Faculty of Information and Communication Technology

A dissertation presented and submitted in partial fulfilment of the requirements for the degree of a Bachelor of Science in Computer Science

Titile

Development of a Skin Lesion Detection and Classification System using Convolutional Neural Networks (CNNs)

By

Ngane Emmanuel

Registration Number: ICTU20222972

Supervised by: Engr. Nkiamboh Tanwi

August 11, 2025

DECLARATION

I declare that the work entitled **"Development of a Skin Lesion Detection and Classification System using Convolutional Neural Networks (CNNs)"** is my own original work, conceived and presented in the partial fulfilment of the requirement for the degree of a Bachelor of Science in Computer Science at ICT University. This work has not been submitted for any degree or examination in any other university, and that all the sources I have used or quoted have been indicated and acknowledged as complete references.

Signed _____ Date: _____

Name: _____

Registration Number: _____

CERTIFICATION

This work entitled **"Development of a Skin Lesion Detection and Classification System using Convolutional Neural Networks (CNNs)"** has been submitted for examination with my approval as the Research Supervisor.

Signed _____ Date: _____

Name: _____

DEDICATION

I dedicate this work to my parents, Rev. Dr. Ntoko Samuel Eseh and Mme Ntoko Grace Melioge, and to my loving sister, Ntoko Racheal Edenge, for their unwavering support, encouragement, and sacrifices throughout my academic journey.

Acknowledgments

I would like to express my sincere appreciation to my project supervisor, Engr. Nkiambo Tanwi, whose guidance, feedback, and encouragement were invaluable throughout the course of this project. His expertise and support helped shape both the direction and quality of this research.

I am also grateful to the faculty and staff of the Department of Computer Science, ICT University, for providing the academic foundation and resources necessary for the completion of this work. Their commitment to academic excellence has been instrumental in my development.

Special thanks go to my classmates and friends for their collaboration, feedback, and moral support during the challenging phases of this project. Their insights and motivation helped me stay focused and persistent.

Lastly, I appreciate my family for their continuous encouragement and understanding, which enabled me to dedicate the necessary time and energy to this project.

This project has been a significant learning experience, and I am thankful to all who contributed to its successful completion.

FACULTY APPROVAL

This dissertation has been duly reviewed by the Department and the Faculty and is ready for examination with our approval.

Approved by

Signature Date

_____

Engr. Nkiamboh Tanwi

Supervisor

Signature Date

_____

Dr. Abdallah Ziraba

Head of Department

Signature Date

_____

Dr. Luc Einstein Ngend

Dean

ABSTRACT

Skin lesions encompass a wide spectrum of conditions, ranging from benign disorders such as eczema, fungal infections, and psoriasis to malignant forms like melanoma and squamous cell carcinoma. Globally, these conditions impose significant public health and economic burdens, with their impact being particularly acute in low-resource regions where specialized dermatological expertise is scarce. In Cameroon and much of sub-Saharan Africa, access to timely and accurate dermatological diagnosis is hindered by limited specialist availability, under-resourced health facilities, and geographic barriers to care.

This study presents the development of a multi-class skin lesion detection and classification system using Convolutional Neural Networks (CNNs), designed to enhance diagnostic accessibility through automated image analysis. The system was trained on a curated, high-quality dataset compiled from multiple open-access sources, including Kaggle dermatology repositories and the DermNetNZ medical image database. Careful preprocessing and augmentation techniques were employed to normalize image quality, address class imbalance, and improve model robustness across diverse lesion presentations and skin tones.

A fine-tuned ResNet-18 architecture, leveraging transfer learning from the ImageNet dataset, was implemented to classify lesions into multiple diagnostic categories, covering both malignant and non-malignant conditions. The model was trained using supervised learning with categorical cross-entropy loss, and evaluated using metrics including accuracy, precision, recall, and F1-score. The final system achieved a validation accuracy of 75% and demonstrated consistent performance on the test set, underscoring its potential as a supportive diagnostic aid.

The implementation includes a command-line interface (CLI) for research use and a prototype mobile application framework aimed at offline deployment in rural and underserved settings. By combining computational efficiency with diagnostic breadth, the system offers a cost-effective and scalable approach to skin lesion classification in contexts where traditional healthcare access is limited.

