approach for data balancing

3.1. Introduction

Dermoscopic imaging has proven to be an essential tool for evaluating skin lesions and classifying them as normal, benign, or malignant using a dermatoscope. It is one of the most widely used diagnostic techniques in dermatology due to its non-invasive nature, real-time analysis, portability, and cost-effectiveness.

The primary objective of this chapter is to present a skin cancer diagnosis approach based on deep learning techniques. Specifically, the proposed method leverages Deep Convolutional Neural Networks (DCNNs) and transfer learning using pre-trained models such as VGG16 and MobileNetV2. The developed pipeline consists of four main phases: dataset selection, data preprocessing, feature extraction, and fine-tuning.

3.2. Background

3.2.1. Related works

Application of artificial intelligence (AI), and more particularly deep learning, in the field of medical imaging has been a subject of great interest in the last few years. A vast number of research works have dealt with the automatic analysis of skin lesions from dermoscopic images for assisting the early detection of melanoma and other skin cancers. These advances have been driven largely by the availability of public datasets, the development of convolutional neural networks (CNNs), and growing interest in transfer learning techniques. In this section, we recap the most relevant research that has explored skin lesion classification, image synthesis for data augmentation, and total body photography (TBP) systems. Specific emphasis is placed on comparing their approaches, datasets, model architectures, and performance metrics for highlighting both the progress made and the open issues in the field.

Aljohani and Turki [90] conducted a comprehensive evaluation of eight prominent CNN architectures—including DenseNet201, MobileNetV2, ResNet50V2, ResNet152V2, Xception,

VGG16, VGG19, and GoogleNet—to classify melanoma using the ISIC-2019 dataset. They trained and validated their models on a set of 7146 dermoscopic images, and subsequently they tested its performance on a separate test images. Among all the models tested, GoogleNet achieved the best results, with a test-set accuracy of 76.08% and training accuracy of 74.91%. The study also compared model performance to dermatologists' diagnostic capabilities, revealing that GoogleNet surpassed expert clinicians in sensitivity, detecting melanoma in 84.5% of cases—outperforming dermatologists, whose sensitivity was 73.3%. In summary, their work demonstrates that among several deep CNN models applied to a public dermoscopic image dataset, GoogleNet offers the strongest balance of accuracy and sensitivity, even surpassing human experts in identifying melanoma lesions.

Djaroudib et al. [91] explored the impact of image quality versus quantity for melanoma diagnosis using transfer learning, where a VGG-16 model pre-trained on ImageNet was fine-tuned on the HAM10000 dataset. They determined that a model trained with a high-quality yet comparatively lower collection of images offered higher accuracy—around 94%—than models trained using larger but not necessarily high-quality datasets. The outcome highlights the importance of image clarity and suitability as crucial determinants toward guaranteeing stable diagnostic performance.

Independently, Sikandar et al. [92] created "SCDet," a more effective, 32-layered convolutional neural network designed to identify suspicious skin lesions efficaciously. SCDet, after training on the smaller DermIS dataset, performed flawlessly with 99.6% accuracy, high sensitivity, and high specificity, and was appropriately geared towards generalization with 85% accuracy when tested on the external HAM10000 dataset. Their conclusions show that tailored, light-weight CNN designs could offer a compelling compromise between diagnostic performance and computational load, particularly for constrained or real-time clinical environments.

Yashwant S. Ingle and Shaikh [93] conducted a comparative evaluation of skin cancer classification on the HAM10000 dataset with both a self-coded convolutional neural network (CNN) and a transfer learning technique with VGG16. They showed that incorporating additional dense layers and dropout regularization into VGG16 significantly improved its performance, achieving comparable accuracy and weighted F1-score of approximately 89%. Their work illustrates the advantage of starting from strong pre trained models instead of training models from scratch.

In a separate study, K et al. [94] proposed a sequential CNN architecture trained on a large, unnamed Kaggle skin lesion dataset. This model demonstrated high effectiveness in distinguishing among seven different lesion types, achieving around 95% accuracy and a 96.3% F1-score, underscoring the ability of well-structured CNNs to support multi-class skin cancer diagnosis.

On the other hand, M'hamedi el al. [95] focused on reducing class imbalance in the SIIM-ISIC 2020 dataset through extensive data augmentation, combined with transfer learning using fine-tuned VGG19 and MobileNetV2 . Their best-performing variant, based on MobileNetV2, reached 95.16% accuracy, demonstrating that thoughtfully applied augmentation and lightweight pre-trained models can significantly improve melanoma classification accuracy. Together, these studies illustrate the consistent benefits of transfer learning and architectural refinement (especially when paired with data balancing techniques) in advancing the performance of skin lesion classification models.

Gouda et al. [96] analyzed the impact of skin cancer categorization from high-resolution images by adding Enhanced Super-Resolution Generative Adversarial Network (ESRGAN) as a preprocessing stage before training various deep learning architectures on the ISIC-2018 dataset. After fine-tuning procedures on models such as ResNet-50, InceptionV3, and Inception-ResNet, InceptionV3 achieved highest accuracy at approximately 85.8%, demonstrating the potential of super-resolution methods to enhance classification performance dramatically when the quality of images is increased. Their work highlights the degree to which intentional preprocessing can improve diagnostic outcomes in skin lesion analysis.

In another study, Mehr and Ameri [97] utilized a multimodal approach by combining lesion images with patient context like age, gender, and lesion location through an Inception-ResNet-v2 backbone. Evaluated on three diverse datasets (ISIC-2019, PAD-UFES-20, and Fitzpatrick 17k), their model achieved an astonishing 94.5% accuracy for benign/malignant classification of lesions. This suggests that integrating patient context into image-based models can potentially increase diagnostic accuracy significantly, which stresses the value of metadata integration in AI-aided dermatology.

Alrabai et al. [98] examined the effectiveness of transfer learning by comparing the accuracy of two of the most popular architectures: InceptionV3 and Xception; on a Kaggle-based skin cancer dataset using fine-tuning techniques. Their results showed that InceptionV3 was more accurate than Xception with an accuracy of approximately 89.1%, a vast improvement from existing benchmarks. In another comparative study, Islam and Panta [99] applied five pre-trained CNNs on a portion of the Kaggle ISIC dataset of 3,297 images. ResNet-50, after tailored optimization of activation units and network layers, had the highest performance with accuracy of 93.5% and F1-score of 0.86, showing the remarkable improvement that can be attained by fine-tuning. Talking about AI diagnostic fairness, Pope et al. [100] examined potential skin color bias with skewed (3,623 images) and balanced (~500 images) datasets. Their findings recognized models performing substantially better for light skin tones, with a difference of selection rate of 27.5% vs. 15.9% in the skewed set, and 50.0% vs. 34.2% in the balanced set, considerably below the accepted fairness threshold of 0.80. These results indicate continued bias in preliminary skin cancer detection technology and underscore the need for more balanced AI technology in dermatology.

Imran et al. [101] used the EfficientNetB0 architecture in combination with Ant Colony Optimization (ACO), a nature-inspired feature extraction algorithm, to enhance classification on a self-created ISIC-derived skin lesion dataset. Their hybrid CB-SVM model achieved a remarkable accuracy of over 98% with the added virtues of speed and efficiency.

Naeem et al. [102] introduced SNC_Net, a hybrid scheme combining hand-crafted image features with deep learning output. Evaluated using the ISIC 2019 benchmark, SNC_Net performed better than baseline networks like EfficientNetB0 and ResNet-101, recording 97.81% accuracy, 98.31% precision, and 98.10% F1-score. The results demonstrate the benefit of combining traditional and learned features to enhance diagnostic performance.

Sedigh et al. [103] bypassed data sparsity through the use of a generative adversarial network (GAN) to create additional skin lesion images for a low ISIC dataset consisting of only 97 images. Augmentation improved CNN classification accuracy to 71% from around 53%, reporting the usefulness of synthetic data to improve model performance.

Finally, Teodoro et al. [104] proposed EfficientAttentionNet, which is a hybrid model that leverages GANs for class balancing, a U-Net for mask region-of-interest, and an attention-augmented CNN backbone. Trained on a few datasets (ISIC 2020, HAM10000, BCN20000, and MKS), the method produced strong results—around 97.9% accuracy, 99.5% recall, and 94.5% precision—emphasizing the roles of data balancing and attention in skin lesion classification.

When assessing our model's performance, the most meaningful comparisons come from studies that also utilize the SLICE-3D dataset, in order to guarantee a fair comparison. One particularly relevant work is by Teymouri *et al.* [105], who implemented a two-stage strategy for skin lesion classification using SLICE-3D. They crafted an ensemble model combining a Light Gradient Boosting Machine (LGBM) with a Data-efficient Image Transformer (DeiT3), and incorporated a comprehensive preprocessing pipeline. This included stratified sampling, synthetic oversampling using SMOTE, and augmentation techniques such as random Gaussian blur, rotations, and flips to address the dataset's class imbalance and improve training quality. This rigorous, multi-phased approach resulted in a classification setup in which the ensemble model achieved a remarkable 89% accuracy and 90% recall in identifying malignant lesions, outperforming the individual constituent models.

In another related study based on the SLICE-3D dataset, Syed and Albalawi [106] presented a more advanced AI solution that combines 3D Total Body Photography (3D-TBP) with the use of a tailored convolutional neural network architecture. Their preprocessing workflow included actions like zooming, normalization, translation, and rotation, all for dataset diversification and model robustness. Using transfer learning in feature extraction and aggressive augmentation techniques, they pre-trained a CNN to study single lesion crops from 3D-TBP images. Evaluated against the ISIC 2024 test set, their model demonstrated well with a high diagnostic performance, achieving a true positive rate in excess of 80% and attaining a partial area under

the ROC curve (pAUC) around 85% With a focus on performance and efficiency, Pintelas et al. [107] proposed ensemble-based modeling that optimizes MobileNet variants. Through a two-stage "Expand-and-Squeeze" mechanism, different incarnations of MobileNet were generated and then pruned based on performance-oriented criteria. Trained on the vast ISIC 2024 dataset, the resulting MobileNet-HeX model achieved an accuracy of 89%, AUC of 90.5%, and geometric mean (GM) of 0.809. This demonstrates its effectiveness as well as its capacity to manage large class imbalances in vast lesion datasets.

3.2.2. Dataset Presentation

In our study, we leveraged the newly released SLICE-3D dataset (Skin Lesion Image Crops Extracted from 3D Total Body Photography) [41], which was first introduced in the ISIC 2024 challenge on Kaggle [108]. It includes 401,059 high-quality, standardized, and de-identified dermoscopic images, each annotated with a diagnostic label, 393 of them are malignant, while the remaining 400,552 are benign. Alongside the images, detailed metadata is provided, which we leveraged during the data cleaning and preprocessing stages. Figure 3.1 offers a visual summary of some of the key attributes contained in this dataset, while the following table shows in details all the metadata offerd in this dataset with the explination.