This project contributes to the growing field of AI-assisted dermatology by emphasizing inclusivity across skin tones and lesion types, while maintaining technical feasibility for low-resource environments. Future improvements will focus on expanding the dataset, integrating explainable AI features, incorporating multi-modal patient data, and validating the system in real-world clinical settings.

Keywords: Skin lesion detection, convolutional neural networks, ResNet-18, transfer learning, dataset diversity, mobile health, Cameroon, medical imaging.

Contents

List of Figures

1 Chapter 1: Introduction

1.1 Introduction

Skin lesions encompass a broad spectrum of abnormalities in the skin, ranging from benign conditions such as moles and warts to malignant manifestations like melanoma and squamous cell carcinoma. Globally, these dermatological disorders present a significant public health challenge, with skin cancers alone accounting for millions of new cases annually and exerting substantial socioeconomic and healthcare burdens (Ogudo et al., 2023), (Kassem et al., 2021). The prevalence of skin lesions has been amplified by factors such as increased ultraviolet (UV) exposure due to climate change, demographic transitions, and lifestyle changes, contributing to both malignant and non-malignant skin disorders (Zafar et al., 2023).

In high-income countries, early diagnosis is often facilitated by access to trained dermatologists, advanced imaging technologies, and well-established referral systems. However, in low- and middle-income countries (LMICs), particularly in sub-Saharan Africa, dermatology services remain sparse, with a severe shortage of qualified dermatologists and specialized diagnostic tools (Hay et al., 2006). This disparity often leads to delayed diagnoses, inappropriate treatments, and preventable morbidity and mortality from conditions that are otherwise treatable when detected early.

Cameroon, like many African nations, faces a dual challenge: the burden of infectious skin diseases such as fungal infections, scabies, and bacterial dermatitis coexists with a growing incidence of non-communicable skin disorders, including actinic keratosis, psoriasis, and malignant lesions (Hees & Naafs, 2005). The limited integration of dermatological screening into primary healthcare, coupled with the absence of large-scale public awareness campaigns, exacerbates the problem.

Recent advancements in artificial intelligence (AI) and machine learning (ML), particularly deep learning-based image analysis, have demonstrated remarkable potential in automating the detection and classification of skin lesions with performance levels approaching or surpassing that of experienced dermatologists (Khan et al., 2021), (Adegun & Viriri, 2020). Such systems leverage large annotated datasets to learn discriminative features directly from dermoscopic or clinical images, enabling fast, scalable, and cost-effective screening.

This research seeks to harness these technological advancements to develop a skin lesion detection and classification system capable of accurately identifying multiple common skin lesion types. The proposed system integrates a convolutional neural network (CNN) trained on diverse, high-quality datasets sourced from publicly available repositories such as Kaggle and curated medical image databases like DermNetNZ. In addition to the AI model, the system will include a graphical user interface (GUI) for clinical use, a command-line interface (CLI) for research purposes, and a mobile application to extend accessibility to rural and underserved communities. This comprehensive approach aims to bridge the diagnostic gap, improve early detection rates, and contribute to the broader objective of reducing the dermatological disease burden in Cameroon and similar contexts.

1.2 Background to the Problem

Dermatological diseases represent one of the most common categories of health complaints worldwide, with conditions ranging from self-limiting skin irritations to life-threatening malignancies. The World Health Organization (WHO) has repeatedly emphasized that skin diseases are among the top ten causes of non-fatal disease burden globally, and in certain regions of sub-Saharan Africa, they are among the most frequent reasons for outpatient consultations (Hay et al., 2006), (Hees & Naafs, 2005). While conditions like melanoma and basal cell carcinoma dominate skin cancer statistics in developed nations, the disease landscape in Africa and particularly Cameroon is more heterogeneous, encompassing a mix of infectious, inflammatory, and neoplastic disorders.

Traditionally, diagnosis of skin lesions has relied heavily on visual inspection by a dermatologist, often supplemented by dermoscopic evaluation or histopathological examination for suspicious cases (Yap et al., 2018). However, this process is inherently subjective and highly dependent on the clinician's expertise and experience. In many LMICs, including Cameroon, the ratio of dermatologists to the population is critically low, often less than one dermatologist per million inhabitants, leading to substantial diagnostic delays. Moreover, general practitioners and nurses, who often serve as first-line healthcare providers, may lack specialized training in dermatology, further compounding the risk of misdiagnosis or missed diagnoses.