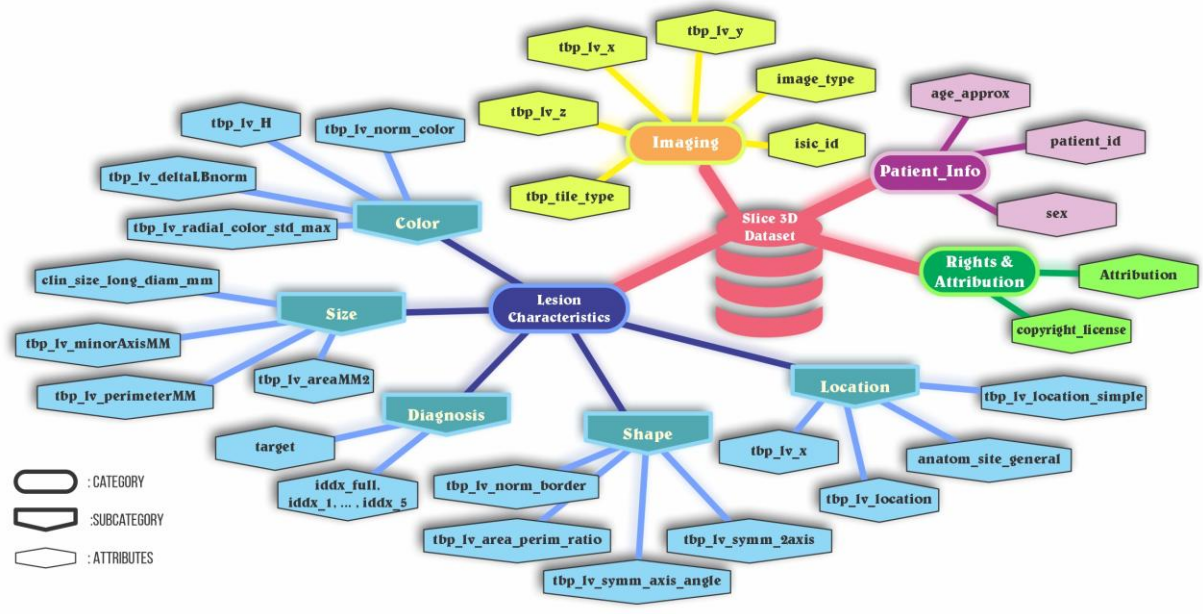


Figure 3-1: The Slice-3D dataset metadata classification (created by the author)

Table 3-1: Data element dictionary [41]

Field(s)	Description
isic_id	Filename without the file extension. This is the unique image ID from the ISIC Archive.
iddx_full, iddx_1, …, iddx_5	Fully classified lesion diagnosis: first- through fifth-level diagnosis.
attribution	Used for usage and licensing purposes; also indicates the medical center from which the image was sourced.
copyright_license	Creative Commons license attributed to the image.
patient_id	Unique patient identifier from the ISIC Archive.
age_approx	Approximate age of the patient at the time of imaging.
sex	Sex of the patient.
lesion_id	Unique lesion identifier from the ISIC Archive. Presence of this field indicates a manual tag.
anatom_site_general	ISIC Archive variable indicating the general anatomical location of the lesion*.
clin_size_long_diam_mm	Maximum diameter of the lesion (mm)*.
mel_mitotic_index	Mitotic index of invasive malignant melanomas.
mel_thick_mm	Thickness (depth) of melanoma invasion.
image_type	Structured ISIC Archive field describing the image type.
tbp_tile_type	Lighting modality of the 3D total body photography (TBP) source image.
tbp_lv_L, tbp_lv_A, tbp_lv_B, tbp_lv_C	L*, A*, B*, and chroma values inside the lesion*.
tbp_lv_Lext, tbp_lv_Aext, tbp_lv_Bext, tbp_lv_Cext	L*, A*, B*, and chroma values outside the lesion*.

tbp_lv_deltaL, tbp_lv_deltaA, tbp_lv_deltaB	Average L*, A*, B* contrast between the lesion and surrounding skin*.
tbp_lv_H	Hue inside the lesion, computed as the angle of A* and B* in L*A*B* color space (typical values: 25–75)*.
tbp_lv_Hext	Hue outside the lesion*.
tbp_lv_areaMM2	Area of the lesion (mm^2)*.
tbp_lv_area_perim_ratio	Border jaggedness, defined as the ratio between lesion perimeter and area (range: 0–10)*.
tbp_lv_color_std_mean	Color irregularity computed as the variance of color values within the lesion (range: 0–10)*.
tbp_lv_deltaLBnorm	Contrast between the lesion and immediate surrounding skin in L*A*B* color space (typical range: 5.5–25)*.
tbp_lv_dnn_lesion_confidence	Lesion confidence score estimated by a deep neural network (0–100 scale)*.
tbp_lv_Eccentricity	Lesion eccentricity*.
tbp_lv_Location	Detailed anatomical location classification (arms/legs subdivided; torso divided into thirds)*.
tbp_lv_location_simple	Simplified anatomical location classification*.
tbp_lv_minorAxisMM	Smallest lesion diameter (mm)*.
tbp_lv_nevi_confidence	Nevus confidence score (0–100), estimated by a CNN trained on dermatologist-labeled lesions*.
tbp_lv_norm_border	Normalized border irregularity score (0–10)*.
tbp_lv_norm_color	Normalized color variation score (0–10)*.
tbp_lv_perimeterMM	Perimeter of the lesion (mm)*.
tbp_lv_radial_color_std_max	Color asymmetry based on radial color variation (range: 0–10)*.

tbp_lv_stdL	Standard deviation of L* inside the lesion*.
tbp_lv_stdLExt	Standard deviation of L* outside the lesion*.
tbp_lv_symm_2axis	Border asymmetry score based on the second axis of symmetry (range: 0–10)*.
tbp_lv_symm_2axis_angle	Angle corresponding to the lesion's second symmetry axis*.
tbp_lv_x, tbp_lv_y, tbp_lv_z	3D coordinates of the lesion in total body photography*.

Most publicly available skin cancer image datasets are made up primarily of dermoscopic photographs, but they often suffer from limitations such as selection bias, a lack of standardization, and an overreliance on expert interpretation, which restricts their usability in broader clinical settings. These valuable datasets, generally focus on high-resolution dermoscopy images and are mostly used to train AI models meant to help dermatology professionals. In contrast, the dataset introduced in our work offers a new and innovative perspective. Developed by members of the International Skin Imaging Collaboration (ISIC a global initiative bringing together academia and industry), this dataset was specifically designed to support the development of open-access AI tools that can make diagnostic assessments from lower-resolution clinical images, similar to those captured by smartphones. Depending on the involved technology, Total Body Photography (TBP) can be used in either two-dimensional (2D) or three-dimensional (3D) mode, and visible light imaging is utilized to capture the skin surface.

In 2D-TBP systems, one commonly places a stationary camera frame with a single camera, requiring several images at different angles to sufficiently image the patient's skin. While this method is relatively affordable and has shown benefits in monitoring high-risk patients, this method has several obstacles. The requirement for readjustment of the camera repeatedly can make image acquisition difficult (especially in patients with mobility disorders) and lead to overlapping areas or missing lesions. The process is also long and needs large clinical resources, taking at least one hour per patient [109].

In contrast, 3D-TBP represents a more advanced and efficient alternative, designed to address many of the limitations inherent in traditional methods. Introduced around 2015, this system uses a rotating platform equipped with 92 synchronized high-resolution cameras and depth sensors to capture nearly the entire skin surface in a matter of seconds. Within approximately ten minutes, the system generates a highly detailed 3D digital model of the patient's body,

offering a comprehensive 360° skin view. This format significantly improves the capture of curved anatomical surfaces and provides a more complete skin assessment. However, some regions such as those covered by clothing, areas between the fingers and toes, the scalp, soles of the feet, scars, and tattoos may still be excluded from imaging. One of the major advantages of 3D-TBP is its integrated software, which accurately maps clinical and dermoscopic images to their precise anatomical locations on the digital avatar. This feature enhances the ability to track the emergence of new lesions or subtle changes over time, making it a powerful tool in long-term skin cancer surveillance [109].

TBP is a valuable tool that can improve the sensitivity for finding early malignant melanoma, particularly in high-risk patients, via the detection of changes noted during follow up visits, The ability to compare the patients' nevi to a set of baseline images obtained at a prior point in time allows clinicians to utilize the comparative recognition process to identify new or subtly changing lesions that may be an early melanoma. Simple baseline TBP plays an important role in the early detection of skin cancer while at the same time reducing the unnecessary removal of clinically and biologically benign lesions [110].

The following figure illustrates the Progression of a Stable Naevus to Melanoma in Situ, where is:

- (A) The initial 3D total body mapping of naevi at the first visit.
- (B) Digitally magnified 3D total body views of the corresponding naevus at the first, second, and third visits, showing sequential changes.
- (C) Dermoscopic images of the same naevus at each follow-up visit.

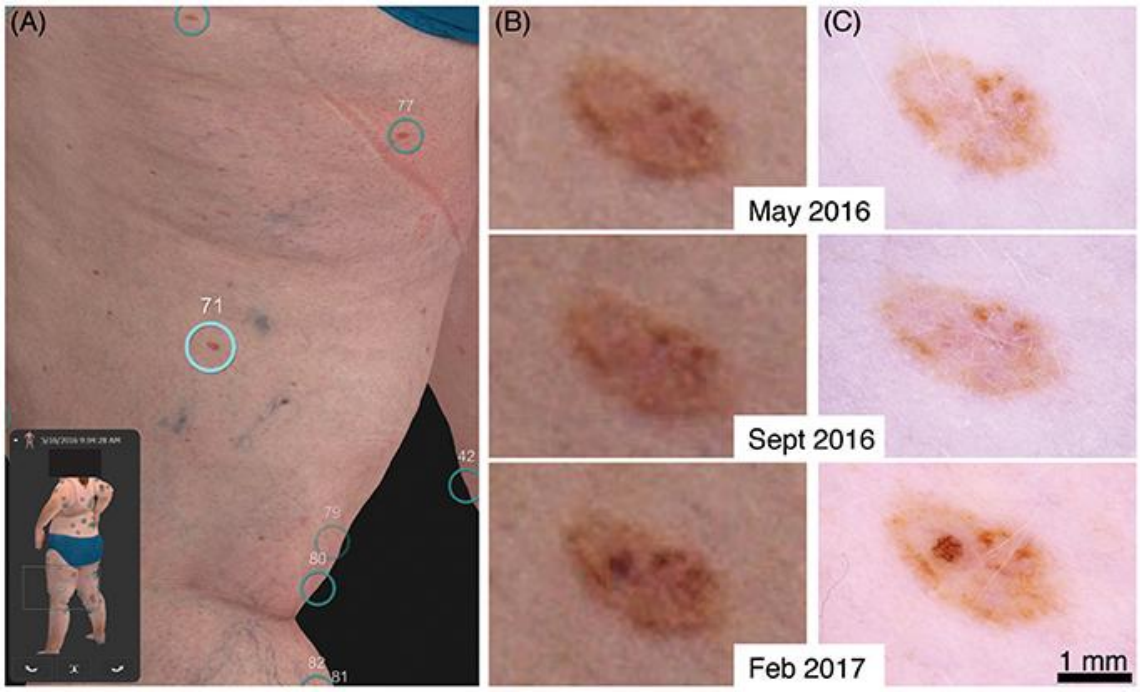


Figure 3-2: Evolution of a stable naevus to a melanoma in situ [111]

TBP involves a specially trained photographer (melanographer) taking clinical photographs of the entire skin surface, which are stored for future reference. These images can be stored digitally, on traditional printed photographs, or via mobile apps, and may also be provided to patients for their own records. The primary purpose is to compare current and past images to monitor for new or evolving lesions.

TBP is mainly conducted using 2D photography, with some centers offering 3D imaging. The procedure is painless and non-invasive, taking approximately 10-15 minutes in a photography studio. Patients are asked to remove all clothing except underwear to ensure maximum skin visibility. Before the session, consent is obtained, and images are securely stored following protocols at the treating center. Depending on individual risk factors, images may be taken annually or only once as a baseline for future comparison [112].

The SLICE-3D dataset includes individual skin lesion images collected from more than 1,000 patients seen between 2015 and 2024 across seven high-risk dermatologic centers on three continents. Unlike previous datasets that focused on isolated lesions, this one provides a holistic view of each patient's skin, aiming to improve algorithms by incorporating broader context in skin cancer detection. Importantly, all images were captured using a consistent imaging device and standardized field of view, which helps reduce the variability often caused by differences in lighting, camera settings, or color balance. Each image, or "tile", covers a 15mm-by-15mm area and is centered on a lesion—either automatically detected if larger than 2.5mm or manually annotated regardless of size—with an average resolution of 133x133 pixels. Accompanying metadata files link each tile to its source 3D image and include information generated by automated lesion detection software as well as manual clinical annotations, allowing for detailed patient-level analysis.

Overall, this dataset has been carefully structured to support the analysis of skin lesion images, while accounting for potential variability introduced by differences in lighting conditions and patient positioning during image capture. However, researchers should remain aware of certain limitations—particularly the lack of detailed documentation regarding skin tone and the dataset’s unique imaging characteristics—which may influence the generalizability of their findings. These factors should be thoughtfully considered when interpreting results or designing experimental protocols based on this data.

3.2.3. Dataset Preprocessing

Before beginning any modeling attempts, a comprehensive data preprocessing and reduction procedure was necessary due to the dataset's notable class imbalance, which included over 400,000 benign photos compared to just 393 malignant instances. We used a number of systematic preprocessing procedures to overcome this discrepancy and guarantee a more controllable and representative dataset. These allowed us to down sample the benign class to 4084 carefully selected images, while preserving all malignant images, due to their limited number and critical diagnostic value.

Using the binary target attribute (target), where 0 denotes a benign lesion and 1 a malignant one, we first separated the dataset into two subsets. where 0 indicates a benign lesion and 1 a malignant one. From that point forward, we focused exclusively on the benign class for data refinement with one goal: Reduce the number of benign images, intentionally leaving the malignant samples untouched to avoid further reducing their already low representation.

The cleaning process began by eliminating rows that contained missing values in a set of key columns considered essential for downstream analysis and model performance. These columns included age_approx, sex, lesion_id, anatom_site_general and tbp_tile_type. Subsequently, columns with more than 90% missing values were discarded, as they offered little informative value and introduced noise. We also removed duplicate entries, retaining only the first occurrence per patient and anatomical region to ensure data independence and avoid model bias.