The advent of AI in medical imaging, particularly the application of convolutional neural networks (CNNs) to dermoscopic and clinical images, has opened new possibilities for addressing these

challenges (Harangi, 2018), (Mahbod et al., 2019). CNNs can automatically extract hierarchical features from raw image data, allowing them to differentiate between lesion types with minimal human intervention. Studies have demonstrated that AI-based diagnostic tools can achieve diagnostic accuracy comparable to or exceeding that of experienced dermatologists in controlled settings (Albahar, 2019), (Jinnai et al., 2020).

Despite these advancements, most existing AI models for skin lesion classification have been trained and validated on datasets that predominantly contain images from lighter-skinned populations in Europe, North America, and Australia (Gouda et al., 2022). This raises concerns about model generalizability to populations with darker skin tones, such as those in sub-Saharan Africa. Additionally, existing research has often focused on detecting malignant lesions, particularly melanoma, with limited emphasis on the broader spectrum of common lesions prevalent in African contexts, such as eczema, tinea infections, or Kaposi's sarcoma.

This project responds to these gaps by developing a multi-class lesion classification system specifically curated to include a diverse set of common skin conditions affecting African populations. The dataset compilation process involves selecting high-quality images from multiple Kaggle repositories and supplementing them with clinically verified images from DermNetNZ, ensuring representation across various lesion categories and skin tones. Through this approach, the proposed system aims to offer a clinically relevant, context-sensitive diagnostic support tool that addresses both the technical challenges of AI implementation and the pressing healthcare needs of the target population.

1.3 Problem Statement

The early and accurate diagnosis of skin lesions remains a critical public health challenge in Cameroon and across much of sub-Saharan Africa. While skin lesions may appear trivial in their initial stages, certain forms such as malignant melanomas, squamous cell carcinomas, and basal cell carcinomas, can rapidly progress to life-threatening stages if left untreated. In parallel, other non-malignant but common conditions, such as fungal infections, eczema, and psoriasis, can cause significant discomfort, social stigma, and economic loss due to chronicity and recurrence.

In many urban centers of high-income countries, advanced diagnostic services and dermatology

specialists are accessible, enabling early detection and management. However, in Cameroon, the number of dermatologists is grossly inadequate relative to the population size, with some regions having no specialist at all (Hay et al., 2006), (Hees & Naafs, 2005). This shortage forces patients to rely on general practitioners or traditional healers, which often leads to delayed diagnoses, mis-management, or complete neglect of skin-related health issues.

Furthermore, conventional diagnostic workflows are limited by two key barriers:

1. **Geographic accessibility** – Many patients in rural areas must travel long distances to reach healthcare facilities, resulting in missed opportunities for early diagnosis.
2. **Resource limitations** – Even in urban hospitals, dermoscopic imaging equipment and histopathological services are often scarce or prohibitively expensive.

Recent developments in artificial intelligence, specifically convolutional neural networks (CNNs), have shown promise in automating skin lesion detection and classification with high accuracy (Khan et al., 2021), (Adegun & Viriri, 2020). Nonetheless, most existing models are trained on datasets from lighter-skinned populations and focus primarily on malignant lesions, limiting their applicability in African contexts where lesion diversity and skin tone variations differ significantly (Gouda et al., 2022).

The central problem, therefore, is the absence of an accessible, accurate, and context-specific diagnostic tool for multiple types of skin lesions prevalent in Cameroon. There is a need for an integrated system that combines AI-powered lesion detection with user-friendly interfaces, ranging from clinical desktop applications to mobile apps that can bridge the diagnostic gap between rural patients, urban hospitals, and research institutions.

1.4 Objectives of the Study

The overarching aim of this study is to develop a comprehensive skin lesion detection and classification system that leverages deep learning techniques to address the diagnostic challenges faced in Cameroon and similar low-resource settings.

1.4.1 General Objective

To design and implement an AI-driven system capable of accurately detecting and classifying multiple types of common skin lesions, with deployment across desktop, command-line, and mobile platforms to improve diagnostic accessibility.