One of the most impactful filtering steps we applied was selecting images captured under a one specific lighting modality, the attribute “tbp-tile-type” was responsible for this value and among its 2 values “3D: white” and “3D: Xp” we chose columns identified as '3D: XP'. This choice was motivated by the fact that lighting conditions represent a major technical variable influencing color perception—affecting hue, tone, and contrast—even among otherwise standardized images. Keeping this modality constant, we were then able to keep more homogeneous image features within samples, which is particularly important in medical imaging tasks that consist of intricate analysis of skin texture and pigmentation.

Ultimately, this rigorous cleaning pipeline reduced the benign dataset from over 400,000 to 4084 images, aligned with the '3D: XP' lighting modality, while retaining all 393 malignant cases. This curated and balanced subset forms the basis of our experiments and ensures a more robust and controlled training process. A detailed breakdown of the data cleaning workflow is illustrated in Figure 3.3

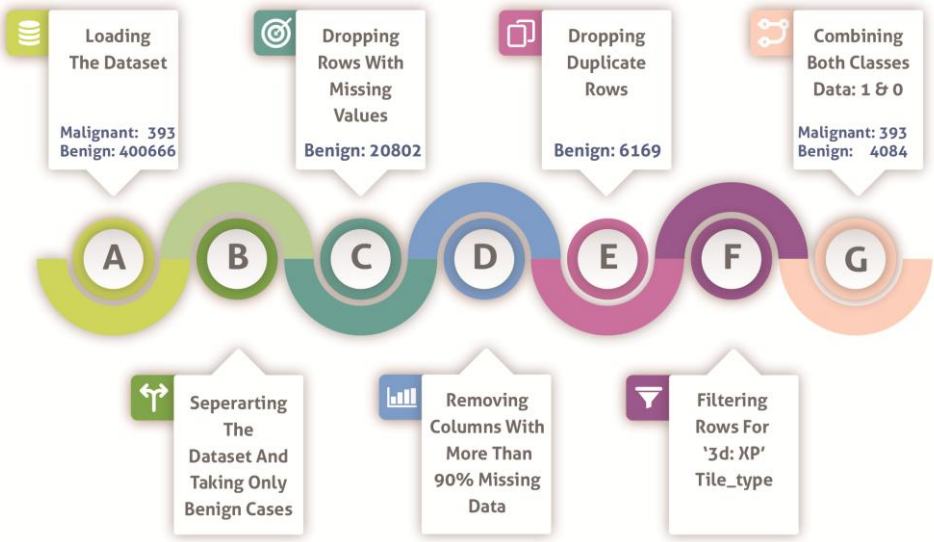


Figure 3-3: Data preprocessing steps

3.2.4. Data Augmentation using DCGAN

Now that our benign data is manageable, we proceeded to the data balancing, where at this step the number of benign images is equal to 4084 images and malignant images are equal to 393. There are several approaches to handle such data imbalance, like geometric transformations, mixing images, kernel filters, random erasing, feature space augmentation, color space augmentations, adversarial training, neural style transfer, generative adversarial networks, and meta-learning [113]; But we chose to use data augmentation via Deep Convolutional Generative Adversarial Networks (DCGAN) which is a widely adopted approach for synthetic image generation in medical datasets, as it was proven to handle datasets imbalance very well. Figure 3.4 depicts a taxonomy of image data augmentation techniques according to Shorten et al. [113]

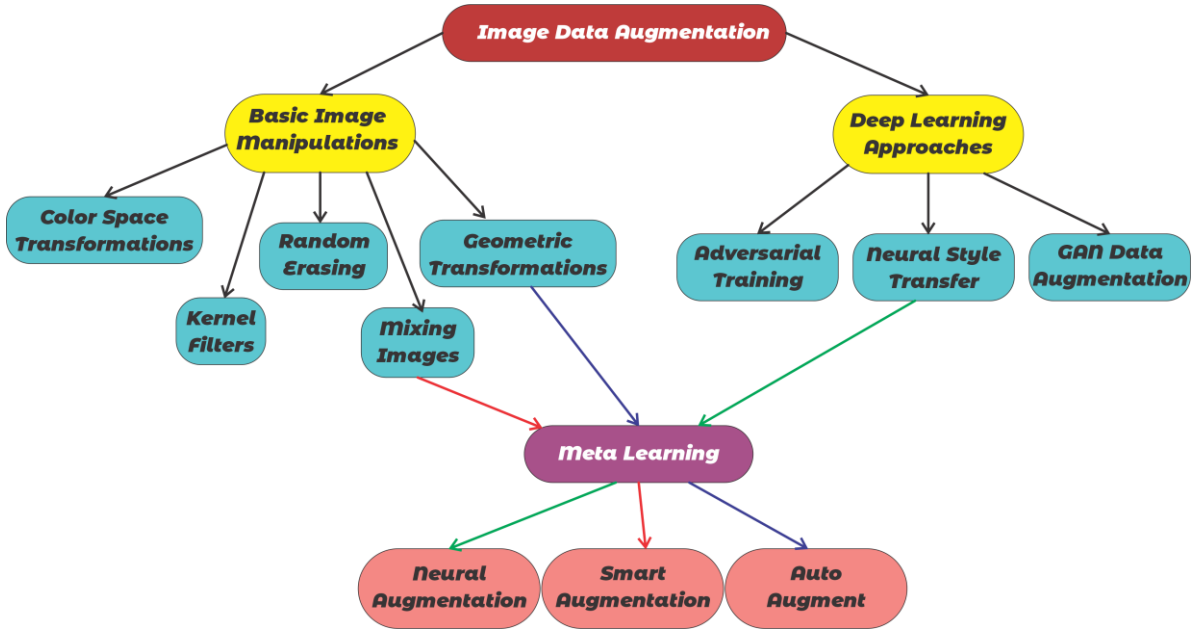


Figure 3-4: Data augmentation techniques [114] [113]

In the 2016 paper, Radford et al [115] define Deep Convolutional Generative Adversarial Networks as a class of Convolutional Neural Networks (CNNs) with specific architectural constraints, proving their effectiveness in unsupervised learning; it is a subclass of Generative Adversarial Networks (GANs) that leverage convolutional layers, batch normalization, and specific activation functions to generate high-quality images.
Our proposed DCGAN architecture was designed to generate high-quality synthetic dermoscopic images for the malignant class, addressing the dataset imbalance. The model consists of two adversarial networks: a generator and a discriminator, trained in opposition to one another.

The generator takes as input a 224-dimensional latent noise vector sampled from a standard normal distribution and progressively up samples it through a sequence of five transposed convolutional layers. Each layer is followed by batch normalization and ReLU activation, except for the output layer, which applies a Tanh function to produce RGB images normalized in the range [–1, 1] and of spatial size 64×64 pixels.

The discriminator serves as a binary classifier that distinguishes real images from generated ones. It comprises five convolutional layers with LeakyReLU activations (α=0.2) and batch normalization (except in the first layer), progressively reducing the spatial resolution until a single scalar output is obtained through a Sigmoid activation.

To enhance robustness and diversity, the training set of 393 malignant lesions was augmented with random horizontal and vertical flips and normalization following VGG16 preprocessing statistics. Both networks were optimized using the Adam optimizer with a learning rate of 0.0002 and parameters β_1=0.5, β_2=0.999. The adversarial training was conducted for 300

epochs, with losses monitored to ensure stability and to avoid mode collapse. Synthetic samples were saved every ten epochs for qualitative evaluation. The following figure represents the architecture of the DCGAN used in our code:

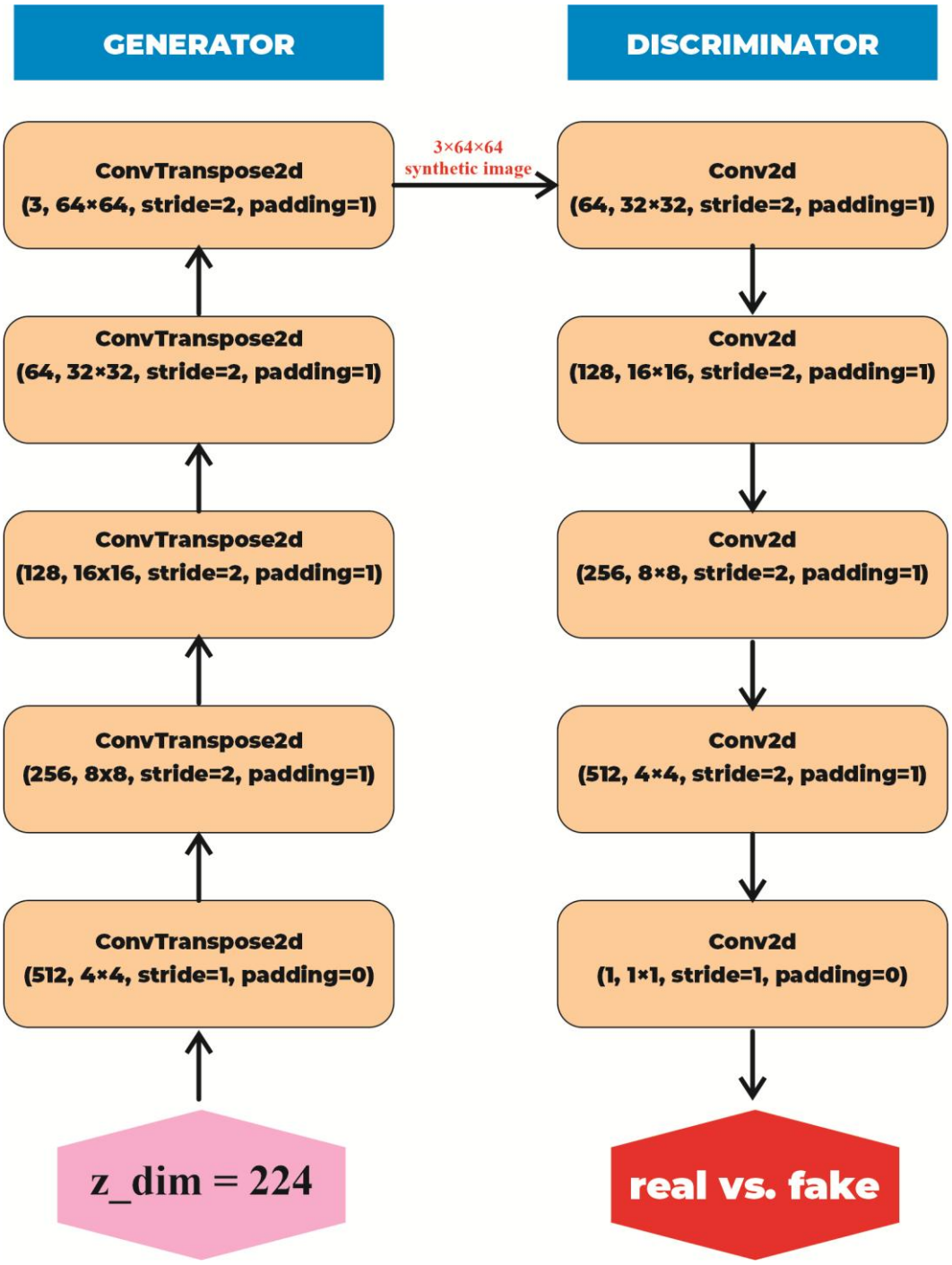


Figure 3-5: Architecture of the DCGAN used for malignant lesion image synthesis

Following the conclusion of training, we employed the generator to create 3,691 synthetic malignant photos, thereby achieving a balanced dataset by augmenting the minority class. These artificially generated images exhibited visual fidelity and preserved the morphological

characteristics of real melanoma lesions. Figure 3.6 presents examples of the generated images, showcasing their resemblance to the original samples used for model training.

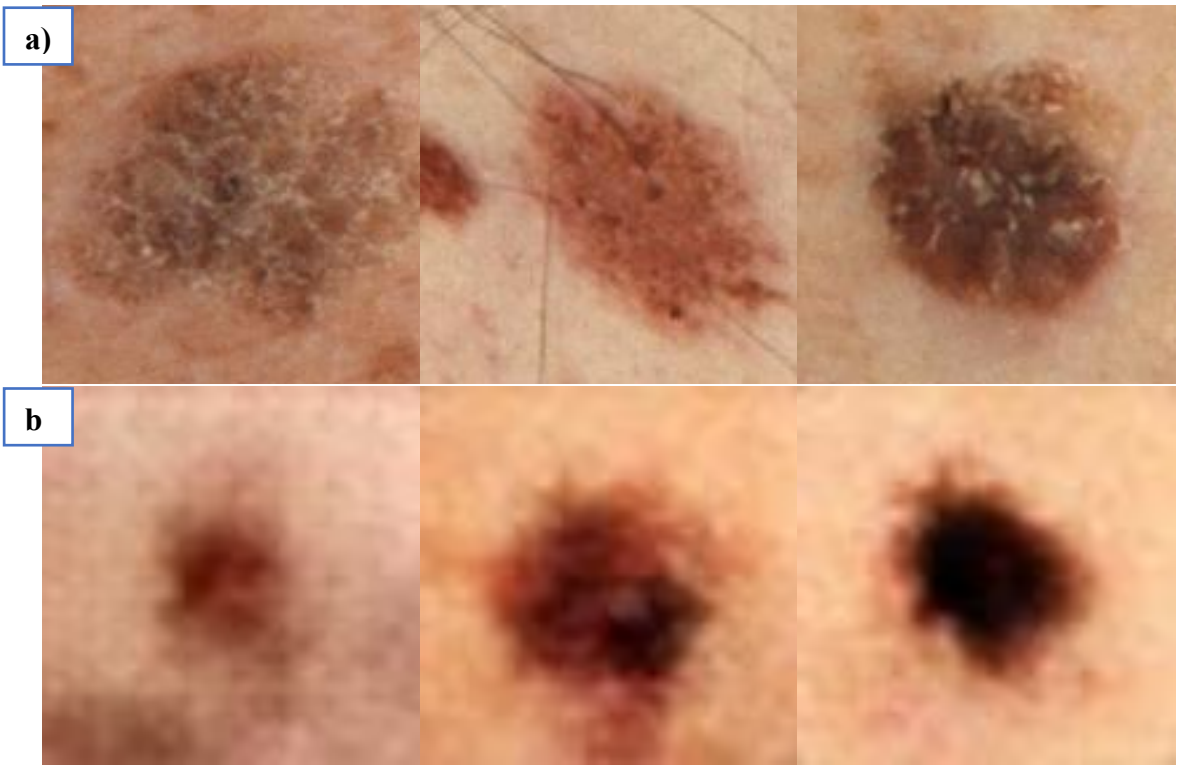


Figure 3-6: Visual differences between real vs synthetic skin lesions images a: original lesion images / b: generated lesion images)

3.2.5. Dataset Splitting

Following appropriately the class balance establishment—having the same number of benign and malignant cases (4,084 images per class for a total of 8,168 images)—we proceeded with a systematic partitioning of data to enable robust and unbiased model estimation. The collection was first split into two main subsets: 80% to be used for training and 20% to be reserved for the test set, which translates to 6,534 images for training and 1,634 images reserved for final performance evaluation.

To further fine-tune the learning process and prevent overfitting, we subdivided the training portion by allocating 85% of it to actual model training and 15% to validation. This resulted in 4,900 images for training and 1634 images for validation. The validation set was used exclusively during training to monitor generalization performance and guide hyperparameter tuning, such as learning rate adjustments, early stopping criteria, and model checkpointing. Table 3.2 demonstrates the exact number of images used for each phase.

Table 3-2: Data splitting values

Dataset 100%	Benign	Malignant	Total
Training 60%	2,450	2,450	4,900
Validation 20%	817	817	1,634
Testing 20%	817	817	1,634

This three-way data split (training, validation, test) was stratified carefully so that each subset would retain the same distribution of classes—upholding the 1:1 ratio of benign and malignant samples. This stratification is important in medical datasets, particularly when dealing with synthetic data or otherwise imbalanced sources, to prevent model evaluation bias as well as to uphold clinical significance

3.2.6. Transfer Learning Application with Custom Architectures: Modified VGG16 and MobileNetV2

CNN is a deep learning image and video processing algorithm. A CNN consists of several layers: an input layer, hidden layers (for feature extraction), a fully connected layer (for classification), and an output layer. As deep learning models evolved quickly, in practice, various pre-trained classification models (previously discussed in Chapter 1) are used for various datasets, for instance, AlexNet, VGG, GoogLeNet, Tom, Inception, ResNet, and MobileNet, for disease classification.

With our dataset prepared, we now advance to the final phase: training. we Leveraged transfer learning using two convolutional neural network architectures: VGG16 and MobileNetV2. Both networks were pre-trained using the ImageNet weights; plus, we further fine-tuned them on our binary skin cancer classification task (benign/malignant) with architectural and training modifications.

The experiments were carried out on the Kaggle platform, which offers a powerful and widely available machine learning development platform. Kaggle offers access to GPU-boosted runtimes at no cost, and it is particularly geared towards training computationally heavy deep learning models. Model building was done through the use of Python with the PyTorch framework being selected for its extensibility, dynamic computation graph, and abundant support for cutting-edge neural network architectures.

- ***Modified VGG16 Architecture***

We employed a modified version of VGG16 **architecture** originally proposed by Simonyan and Zisserman [76], renowned for its deep convolutional hierarchy and large representational ability. The original VGG16 architecture that was initially created to perform large-scale image classification was particularly adapted to our needs; the convolutional feature extractor of the

model was retained, and batch normalization layers between each convolution were added for training stabilization and acceleration.

The classifier section was redesigned into a more compact configuration composed of:

- An adaptive average pooling layer to aggregate spatial features,
- A flattening layer,
- A fully connected layer with 128 neurons and ReLU activation,
- A dropout layer (p = 0.6) for regularization,
- And a final linear layer producing a single output for binary classification.

The model was trained using the Adam optimizer (learning rate = 1e-4) with BCEWithLogitsLoss, and early stopping was employed to mitigate overfitting. This modified architecture achieved robust performance while maintaining a balance between model complexity and generalization ability.

- *Modified MobileNetV2 Architecture*

As a light-weight and computationally economical option, we modified MobileNetV2, introduced by Sandler et al. [82], which is based on depth wise separable convolutions and inverted residuals.

In our implementation, the entire convolutional feature extractor (features block) was frozen to preserve the pretrained feature representations. A custom classifier head was then appended, consisting of the following components:

- An adaptive average pooling layer to aggregate spatial features into a compact representation.
- Two fully connected layers with 1024 and 512 neurons, each followed by batch normalization, ReLU activation, and dropout (p = 0.5) to enhance regularization and prevent overfitting.
- A final linear output neuron producing the logits for binary classification.

Training was performed using the Adam optimizer (learning rate = 1×10^{-4}) and the BCEWithLogitsLoss function. This adaptation retained MobileNetV2's efficiency while significantly improving its discriminative capability for skin lesion classification.

The 2 network architectures have been summarized respectively in Figures 3.6 and 3.7 below.

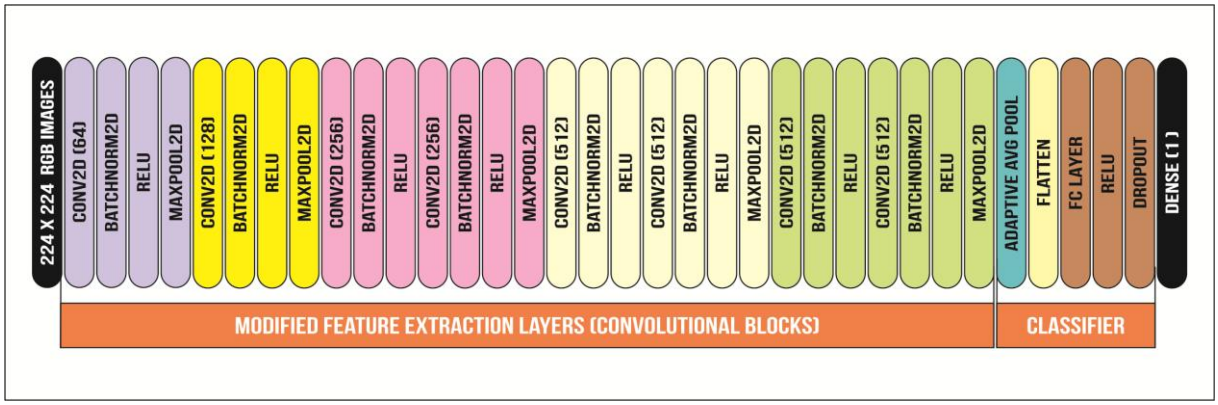


Figure 3-7: Architecture or the VGG-16-based model with transfer learning

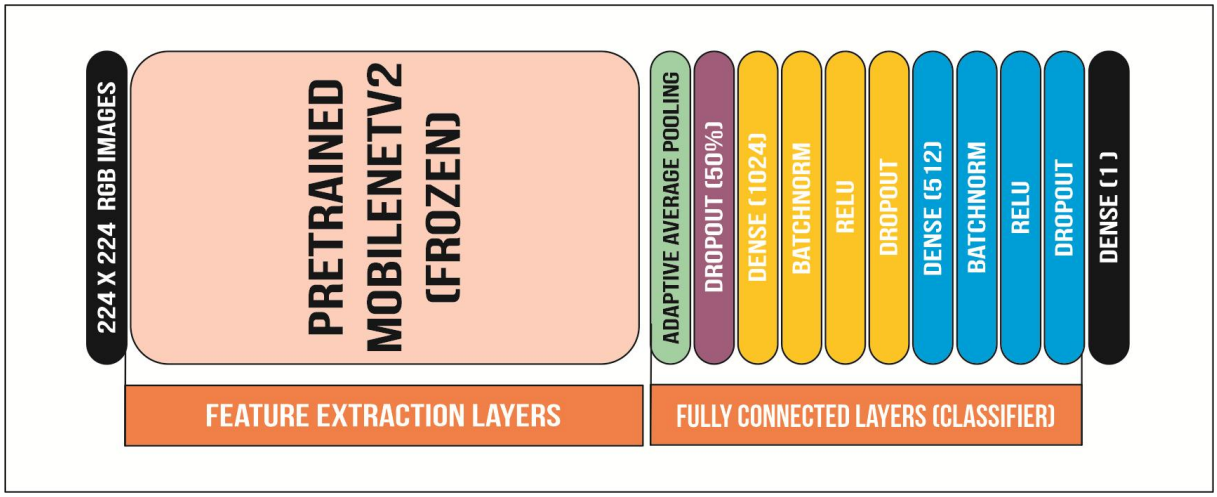


Figure 3-8: Architecture or the MobileNetV2-based model with transfer learning

Despite these variations, the two models share the same input shape (224x224x3), the same cleaned-augmented dataset, and the same overall training methodology to enable a fair comparison of efficiency and performance.

3.3. Results and discussion

Training Convolutional Neural Networks (CNNs) requires huge amounts of data and is a computation-heavy process. In order to handle this, high-performance computational resources—more specifically Graphics Processing Units (GPUs)—are required. In this research study, the experiments were executed on the Kaggle platform, which offers the facility of a free GPU-accelerated environment. This not only speeds up the training process incredibly but also enables larger datasets to be handled, leading to greater accuracy and model performance.

Owing to pragmatic constraints of computation time and resource availability, training was restricted to a maximum of 100 epochs. Nevertheless, the use of GPU acceleration provided an

opportunity for quicker convergence and more effective experimentation even within this limited training period. Greater classification accuracy could possibly be achieved by increasing the training epochs or utilizing more advanced hardware.

Our experiments relied on two popular pre-trained CNN architectures: VGG16 and MobileNetV2, implemented in PyTorch. The two models were chosen due to their established performance in image classification and for being suitable for transfer learning, a particular benefit when dealing with comparatively small datasets. Both architectures take RGB images resized to 224x224 pixels as input, which makes them highly compatible with standardized image preprocessing pipelines.

Hyperparameter optimization is a crucial part of CNN training. In our experiments, significant parameters were determined empirically through trial and error to realize stable and efficient learning. The batch size was 32, a compromise between stability during training and convergence speed. We chose to put the learning rate value to 0.0001, as it is a very popular value that favors slow learning and discourages overshooting during optimization.

We used the Adam optimizer to train the model, as it possesses adaptive learning rate characteristics and is effective in handling sparse gradients. For the loss function, we employed BCEWithLogitsLoss, a standard choice for binary classification problems. Implementations were carried out in Python using the PyTorch deep learning framework, which allows flexibility and modularity in building and experimenting with models.