1.4.2 Specific Objectives

1. To compile and curate a diverse, high-quality dataset of common skin lesion images from multiple open-access sources, including Kaggle repositories and the DermNetNZ medical image database.
2. To preprocess and augment image data to enhance model robustness..
3. To train a convolutional neural network (CNN) model optimized for multi-class lesion classification, ensuring high performance on both malignant and non-malignant categories.
4. To develop a command-line interface (CLI) for research and model testing purposes.
5. To design and implement a mobile application capable of performing on-device lesion classification for remote and rural healthcare access.
6. To evaluate the system's performance against standard diagnostic accuracy metrics, including sensitivity, specificity, and overall classification accuracy.
7. To assess the feasibility of integrating the system into existing telemedicine frameworks in Cameroon.

1.5 Scope of the Study

This study focuses on the development of a skin lesion detection and classification system tailored for the Cameroonian healthcare context, while incorporating global best practices in AI-assisted dermatological diagnosis. The scope of the research spans four key dimensions: the dataset, the technical solution, the deployment platforms, and the evaluation process.

From a **data perspective**, the system will be trained and validated using a carefully curated dataset compiled from multiple publicly available repositories on Kaggle, supplemented with high-quality dermatological images sourced from DermNetNZ. These sources offer a broad spectrum of lesion categories, ranging from malignant cancers such as melanoma and squamous cell carcinoma to non-malignant conditions like eczema, psoriasis, and fungal infections. Selection criteria will prioritize image clarity, correct annotation, diversity of skin tones, and representation of lesion variations to ensure the model's applicability across a wide demographic.

From a **technical perspective**, the system will employ a convolutional neural network (CNN) architecture fine-tuned for multi-class classification. Advanced preprocessing techniques, such as data augmentation and normalization, will be applied to enhance the model's generalizability across different imaging conditions. While the primary focus is on image-based diagnosis, the system is not intended to replace histopathological confirmation, which remains the gold standard for definitive lesion classification.

From a **deployment perspective**, the project encompasses three interfaces:

1. A **Command-Line Interface (CLI)**, primarily intended for researchers and developers for testing and evaluating the AI model.

2. A **Mobile Application**, optimized for both offline and online usage, to enable healthcare workers and patients in rural areas to perform preliminary lesion assessments without constant internet connectivity.

From an **evaluation perspective**, the model will be tested using standard performance metrics, including accuracy, sensitivity, specificity, and confusion matrix analysis. The evaluation will emphasize the system's ability to handle variations in lesion presentation due to differences in skin pigmentation, image quality, and environmental lighting.

The project does not aim to cover the entire spectrum of dermatological conditions, rare or highly complex lesions requiring specialized diagnostic equipment fall outside the intended operational scope. Instead, the focus is on delivering a reliable, accessible, and scalable solution for the most

common and clinically significant skin lesions encountered in Cameroon and similar resource-limited environments.

1.6 Significance of the Study

Skin diseases are among the most common health concerns globally, with billions of people affected each year, spanning all socioeconomic groups (Hay et al., 2006). In Cameroon and much of sub-Saharan Africa, the impact of skin lesions is amplified by a combination of high prevalence, limited specialist care, and widespread misinformation about skin health. Misdiagnosed or untreated lesions, whether malignant or non-malignant, can lead to severe health complications, disfigurement, psychological distress, and in the case of cancers, increased mortality rates.

The significance of this study lies in its potential to bridge the diagnostic gap through a context-specific, AI-driven approach. By leveraging deep learning algorithms and multi-platform deployment, the system offers a means of providing timely and accurate lesion assessment to populations that currently lack access to dermatological expertise. For healthcare providers, this could translate to earlier interventions, better patient outcomes, and a reduction in the burden on tertiary healthcare facilities.

From a **public health perspective**, the project addresses a critical need in preventive care. Early detection and classification can significantly reduce treatment costs and improve survival rates in malignant cases, while minimizing chronic complications in non-malignant cases. Moreover, the system's mobile deployment makes it a practical tool for community health workers conducting outreach in rural and peri-urban areas.