3.3.1. Training and validation result

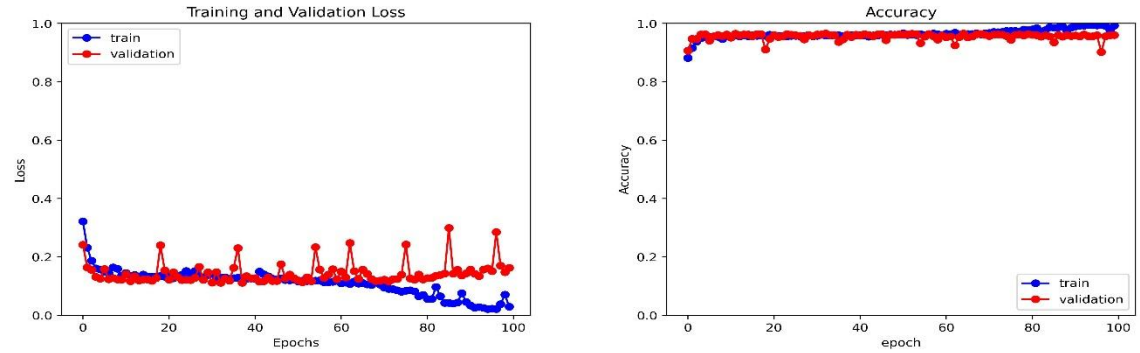
The experimental outcome verifies that the MobileNetV2-based model was a bit superior to its VGG16-based counterpart. Figure 3.9 illustrates the training and validation history of the optimized models, accuracy and loss.

Since other metrics of importance like test loss, accuracy, precision, recall, and AUC were also determined, the initial model that used a pre-modified VGG16 architecture recorded a test accuracy of 96.20%, which is a high overall classification rate. An indication of a comparatively small prediction mistake is provided by a test loss of 0.113. Furthermore, a 97.4% accurate precision suggests a minimal number of false positives, and a recall of 94.8% indicates the model's ability to classify the overwhelming majority of true positives accurately. This is followed by an AUC value of 0.99 that also supports the model's extremely high ability to distinguish between the two classes.

The second model, with the MobileNetV2 architecture, was also tried on the same test set. The test loss was 0.0926, which represents a slightly lower prediction error and confirms the strength of the model. Its strong capability to correctly classify cases is also accompanied by a test precision that was noted at 96.88%. The accuracy of the model at 98.98% and 94.74% recall further signify that it strikes an amazing balance between predicting positive cases correctly

and avoiding false positives and false negatives. The 99.13% AUC also shows to what extent the model can differentiate between classes.

a)



b)

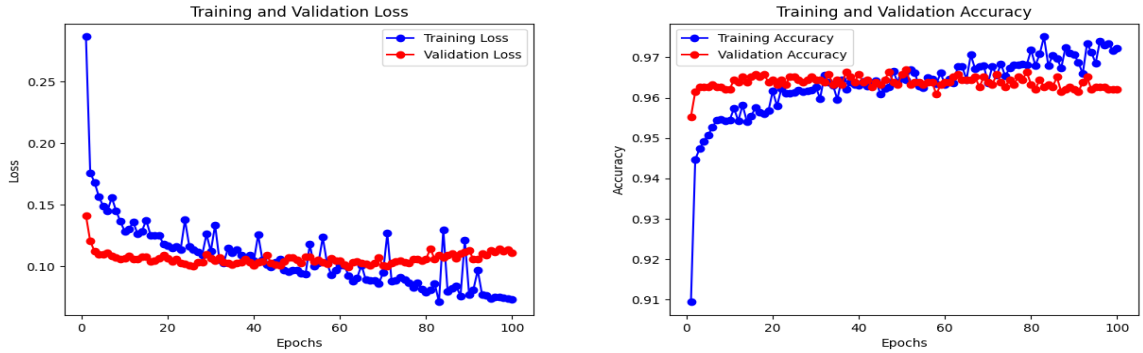


Figure 3-9: The accuracy & loss functions vs epochs of: a) VGG-16, b) MobileNetV2 based models

In summary, both models' performance was highly impressive; however, for loss, accuracy, and precision, the MobileNetV2-based model outperformed that of VGG16. Additionally, the highly impressive classification accuracy of both models is evidenced by the fact that AUC scores were highly impressive.

3.3.2. Testing and Evaluation

The following section discusses the evaluation process employed to determine the performance of the model, with the respective results.

Confusion matrix is a vital machine learning method for estimating key evaluation and testing measurements, such as accuracy, specificity, recall (sensitivity), AUC, and precision. It gives both a graphical and quantitative way of assessing the performance of the two classification

models based on VGG16 and MobileNetV2 in classifying benign (negative) and malignant (positive) skin lesion images.

Correctly classified samples are displayed in blue, and misclassified points appear in yellow in Figures 3.9 and 3.10. These plots, along with the corresponding confusion matrices, clearly show the excellent classification performance of both models.

Particularly, VGG16-based model possesses powerful classification ability. As the confusion matrix of Figure 3.10 shows, it correctly classified 97.6% of benign cases and 94.9% of malignant cases, with misclassification rates of 2.4% and 5.1%, respectively. Which is a very good balance of sensitivity and specificity, and it is critical in medical diagnostic situations where false positives and false negatives can cause serious problems. Similarly, the MobileNetV2 model also demonstrated slightly higher accuracy, as it correctly identified 99% of benign and 94.7% of malignant samples. The proportion of false positives was limited to 1%, and that of false negatives to only 5.3%. This impressive high detection precision rate for malignant lesions is especially beneficial because this indicates very high reliability on the model's part in marking potential skin cancers.

In general, it is noteworthy that both models were efficient, with MobileNetV2 showing a better improvement in benign classification, but maintaining comparable sensitivity to VGG16 for malignancy detection. These results confirm the effectiveness of transfer learning in binary skin lesion classification

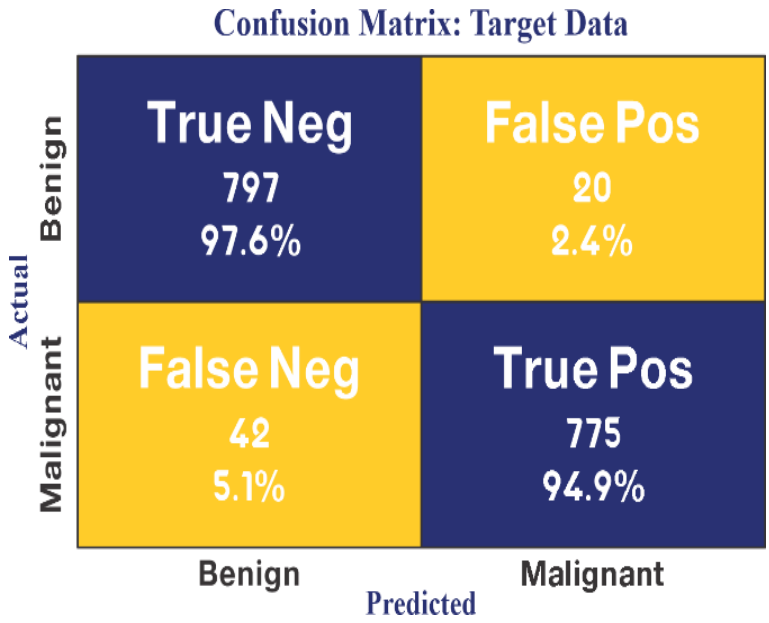


Figure 3-10: Confusion matrix of VGG-16-based model

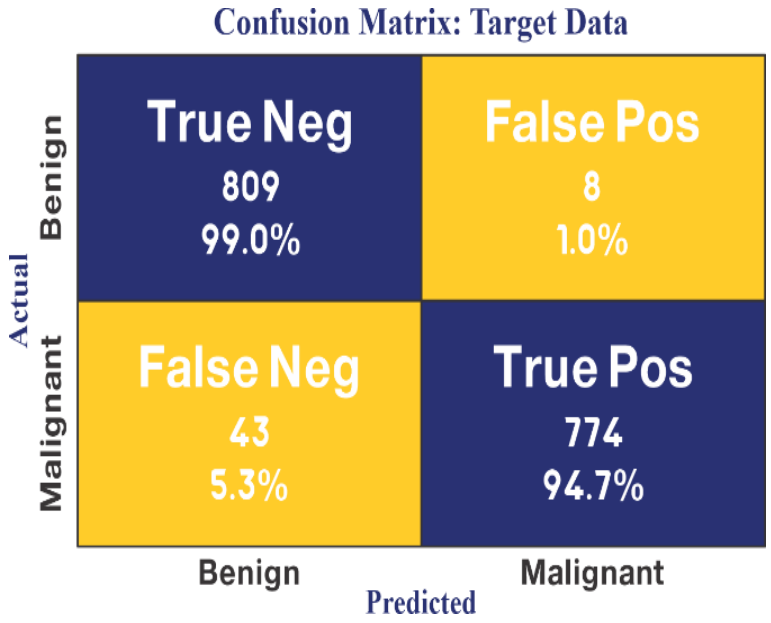


Figure 3-11: Confusion matrix of MobileNetV2-based model

Furthermore, the test results, presented in terms of performance metrics in Figure 3.12, reveal the closely matched performance of both models. This further demonstrates the effectiveness of transfer learning in the context of binary skin cancer classification

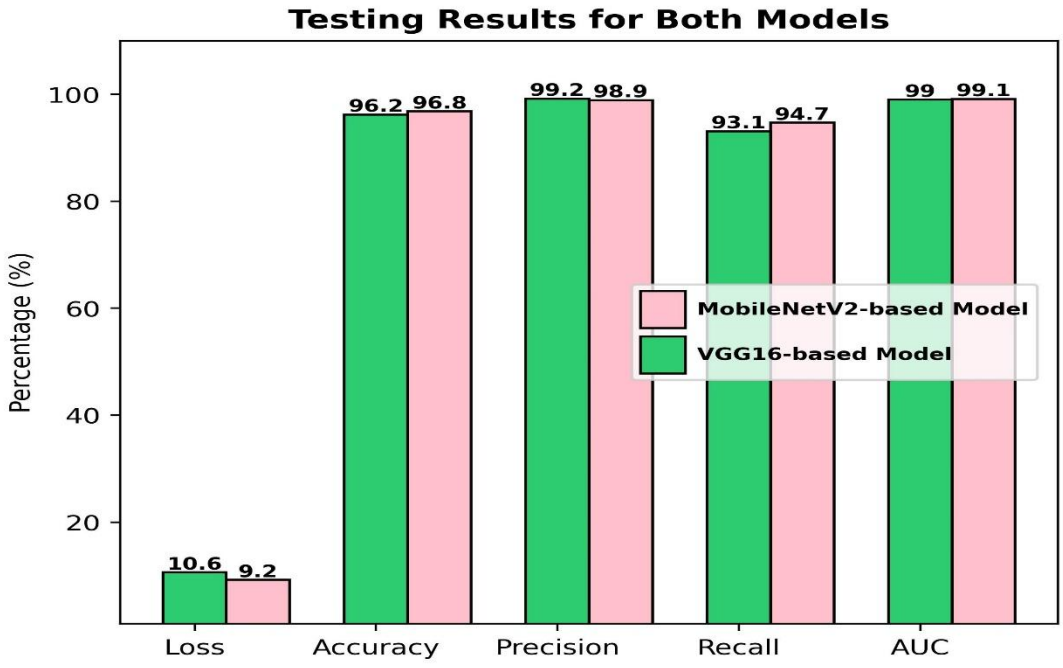


Figure 3-12: Testing results of our proposed models: VGG16-based model vs. MobileNetV2-based model

In order to adequately assess our models, we categorized their test findings in Table 3.3 along with the previously reported state-of-the-art results in the context of our selected ISIC-2024 dataset.

Table 3-3: State-of-the-art Models and their Test results Vs Our 2 models

MODEL	ACCURACY (%)	PRECISION (%)	TRUE POSITIVES (%)	AUC (%)	RECALL (%)
PINTELAS ET AL [107]	*89*	-	-	*90*	-
SYED & ALBAWI [106]	*87.5*	-	*80*	*85*	-
TEYMOURI ET AL [105]	*89*	-	-	-	*90*
BASED ON VGG-16	***96,20***	***99,2***	***93,1***	***99***	***93,1***
BASED ON MOBILENETV2	***96,87***	***98,9***	***94.7***	***99,1***	***94,7***

3.4. Conclusion

In this chapter, we investigated the significant role of deep learning combined with transfer learning in the early diagnosis of skin cancer, with a particular focus on leveraging transfer learning and advanced data augmentation techniques to enhance model performance. Giving the lack of balanced and rich skin cancer datasets, our research was amongst the firsts, if not the best, to improve the quality of the novel SLICE-3D dataset provided by the ISIC-2024 challenge and make it suitable for AI models training, we tackled the issue of severe class imbalance—a common challenge in medical datasets—by integrating a Deep Convolutional Generative Adversarial Network (DCGAN). This approach enabled the generation of synthetic malignant samples, contributing to a more equitable distribution of classes and improving the representativeness of the dataset.

We implemented and evaluated two widely recognized convolutional neural network architectures, VGG16 and MobileNetV2, both of which were adapted through architectural enhancements aimed at improving their generalization and predictive accuracy. Experimental results confirmed that both models performed effectively on the binary classification task, achieving high levels of accuracy. Among the two, the MobileNetV2-based model exhibited slightly better performance in key metrics such as area under the curve (AUC), recall, and precision, while also offering computational efficiency. This balance between accuracy and resource consumption suggests that MobileNetV2 could be particularly suitable for practical deployment in clinical environments where real-time processing and limited hardware are considerations.

Our study focuses on the importance of training deep learning models for medical image analysis on carefully balanced, high-quality datasets. The results further illustrate how transfer

learning, when combined with careful architectural tuning, can yield robust models that can confront the difficulties that are inherent to dermatological diagnosis.

Down the line, further improvements can be done by exploring more advanced techniques such as ensemble methods, model interpretability frameworks, or the integration of complementary data modalities (e.g., clinical metadata or dermoscopic features). Such improvements can result in even more precise and generalizable diagnostic systems, which can contribute to early detection and treatment planning in dermatology.

In the next chapter, we will go deeper into the field of early skin cancer detection, where we will explore different factors that can affect the robustness of the skin cancer model.

General conclusion

Skin cancer is a growing global health concern and remains a major global health challenge due to its aggressive nature and the complexity of early diagnosis. Early diagnosis is the secret to improved patient outcomes, especially when certain subtypes, such as nodular melanoma or atypical nevi, continue to pose classification difficulties even for experienced clinicians. The increasing worldwide incidence of skin cancer highlights the pressing need for accurate, accessible, and early diagnostic tools that can support clinical decision-making and improve patient outcomes.

In recent years, artificial intelligence (and more specifically, deep learning) has emerged as a powerful ally in medical imaging. Within dermatology, deep convolutional neural networks (CNNs) have shown great promise in assisting with skin lesion analysis and improving diagnostic performance. The motivation behind this thesis stems from the goal of enhancing public health through the development of intelligent, efficient, and cost-effective classification tools that harness technological innovation.

Contributions Synthesis

Our work is aimed to improve the field of computer-aided dermatological diagnosis by exploring and optimizing deep learning-based classification systems for skin cancer, with specific focus on binary discrimination between malignant and benign lesions. Through a multi-faceted approach grounded in real-world imaging challenges and modern neural network architectures, we addressed critical limitations in dataset quality, class imbalance, captured images configuration bias, and diagnostic variability.

We took advantage of the SLICE-3D dataset, a recent dermoscopic image database revealed in the ISIC 2024 challenge. Among the core challenges inherent in this dataset; and indeed most clinical imaging databases; was the extreme class imbalance with malignant samples being an incredibly small minority. To offset this, we employed a DCGAN-based image synthesis pipeline that successfully augmented malignant cases to balance the dataset. The synthetic images, when validated by both visual quality and downstream classification performance, were found to be an effective method for improving data diversity without compromising quality.

Yet another notable aspect of this thesis was the particular focus devoted to images' lighting modality, which is usually given short shrift but has quite a high degree of influence in dermoscopic imaging especially after researching and finding only few AI studies in this domain, they were all purely medical researchers conducted with dermatologists and not AI models. So, with the assistance of metadata fields such as tbp_tile_type, we investigated this relatively unexplored area and separated and tested independently three lighting conditions: non-polarized white light (3D:white), Cross-polarized light (3D:XP), and an intermediate set

combination. This allowed us to explore systematically the effect of light on automated diagnostic model performance; a study that, as far as we know, had not been previously investigated with such extent of detailed comparison in Deep learning filed.

Four pre-trained convolutional neural networks, DenseNet121, ResNet50, MobileNetV2, and EfficientNet-B0, were tested to perform binary classification on the preprocessed datasets. The experimental results consistently demonstrate that dermoscopic lighting conditions have a decisive influence on CNN-based skin cancer classification performance. Across all evaluated architectures, non-polarized (white-light) dermoscopy yielded the highest Recall, Specificity, and AUC, confirming its superiority for detecting malignant lesions while maintaining reliable benign exclusion. This finding is clinically significant, as high Recall is essential for screening-oriented diagnostic systems.

Among the tested models, EfficientNet-B0 achieved the best overall balance between accuracy and discriminative performance, particularly under white-light illumination, while MobileNetV2 and ResNet50 exhibited exceptionally high Recall and Specificity, making them well suited for sensitivity-critical clinical scenarios. These results align with dermatological knowledge, as white-light dermoscopy enhances superficial epidermal and pigmentation-related structures that are highly informative for malignancy discrimination.

The key contributions of the thesis are:

- Successful implementation of DCGAN to address class imbalance in our clinical dataset and crafting a transfer learning model that achieved a very high and impressive level of accuracy.
- Benchmarking dermoscopic lighting conditions (XP vs. White Light) on DCNN performance.
- Methodologically rigorous comparison of multiple architectures across balanced datasets, supported with good metrics.
- Demonstrating that metadata-aware experimentation can produce more clinically valuable models.

Overall, this work emphasizes the power of deep learning especially when combined with synthetic data generation and imaging metadata to improve early detection of skin cancer.

Future Research Directions and Perspectives

The findings and contributions of this thesis open several promising avenues for further research aimed at advancing deep learning methodologies for reliable and clinically applicable skin cancer detection:

- **Exploration of More Advanced Generative Models for Data Augmentation:** While this thesis investigated DCGAN for synthetic image generation, future work

could explore other GANs like StyleGAN or diffusion-based generative models and hybrid architectures. These approaches may yield higher-fidelity lesion synthesis, more diverse feature representation, and improved generalization when integrated into training pipelines.

- **Multi-Modal and Cross-Domain Learning:**

  Extending the analysis beyond dermoscopy images by incorporating additional modalities such as clinical photographs, histopathology, or metadata (e.g., patient age, lesion location) could significantly enhance diagnostic robustness. Research into multi-modal fusion strategies and domain adaptation techniques would support broader generalizability across institutions and imaging devices.

- **Systematic Investigation of Lighting Modalities:**

  This thesis highlighted the impact of lighting conditions (white light, cross-polarized, mixed) on model performance. Future work should expand this study by including standardized multi-center datasets, quantifying robustness under diverse acquisition protocols, and developing adaptive models capable of automatically normalizing or correcting lighting variations.

- **Explainability and Clinical Trustworthiness of Models:**

  Future studies should place stronger emphasis on explainable AI (XAI) methods tailored to dermatology. Designing interpretable saliency maps, lesion-specific feature attribution, or prototype-based reasoning can help bridge the gap between model predictions and dermatologists' diagnostic reasoning, thus enhancing trust in clinical settings.

- **Transferability and Federated Learning Across Institutions:**

  To overcome data scarcity and privacy concerns, research into federated learning and privacy-preserving training approaches for skin lesion datasets is recommended. This would allow collaboration across hospitals without direct data sharing, while ensuring that models remain robust and clinically validated across diverse populations.

- **Robustness under Data Shifts and Longitudinal Generalization:**

  Future investigations should examine how models behave under domain shifts such as population diversity, new lesion subtypes, or evolving imaging technologies. Research into continual learning and model adaptation will be crucial for ensuring that AI systems maintain high diagnostic accuracy in dynamic real-world clinical environments.

Overall, this work emphasizes the power of deep learning, especially when combined with synthetic data generation and imaging metadata, to improve early detection of skin cancer. As

encouraging as the results are, there is potential future work that can enhance diagnostic aid systems.

List of Published Work

1. **International Journal Article**

- Mecifi, Y.Z.., Merzoug, M., Etchiali, A., M'hamedi, M., Hadjila, F. & Bekkouche, A. (2025). Early Skin Cancer Detection Using the SLICE-3D Dataset: A Transfer Learning Model and DCGAN Approach to Address Data Imbalance. Cybernetics and Information Technologies, 25(3), 2025. 39-53. https://doi.org/10.2478/cait-2025-0021

2. **International Conference Proceedings**

- Youssera Zoukha Mecifi, Mohammed Merzoug, Fethallah Hadjila. Recent Advances in Deep Multimodal Fusion in Computer-Aided Diagnosis Systems: A literature Review. International Conference on Artificial Intelligence and Sustainable Development ICAISD'25, APRIL 12-13, 2025 IN RELIZANE UNIVERSITY, ALGERIA. p. 94.
- Youssera Zoukha Mecifi. Bridging Disciplines with Artificial Intelligence: Insights from Deep Multimodal Fusion for Humanities Research. The 1st International Colloque: "The Use of Artificial Intelligence in Scientific Research and Higher Education in the Humanities", April 2025, Tlemcen, Algeria

3. **In National Conference Proceedings**

- Youssera Zoukha Mecifi, Mohammed Merzoug, Fethallah Hadjila, Recent Advances in Deep Multimodal Fusion in Computer-Aided Diagnosis Systems: A literature Review, Poster presentation in: THE FIRST NATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND ITS APPLICATIONS (NCAIA'2024), December 2024, Constantine, Algeria

Bibliography

[1] World Health Organisation, “The top 10 causes of death.” Accessed: Jun. 24, 2025. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

[2] A. H. Roky *et al.*, “Overview of skin cancer types and prevalence rates across continents,” *Cancer Pathogenesis and Therapy*, vol. 3, no. 2, pp. 89–100, Mar. 2025, doi: 10.1016/j.cpt.2024.08.002.

[3] “Melanoma Skin Cancer Statistics.” Accessed: Feb. 24, 2026. [Online]. Available: https://www.cancer.org/cancer/types/melanoma-skin-cancer/about/key-statistics.html

[4] T. Regan, “Skin cancer misdiagnosis; your guide to making a medical negligence claim,” wake smith solicitors. Accessed: Jun. 23, 2025. [Online]. Available: https://www.wake-smith.co.uk/advice-centre/medical-negligence/hospital-negligence/misdiagnosis/skin-cancer-misdiagnosis.

[5] 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.

[6] S. Barth, “AI in Healthcare.” Accessed: Jun. 23, 2025. [Online]. Available: https://www.foreseemed.com/artificial-intelligence-in-healthcare

[7] Y. Z. Mecifi, M. Merzoug, A. Etchiali, M. M’hamedi, F. Hadjila, and A. Bekkouche, “Early Skin Cancer Detection Using the SLICE-3D Dataset: A Transfer Learning Model and DCGAN Approach to Address Data Imbalance,” *Cybernetics and Information Technologies*, vol. 25, no. 3, pp. 39–53, Sep. 2025, doi: 10.2478/cait-2025-0021.

[8] “ACS Annual Statistics Report: Milestone 70 Percent 5-Year Survival Rate for all Cancers Combined; Largest Gains for Advanced and Fatal Cancers,” American Cancer Society MediaRoom. Accessed: Feb. 24, 2026. [Online]. Available: https://pressroom.cancer.org/cancer-statistics-report-2026

[9] “Cancer Statistics,” National Cancer Institute. Accessed: Jun. 12, 2025. [Online]. Available: https://www.cancer.gov/about-cancer/understanding/statistics.

[10] P. Boyle, International Agency for Research on Cancer, and Weltgesundheitsorganisation, Eds., *World cancer report 2008*. Lyon: IARC Press, 2008.

[11] E. Craythorne and F. Al-Niami, “Skin cancer,” *Medicine*, vol. 45, no. 7, pp. 431–434, Jul. 2017, doi: 10.1016/j.mpmed.2017.04.003.

[12] A. G. Marzuka and S. E. Book, “Basal cell carcinoma: pathogenesis, epidemiology, clinical features, diagnosis, histopathology, and management,” *Yale J Biol Med*, vol. 88, no. 2, pp. 167–179, Jun. 2015.

[13] “Skin Lesions,” Cleveland Clinic. Accessed: Jun. 14, 2025. [Online]. Available: https://my.clevelandclinic.org/health/diseases/24296-skin-lesions

[14] “What Are Benign Skin Growths and How Can They Be Treated?,” Nature of Skin. Accessed: Jun. 14, 2025. [Online]. Available: https://www.natureofskinatx.com/what-are-benign-skin-growths-and-how-can-they-be-treated/

[15] N. C. Sathe and P. M. Zito, “Skin Cancer,” Treasure Island (FL): StatPearls Publishing; Accessed: Jun. 14, 2025. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK441949/

[16] V. P. Chinem and H. A. Miot, “Epidemiology of basal cell carcinoma,” *An Bras Dermatol.*, vol. 86, no. 2, pp. 292–305, 2011.