From a **technological perspective**, the research contributes to the growing field of AI in medical imaging by demonstrating how existing machine learning methods can be adapted to underrepresented populations. Many existing AI models for skin lesion classification are trained predominantly on lighter-skinned datasets, making them less effective for darker skin tones (Gouda et al., 2022). By incorporating diverse skin tone representation from datasets like DermNetNZ and Kaggle repositories, this study advances the inclusivity and fairness of medical AI systems.

From an **academic perspective**, the project provides a valuable reference for future studies in both

computer vision and health informatics within the African context. The integration of a research-oriented CLI, a mobile health application, and a clinical GUI sets a precedent for multi-platform medical AI systems designed for low-resource settings.

Ultimately, the study seeks to contribute toward Sustainable Development Goal 3 (Good Health and Well-being) by enabling accessible, affordable, and high-quality dermatological diagnostics in regions where such services are currently limited or absent.

1.6.1 1.7 Limitations of the Study

While this study seeks to develop a comprehensive, AI-powered skin lesion detection and classification system, several constraints inevitably shape the scope and potential impact of the research. First, the dataset, although carefully curated from high-quality open-access sources such as Kaggle repositories and the DermNetNZ medical image database, may not capture the full diversity of skin lesion presentations across all ethnic groups and age categories. This limitation is particularly relevant for sub-Saharan African populations, where variations in skin pigmentation can influence lesion visibility and morphology, potentially impacting model generalization (Hay et al., 2006), (Hees & Naafs, 2005).

Second, although image preprocessing and augmentation techniques are employed to mitigate overfitting and improve robustness, the absence of histopathological confirmation for all dataset images introduces an inherent diagnostic uncertainty. Clinical image-based diagnosis, while valuable, cannot entirely substitute for biopsy-confirmed ground truth, especially in differentiating visually similar lesion types (Zafar et al., 2023).

Third, computational resource limitations constrain the complexity and scale of model experimentation. While transfer learning with established CNN architectures, such as ResNet, offers strong baseline performance, exploring more computationally intensive models or ensemble methods may be restricted due to hardware constraints. This also impacts the breadth of hyperparameter optimization that can be conducted within the project timeline.

Fourth, the current implementation focuses primarily on classification accuracy and does not yet incorporate a complete clinical decision support framework, such as integration with electronic

health records (EHRs) or automated referral systems. Similarly, the system's deployment is confined to a mobile application and a research-oriented command-line interface, without the immediate inclusion of a graphical interface for clinical integration, an enhancement reserved for future development phases.

Finally, while the study includes performance evaluation using standard diagnostic metrics (e.g., sensitivity, specificity, overall accuracy), real-world clinical validation with dermatologists or in live telemedicine settings is beyond the scope of the present work. This limits immediate translation into routine healthcare workflows but provides a foundation for subsequent validation studies.

1.6.2 1.8 Organization of the Study

The remainder of this thesis is organized into five chapters, each systematically addressing a core component of the research.

- **Chapter 1 – Introduction**: Provides an overview of the research problem, contextual background, study objectives, scope, significance, and limitations, setting the stage for the investigation.
- **Chapter 2 – Literature Review**: Presents a comprehensive examination of existing studies on skin lesion detection and classification, machine learning and deep learning techniques applied to medical imaging, and relevant mobile health (mHealth) application frameworks. This chapter critically evaluates the strengths and weaknesses of prior approaches, identifying gaps that this study seeks to address.
- **Chapter 3 – Methodology**: Describes the research design, dataset collection and curation process, preprocessing and augmentation strategies, CNN architecture selection, training procedures, and performance evaluation metrics. It also outlines the mobile application and CLI implementation details.
- **Chapter 4 – Results and Discussion**: Reports and analyzes experimental results, including classification performance, error analysis, and comparative evaluation against existing methods. The discussion interprets these findings in the context of the study objectives and broader literature.

- **Chapter 5 – Conclusion and Future Work**: Summarizes key findings, reiterates the contributions of the study, and outlines limitations alongside recommended directions for future research, including potential clinical integration and large-scale deployment strategies.

This structured organization ensures a logical progression from problem identification through methodological implementation to empirical validation and future considerations, facilitating clarity and coherence for both technical and non-technical readers.