[17] B. McDaniel and R. Steele, “Basal Cell Carcinoma.,” In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; Accessed: Jun. 14, 2025. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK482439/

[18] C. E. M. Griffiths, D. Creamer, R. J. G. Chalmers, T. O. Bleiker, and J. Barker, Eds., *Rook’s textbook of dermatology*, Ninth edition. Chichester, West Sussex Hoboken, NJ: John Wiley & Sons Inc, 2016.

[19] M. Cives *et al.*, “Non-Melanoma Skin Cancers: Biological and Clinical Features,” *IJMS*, vol. 21, no. 15, p. 5394, Jul. 2020, doi: 10.3390/ijms21155394.

[20] S. Kang, M. Amagai, and A. Bruckner, Eds., *Fitzpatrick’s Dermatology in General Medicine*, 9th ed. McGraw-Hill Education, 2019.

[21] “Non-melanoma skin cancer risk factors,” Together we are beating cancer. [Online]. Available: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/non-melanoma-skin-cancer/risk-factors#heading-Four

[22] M. Khalesi, D. C. Whiteman, B. Tran, M. G. Kimlin, C. M. Olsen, and R. E. Neale, “A meta-analysis of pigmentary characteristics, sun sensitivity, freckling and melanocytic nevi and risk of basal cell carcinoma of the skin,” *Cancer Epidemiology*, vol. 37, no. 5, pp. 534–543, Oct. 2013, doi: 10.1016/j.canep.2013.05.008.

[23] T. B. Fitzpatrick, “The Validity and Practicality of Sun-Reactive Skin Types I Through VI,” *Arch Dermatol*, vol. 124, no. 6, p. 869, Jun. 1988, doi: 10.1001/archderm.1988.01670060015008.

[24] H. Sugiyama *et al.*, “Skin Cancer Incidence among Atomic Bomb Survivors from 1958 to 1996,” *Radiation Research*, vol. 181, no. 5, pp. 531–539, May 2014, doi: 10.1667/RR13494.1.

[25] A. Arafa, A. Mostafa, A. A. Navarini, and J.-Y. Dong, “The association between smoking and risk of skin cancer: a meta-analysis of cohort studies,” *Cancer Causes Control*, vol. 31, no. 8, pp. 787–794, Aug. 2020, doi: 10.1007/s10552-020-01319-8.

[26] S.-P. Sinikumpu, J. Jokelainen, S. Keinänen-Kiukaanniemi, and L. Huilaja, “Skin cancers and their risk factors in older persons: a population-based study,” *BMC Geriatr*, vol. 22, no. 1, p. 269, Dec. 2022, doi: 10.1186/s12877-022-02964-1.

[27] R. Popp *et al.*, “Impact of Surgical Refusal on Overall Survival in Patients With Melanoma: A Comprehensive Analysis,” *Anticancer Res*, vol. 45, no. 3, pp. 1063–1070, Mar. 2025, doi: 10.21873/anticanres.17493.

[28] M. A. Weinstock, “Nonmelanoma Skin Cancer Mortality: A Population-Based Study,” *Arch Dermatol*, vol. 127, no. 8, p. 1194, Aug. 1991, doi: 10.1001/archderm.1991.01680070094012.

[29] M. Perez, J. A. Abisaad, K. D. Rojas, M. A. Marchetti, and N. Jaimes, “Skin cancer: Primary, secondary, and tertiary prevention. Part I,” *Journal of the American Academy of Dermatology*, vol. 87, no. 2, pp. 255–268, Aug. 2022, doi: 10.1016/j.jaad.2021.12.066.

[30] F. Nachbar *et al.*, “The ABCD rule of dermatoscopy,” *Journal of the American Academy of Dermatology*, vol. 30, no. 4, pp. 551–559, Apr. 1994, doi: 10.1016/S0190-9622(94)70061-3.

[31] I. Muradia, N. Khunger, and A. K. Yadav, “A Clinical, Dermoscopic, and Histopathological Analysis of Common Acquired Melanocytic Nevi in Skin of Colo,” *Journal of Clinical and Aesthetic Dermatology*, vol. 15, no. 10, pp. 41–51, 2022.

[32] H. P. Soyer *et al.*, “Three-Point Checklist of Dermoscopy,” *Dermatology*, vol. 208, no. 1, pp. 27–31, 2004, doi: 10.1159/000075042.

[33] G. Argenziano, G. Fabbrocini, P. Carli, V. De Giorgi, E. Sammarco, and M. Delfino, “Epiluminescence Microscopy for the Diagnosis of Doubtful Melanocytic Skin Lesions: Comparison of the ABCD Rule of Dermatoscopy and a New 7-Point Checklist Based on Pattern Analysis,” *Arch Dermatol*, vol. 134, no. 12, Dec. 1998, doi: 10.1001/archderm.134.12.1563.

[34] S. W. Menzies, “Frequency and Morphologic Characteristics of Invasive Melanomas Lacking Specific Surface Microscopic Features,” *Arch Dermatol*, vol. 132, no. 10, p. 1178, Oct. 1996, doi: 10.1001/archderm.1996.03890340038007.

[35] H. Pehamberger, A. Steiner, and K. Wolff, “In vivo epiluminescence microscopy of pigmented skin lesions. I. Pattern analysis of pigmented skin lesions,” *Journal of the American Academy of Dermatology*, vol. 17, no. 4, pp. 571–583, Oct. 1987, doi: 10.1016/S0190-9622(87)70239-4.

[36] E.-G. Dobre *et al.*, “Skin Cancer Pathobiology at a Glance: A Focus on Imaging Techniques and Their Potential for Improved Diagnosis and Surveillance in Clinical Cohorts,” *IJMS*, vol. 24, no. 2, p. 1079, Jan. 2023, doi: 10.3390/ijms24021079.

[37] M. Berseth, “ISIC 2017 - Skin Lesion Analysis Towards Melanoma Detection,” 2017, *arXiv*. doi: 10.48550/ARXIV.1703.00523.

[38] P. Tschandl, C. Rosendahl, and H. Kittler, “The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions,” *Sci Data*, vol. 5, no. 1, p. 180161, Aug. 2018, doi: 10.1038/sdata.2018.161.

[39] “ISIC-2019.” [Online]. Available: https://challenge.isic-archive.com/landing/2019/

[40] International Skin Imaging Collaboration, “SIIM-ISIC 2020 Challenge Dataset.” International Skin Imaging Collaboration, 2020. doi: 10.34970/2020-DS01.

[41] N. R. Kurtansky *et al.*, “The SLICE-3D dataset: 400,000 skin lesion image crops extracted from 3D TBP for skin cancer detection,” *Sci Data*, vol. 11, no. 1, p. 884, Aug. 2024, doi: 10.1038/s41597-024-03743-w.

[42] A. Yilmaz, S. P. Yasar, G. Gencoglan, and B. Temelkuran, “DERM12345: A Large, Multisource Dermatoscopic Skin Lesion Dataset with 40 Subclasses,” *Sci Data*, vol. 11, no. 1, p. 1302, Nov. 2024, doi: 10.1038/s41597-024-04104-3.

[43] S. Legg and M. Hutter, “A Collection of Definitions of Intelligence,” 2007, doi: 10.48550/ARXIV.0706.3639.

[44] *Daniel Kahneman’s Thinking, Fast and Slow*. Macat Library, 2018. doi: 10.4324/9781912453207.

[45] A. M. Turing, “COMPUTING MACHINERY AND INTELLIGENCE,” *Mind*, vol. LIX, no. 236, pp. 433–460, Oct. 1950, doi: 10.1093/mind/LIX.236.433.

[46] H. Hassani, E. S. Silva, S. Unger, M. TajMazinani, and S. Mac Feely, “Artificial Intelligence (AI) or Intelligence Augmentation (IA): What Is the Future?,” *AI*, vol. 1, no. 2, pp. 143–155, Apr. 2020, doi: 10.3390/ai1020008.

[47] C. Janiesch, P. Zschech, and K. Heinrich, “Machine learning and deep learning,” *Electron Markets*, vol. 31, no. 3, pp. 685–695, Sep. 2021, doi: 10.1007/s12525-021-00475-2.

[48] T. Talaei Khoei, H. Ould Slimane, and N. Kaabouch, “Deep learning: systematic review, models, challenges, and research directions,” *Neural Comput & Applic*, vol. 35, no. 31, pp. 23103–23124, Nov. 2023, doi: 10.1007/s00521-023-08957-4.

[49] X. Wang, Y. Zhao, and F. Pourpanah, “Recent advances in deep learning,” *Int. J. Mach. Learn. & Cyber.*, vol. 11, no. 4, pp. 747–750, Apr. 2020, doi: 10.1007/s13042-020-01096-5.

[50] S. Amraee, M. Chinipardaz, and M. Charoosaei, “Analytical study of two feature extraction methods in comparison with deep learning methods for classification of small metal objects,” *Vis. Comput. Ind. Biomed. Art*, vol. 5, no. 1, p. 13, Dec. 2022, doi: 10.1186/s42492-022-00111-6.

[51] V. Rajaraman, “JohnMcCarthy — Father of artificial intelligence,” *Reson*, vol. 19, no. 3, pp. 198–207, Mar. 2014, doi: 10.1007/s12045-014-0027-9.

[52] J. E. (James E.) Tomayko, “Behind Deep Blue: Building the Computer that Defeated the World Chess Champion (review),” *tech*, vol. 44, no. 3, pp. 634–635, Jul. 2003, doi: 10.1353/tech.2003.0140.

[53] A. Pannu and M. Tech Student, “Artificial Intelligence and its Application in Different Areas,” 2015. [Online]. Available: https://api.semanticscholar.org/CorpusID:113451322

[54] I. H. Sarker, “Machine Learning: Algorithms, Real-World Applications and Research Directions,” *SN COMPUT. SCI.*, vol. 2, no. 3, p. 160, May 2021, doi: 10.1007/s42979-021-00592-x.

[55] I. H. Sarker, “Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions,” *SN COMPUT. SCI.*, vol. 2, no. 6, p. 420, Nov. 2021, doi: 10.1007/s42979-021-00815-1.

[56] F. Rida, “Smart Business : repenser le business de votre entreprise avec l’IA, le Machine Learning et le Deep Learning,” Cellenza Blog. Accessed: Oct. 04, 2025. [Online]. Available: https://blog.cellenza.com/data/smart-business-repenser-le-business-de-votre-entreprise-avec-lia-le-machine-learning-et-le-deep-learning.

[57] R. Schalkoff, *Artificial Neural Networks*. McGraw-Hill, New York., 1997.

[58] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[59] K. O’Hara, W. Hall, K. Van Rijsbergen, and N. Shadbolt, “Memory, Reasoning and Learning,” in *Cognitive Systems - Information Processing Meets Brain Science*, Elsevier, 2006, pp. 236–260. doi: 10.1016/B978-012088566-4/50016-7.

[60] Q. Zhang, H. Yu, M. Barbiero, B. Wang, and M. Gu, “Artificial neural networks enabled by nanophotonics,” *Light Sci Appl*, vol. 8, no. 1, p. 42, May 2019, doi: 10.1038/s41377-019-0151-0.

[61] S. Agatonovic-Kustrin and R. Beresford, “Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research,” *Journal of Pharmaceutical and Biomedical Analysis*, vol. 22, no. 5, pp. 717–727, Jun. 2000, doi: 10.1016/S0731-7085(99)00272-1.

[62] G. Bebis and M. Georgiopoulos, “Feed-forward neural networks,” *IEEE Potentials*, vol. 13, no. 4, pp. 27–31, Oct. 1994, doi: 10.1109/45.329294.

[63] H. Taud and J. F. Mas, “Multilayer Perceptron (MLP),” in *Geomatic Approaches for Modeling Land Change Scenarios*, M. T. Camacho Olmedo, M. Paegelow, J.-F. Mas, and F. Escobar, Eds., in Lecture Notes in Geoinformation and Cartography. , Cham: Springer International Publishing, 2018, pp. 451–455. doi: 10.1007/978-3-319-60801-3_27.

[64] L. R. Medske and L. Jain, *Recurrent neural networks. Design and Applications*, vol. 5(64-67), 2. 2001.

[65] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[66] N. Ketkar and J. Moolayil, “Convolutional Neural Networks,” in *Deep Learning with Python*, Berkeley, CA: Apress, 2021, pp. 197–242. doi: 10.1007/978-1-4842-5364-9_6.

[67] G. Corso, H. Stark, S. Jegelka, T. Jaakkola, and R. Barzilay, “Graph neural networks,” *Nat Rev Methods Primers*, vol. 4, no. 1, p. 17, Mar. 2024, doi: 10.1038/s43586-024-00294-7.

[68] Department of Computer Sciences, Ajayi Crowther University, Oyo, Nigeria. and J. A. Ayeni, “Convolutional Neural Network (CNN): The architecture and applications,” *Appl. J. Phys. Sci.*, vol. 4, no. 4, pp. 42–50, Dec. 2022, doi: 10.31248/AJPS2022.085.

[69] Department of Computer Sciences, Ajayi Crowther University, Oyo, Nigeria. and J. A. Ayeni, “Convolutional Neural Network (CNN): The architecture and applications,” *Appl. J. Phys. Sci.*, vol. 4, no. 4, pp. 42–50, Dec. 2022, doi: 10.31248/AJPS2022.085.

[70] S. Albelwi and A. Mahmood, “A Framework for Designing the Architectures of Deep Convolutional Neural Networks,” *Entropy*, vol. 19, no. 6, p. 242, May 2017, doi: 10.3390/e19060242.

[71] L. Torrey and J. Shavlik, “Transfer Learning:,” in *Handbook of Research on Machine Learning Applications and Trends*, E. S. Olivas, J. D. M. Guerrero, M. Martinez-Sober, J. R. Magdalena-Benedito, and A. J. Serrano López, Eds., IGI Global, 2010, pp. 242–264. doi: 10.4018/978-1-60566-766-9.ch011.

[72] K. Weiss, T. M. Khoshgoftaar, and D. Wang, “A survey of transfer learning,” *J Big Data*, vol. 3, no. 1, p. 9, Dec. 2016, doi: 10.1186/s40537-016-0043-6.

[73] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, “How transferable are features in deep neural networks?,” 2014, doi: 10.48550/ARXIV.1411.1792.

[74] M. Iman, H. R. Arabnia, and K. Rasheed, “A Review of Deep Transfer Learning and Recent Advancements,” *Technologies*, vol. 11, no. 2, p. 40, Mar. 2023, doi: 10.3390/technologies11020040.

[75] F. Chollet, *Deep learning with Python. Simon and Schuster.*, Simon and Schuster. 2021.

[76] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” 2014, *arXiv*. doi: 10.48550/ARXIV.1409.1556.

[77] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” 2015, *arXiv*. doi: 10.48550/ARXIV.1512.03385.

[78] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the Inception Architecture for Computer Vision,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA: IEEE, Jun. 2016, pp. 2818–2826. doi: 10.1109/CVPR.2016.308.

[79] C. Szegedy *et al.*, “Going deeper with convolutions,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA: IEEE, Jun. 2015, pp. 1–9. doi: 10.1109/CVPR.2015.7298594.

[80] M. Tan and Q. V. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” 2019, doi: 10.48550/ARXIV.1905.11946.

[81] A. G. Howard *et al.*, “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications,” Apr. 17, 2017, *arXiv*: arXiv:1704.04861. doi: 10.48550/arXiv.1704.04861.

[82] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “MobileNetV2: Inverted Residuals and Linear Bottlenecks,” 2018, doi: 10.48550/ARXIV.1801.04381.

[83] F. Chollet, “Xception: Deep Learning with Depthwise Separable Convolutions,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI: IEEE, Jul. 2017, pp. 1800–1807. doi: 10.1109/CVPR.2017.195.

[84] J. Deng, W. Dong, R. Socher, L.-J. Li, Kai Li, and Li Fei-Fei, “ImageNet: A large-scale hierarchical image database,” in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, Miami, FL: IEEE, Jun. 2009, pp. 248–255. doi: 10.1109/CVPR.2009.5206848.

[85] A. Mumuni and F. Mumuni, “Data augmentation: A comprehensive survey of modern approaches,” *Array*, vol. 16, p. 100258, Dec. 2022, doi: 10.1016/j.array.2022.100258.

[86] V. Kůrková, *Artificial Neural Networks and Machine Learning - ICANN 2018: 27th International Conference on Artificial Neural Networks, Rhodes, Greece, October 4-7, 2018, Proceedings, Part I*. in Lecture Notes in Computer Science Ser, no. v. 11139. Cham: Springer International Publishing AG, 2018.

[87] L. Moles, A. Andres, G. Echegaray, and F. Boto, “Exploring Data Augmentation and Active Learning Benefits in Imbalanced Datasets,” *Mathematics*, vol. 12, no. 12, p. 1898, Jun. 2024, doi: 10.3390/math12121898.

[88] R. Susmaga, “Confusion Matrix Visualization,” in *Intelligent Information Processing and Web Mining*, M. A. Kłopotek, S. T. Wierzchoń, and K. Trojanowski, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 107–116. doi: 10.1007/978-3-540-39985-8_12.

[89] S. V. Stehman, “Selecting and interpreting measures of thematic classification accuracy,” *Remote Sensing of Environment*, vol. 62, no. 1, pp. 77–89, Oct. 1997, doi: 10.1016/S0034-4257(97)00083-7.

[90] K. Aljohani and T. Turki, “Automatic Classification of Melanoma Skin Cancer with Deep Convolutional Neural Networks,” *AI*, vol. 3, no. 2, pp. 512–525, Jun. 2022, doi: 10.3390/ai3020029.

[91] K. Djaroudib, P. Lorenz, R. Belkacem Bouzida, and H. Merzougui, “Skin Cancer Diagnosis Using VGG16 and Transfer Learning: Analyzing the Effects of Data Quality over Quantity on Model Efficiency,” *Applied Sciences*, vol. 14, no. 17, p. 7447, Aug. 2024, doi: 10.3390/app14177447.

[92] S. Sikandar, R. Mahum, A. E. Ragab, S. Y. Yayilgan, and S. Shaikh, “SCDet: A Robust Approach for the Detection of Skin Lesions,” *Diagnostics*, vol. 13, no. 11, p. 1824, May 2023, doi: 10.3390/diagnostics13111824.

[93] Yashwant S. Ingle, “Deep Learning for Skin Cancer Classification: A Comparative Study of CNN and Vgg16 on HAM10000 Dataset,” *cana*, vol. 31, no. 4s, pp. 490–499, Jul. 2024, doi: 10.52783/cana.v31.944.

[94] P. K, S. P, A. I. R, S. K, V. M, and S. N, “Deep learning-based skin lesion classification for cancer detection,” Dec. 30, 2024. doi: 10.21203/rs.3.rs-5675190/v1.

[95] M. M’hamedi, M. Merzoug, M. Hadjila, and A. Bekkouche, “Enhancing melanoma skin cancer classification through data augmentation,” *TELKOMNIKA*, vol. 22, no. 5, p. 1209, Jul. 2024, doi: 10.12928/telkomnika.v22i5.26106.

[96] W. Gouda, N. U. Sama, G. Al-Waakid, M. Humayun, and N. Z. Jhanjhi, “Detection of Skin Cancer Based on Skin Lesion Images Using Deep Learning,” *Healthcare*, vol. 10, no. 7, p. 1183, Jun. 2022, doi: 10.3390/healthcare10071183.

[97] R. A. Mehr and A. Ameri, “Skin Cancer Detection Based on Deep Learning,” *J Biomed Phys Eng*, vol. 12, no. 6, Dec. 2022, doi: 10.31661/jbpe.v0i0.2207-1517.

[98] A. Alrabai, A. Echtioui, and F. Kallel, “Skin Cancer Detection Based on Transfer Learning Techniques,” in *2024 IEEE 7th International Conference on Advanced*

Technologies, Signal and Image Processing (ATSIP), Sousse, Tunisia: IEEE, Jul. 2024, pp. 8–13. doi: 10.1109/ATSIP62566.2024.10638982.

[99] M. S. Islam and S. Panta, "Skin Cancer Images Classification using Transfer Learning Techniques," Jun. 18, 2024, *arXiv*: arXiv:2406.12954. doi: 10.48550/arXiv.2406.12954.

[100] J. Pope, M. Hassanuzzaman, M. Sherpa, O. Emara, A. Joshi, and N. Adhikari, "Skin Cancer Machine Learning Model Tone Bias," Oct. 08, 2024, *arXiv*: arXiv:2410.06385. doi: 10.48550/arXiv.2410.06385.

[101] T. Imran, A. S. Alghamdi, and M. S. Alkatheiri, "Enhanced Skin Cancer Classification using Deep Learning and Nature-based Feature Optimization," *Eng. Technol. Appl. Sci. Res.*, vol. 14, no. 1, pp. 12702–12710, Feb. 2024, doi: 10.48084/etasr.6604.

[102] A. Naeem, T. Anees, M. Khalil, K. Zahra, R. A. Naqvi, and S.-W. Lee, "SNC_Net: Skin Cancer Detection by Integrating Handcrafted and Deep Learning-Based Features Using Dermoscopy Images," *Mathematics*, vol. 12, no. 7, p. 1030, Mar. 2024, doi: 10.3390/math12071030.

[103] P. Sedigh, R. Sadeghian, and M. T. Masouleh, "Generating Synthetic Medical Images by Using GAN to Improve CNN Performance in Skin Cancer Classification," in *2019 7th International Conference on Robotics and Mechatronics (ICRoM)*, Tehran, Iran: IEEE, Nov. 2019, pp. 497–502. doi: 10.1109/ICRoM48714.2019.9071823.

[104] A. A. M. Teodoro *et al.*, "A Skin Cancer Classification Approach using GAN and RoI-Based Attention Mechanism," *J Sign Process Syst*, vol. 95, no. 2–3, pp. 211–224, Mar. 2023, doi: 10.1007/s11265-022-01757-4.

[105] S. Teymouri, A. A. Larki, M. Mohammadi, and S. H. Klidbary, "Hybrid Vision Transformer-Boosting Model for Skin Lesion Classification on SLICE-3D," in *2024 12th RSI International Conference on Robotics and Mechatronics (ICRoM)*, Tehran, Iran, Islamic Republic of: IEEE, Dec. 2024, pp. 586–592. doi: 10.1109/ICRoM64545.2024.10903606.

[106] S. Syed and E. M. Albalawi, "Improved Skin Cancer Detection with 3D Total Body Photography: Integrating AI Algorithms for Precise Diagnosis," Jul. 08, 2024. doi: 10.21203/rs.3.rs-4677329/v1.

[107] E. Pintelas, I. E. Livieris, V. Tampakas, and P. Pintelas, "MobileNet-HeX: Heterogeneous Ensemble of MobileNet eXperts for Efficient and Scalable Vision Model Optimization," *BDCC*, vol. 9, no. 1, p. 2, Dec. 2024, doi: 10.3390/bdcc9010002.

[108] International Skin Imaging Collaboration (ISIC) · Research Code Competition ·, "ISIC 2024 - Skin Cancer Detection with 3D-TBP." Accessed: Jun. 16, 2025. [Online]. Available: https://www.kaggle.com/competitions/isic-2024-challenge

[109] A. Ferreirinha, V. Farricha, and A. L. João, “Melanoma Diagnosis With 3D Total-Body Photography,” *Actas Dermo-Sifiliográficas*, p. S000173102500105X, Feb. 2025, doi: 10.1016/j.ad.2024.09.030.

[110] C. Navarrete-Dechent, K. Liopyris, A. Rishpon, N. G. Marghoob, J. Monnier, and A. A. Marghoob, “Total Body Photography as an Aid for the Early Detection of Skin Cancer,” in *Photography in Clinical Medicine*, P. Pasquali, Ed., Cham: Springer International Publishing, 2020, pp. 253–269. doi: 10.1007/978-3-030-24544-3_18.

[111] “Clinical Perspective of 3D Total Body Photography for Early Detection and Screening of Melanoma.” Accessed: Jun. 16, 2025. [Online]. Available: https://www.frontiersin.org/files/Articles/371716/fmed-05-00152-HTML/image_m/fmed-05-00152-g001.jpg

[112] Melanoma Institute Australia, “Total Body Photography Patient Information,” Melanoma Institute Australia, Sep. 2022. [Online]. Available: https://melanoma.org.au/wp-content/uploads/2022/08/MIA22-TOTAL-BODY-PHOTOGRAPHY_update.pdf

[113] C. Shorten and T. M. Khoshgoftaar, “A survey on Image Data Augmentation for Deep Learning,” *J Big Data*, vol. 6, no. 1, p. 60, Dec. 2019, doi: 10.1186/s40537-019-0197-0.

[114] L. Perez and J. Wang, “The Effectiveness of Data Augmentation in Image Classification using Deep Learning,” 2017, *arXiv*. doi: 10.48550/ARXIV.1712.04621.

[115] A. Radford, L. Metz, and S. Chintala, “Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks,” Jan. 07, 2016, *arXiv*: arXiv:1511.06434. doi: 10.48550/arXiv.1511.06434.