2 Chapter 2: Literature Review

2.1 Introduction

The human skin, as the largest organ of the body, serves as the first line of defense against environmental insults, pathogens, and physical trauma. It performs essential physiological roles including thermoregulation, sensory perception, and immunological protection. However, the skin is susceptible to a wide range of pathological conditions, collectively referred to as skin lesions, which can significantly impair an individual's quality of life and, in severe cases, become life-threatening (Hay et al., 2006). These lesions may arise from infectious agents, autoimmune disorders, genetic anomalies, or prolonged environmental exposure, with their clinical presentation varying from benign and self-limiting forms to aggressive malignancies (Hees & Naafs, 2005).

Globally, skin diseases constitute a substantial public health challenge, ranking among the top causes of non-fatal disease burden. According to the Global Burden of Disease Study, conditions such as eczema, acne, and fungal infections collectively affect billions of individuals worldwide, while skin cancers contribute significantly to morbidity and mortality in certain regions (Zafar et al., 2023). In sub-Saharan Africa, and specifically in Cameroon, the prevalence of skin lesions is amplified by climatic factors such as high ultraviolet (UV) index, tropical humidity, and widespread infectious disease exposure, combined with limited access to dermatological specialists and diagnostic facilities (Hay et al., 2006). Rural communities often face the most acute challenges, where early detection and treatment are hindered by geographical, infrastructural, and socioeconomic constraints.

Recent advances in Artificial Intelligence (AI), particularly in computer vision, have opened promising avenues for automated skin lesion analysis. Machine learning and deep learning techniques, particularly Convolutional Neural Networks (CNNs), have demonstrated superior performance in differentiating between lesion types, sometimes rivalling trained dermatologists in diagnostic accuracy (Kassem et al., 2021), (Ogudo et al., 2023). Leveraging such tools in low-resource settings could significantly enhance early diagnosis, improve treatment outcomes, and reduce the long-term health and economic burdens associated with skin diseases.

This chapter reviews the existing literature on skin lesion types, their epidemiology, and the role of AI in dermatological diagnostics, with particular emphasis on the African context. It also outlines the technical underpinnings of CNN-based approaches, dataset challenges, and existing AI-enabled systems for skin lesion detection.

2.2 Overview of Skin Lesions and Skin Cancer

Skin lesions represent a broad spectrum of structural or functional abnormalities affecting the skin's epidermal, dermal, or subcutaneous layers. They are typically classified according to their clinical appearance, etiology, and pathological significance, ranging from benign, non-cancerous conditions to malignant cancers that pose significant health risks (Zafar et al., 2023). Understanding the characteristics, prevalence, and diagnostic challenges associated with each category is crucial for designing AI-assisted diagnostic systems capable of reliable, multi-class classification across diverse lesion types.

2.2.1 Non-Cancerous Lesions

Non-cancerous lesions comprise the majority of dermatological cases encountered in both primary and specialized care. These include inflammatory conditions such as eczema and psoriasis, infectious diseases like fungal infections and bacterial dermatoses, pigmentary disorders such as vitiligo, and benign growths including seborrheic keratoses and skin tags (Hay et al., 2006). While these lesions are typically non-fatal, they can cause chronic discomfort, psychological distress, and social stigma, especially when affecting visible body areas.

In Africa, the high prevalence of infectious dermatoses is closely linked to environmental and socioeconomic conditions. Fungal infections such as tinea capitis are particularly common among children in rural and peri-urban areas due to overcrowding, limited hygiene facilities, and humid climatic conditions (Hees & Naafs, 2005). Similarly, bacterial skin infections like impetigo and cellulitis are widespread, often arising as secondary infections in individuals with compromised skin barriers from insect bites, eczema, or other dermatological conditions.

Chronic inflammatory diseases such as eczema and psoriasis, though less prevalent than infectious

dermatoses, represent a significant health burden due to their recurrent nature and the need for long-term management. In many low-resource settings, limited access to dermatologists and diagnostic tools leads to misdiagnosis, inappropriate treatment, and poor disease control. Pigmentary disorders such as vitiligo and post-inflammatory hyperpigmentation are also of particular concern in African populations, where cultural perceptions and stigma may influence healthcare-seeking behavior (Hay et al., 2006).

From a diagnostic perspective, the visual similarity between non-cancerous and cancerous lesions can pose a challenge, particularly for general practitioners without specialized dermatological training. AI-driven image analysis tools have the potential to assist in differentiating these lesions, improving diagnostic accuracy in settings where specialist input is scarce.

2.2.2 Cancerous Lesions

Cancerous lesions, though less common than non-cancerous forms, are of critical concern due to their potential for metastasis and mortality. The most prevalent malignant skin cancers include melanoma, squamous cell carcinoma (SCC), and basal cell carcinoma (BCC). While melanoma is less frequent in darker skin tones, its prognosis is often poorer in African populations due to late-stage diagnosis (Zafar et al., 2023). SCC and BCC, on the other hand, may occur in sun-exposed or chronically damaged skin regardless of skin tone.

Globally, skin cancers represent a significant proportion of all cancer diagnoses, with an estimated 1.5 million new cases annually (Kassem et al., 2021). In high-income countries, public health campaigns and widespread access to dermatological care have improved early detection rates, leading to better survival outcomes. In contrast, in many African countries, including Cameroon, public awareness of skin cancer remains low, and access to diagnostic biopsies and histopathology services is limited. Consequently, patients often present with advanced disease stages, reducing treatment options and survival prospects.

Risk factors for skin cancer in African populations include albinism, chronic scarring from burns or ulcers, prolonged exposure to UV radiation (especially among outdoor workers), and certain viral infections such as human papillomavirus (HPV). Individuals with albinism, in particular, face a dramatically elevated risk of developing SCC due to the absence of protective melanin in the skin.

Diagnosing skin cancer in low-resource settings is challenging due to both infrastructural and human resource constraints. Dermoscopy, histopathology, and other confirmatory diagnostic modalities may be unavailable outside urban tertiary hospitals. In such contexts, mobile AI-assisted diagnostic tools can play an instrumental role in triaging suspicious lesions, guiding referrals, and facilitating earlier intervention.

2.3 Role of Artificial Intelligence in Dermatology

Artificial Intelligence (AI) has emerged as a transformative force in medical diagnostics, with dermatology standing out as a field particularly well-suited to AI-driven interventions. This suitability stems from the inherently visual nature of dermatological assessment, where clinical diagnosis often relies heavily on visual inspection of lesion morphology, color, texture, and distribution patterns. AI systems, especially those based on computer vision, can replicate and, in some cases, surpass human pattern recognition capabilities by analyzing large volumes of annotated images (Kassem et al., 2021), (Ogudo et al., 2023).

In recent years, deep learning algorithms have demonstrated diagnostic performance comparable to, and occasionally exceeding, that of board-certified dermatologists (Esteva et al., 2017). These advancements have been facilitated by the availability of large-scale annotated image datasets, improvements in computational power, and algorithmic innovations in neural network architectures. AI-powered systems have been successfully developed to identify and differentiate between multiple skin lesion types, including both malignant and non-malignant conditions, with high sensitivity and specificity.

Beyond diagnostic accuracy, AI offers significant advantages in scalability and accessibility. Once trained, AI models can be deployed on a wide range of devices from high-performance clinical workstations to mobile smartphones, making them particularly valuable in low-resource environments. In rural areas of Cameroon, for example, where access to dermatologists is limited, AI-powered mobile applications could enable preliminary screening and triage, directing patients with suspicious lesions to specialized care.

Another important role of AI in dermatology lies in **decision support**. AI systems can assist clinicians by highlighting regions of interest, providing probability scores for different lesion classes,

and integrating clinical metadata (such as patient age, lesion history, and risk factors) into predictive models. This reduces the likelihood of oversight in busy clinical environments and supports more consistent diagnostic outcomes.

Furthermore, AI-based lesion analysis is not limited to classification. Emerging research is exploring its application in lesion segmentation, disease progression tracking, and even predictive analytics for treatment response. Such capabilities could support not only individual patient management but also large-scale epidemiological surveillance.

Despite these strengths, AI adoption in dermatology is not without challenges. Issues such as dataset bias, lack of diverse representation (especially for darker skin tones), and limited interpretability of deep learning models can affect trust and clinical uptake (Gouda et al., 2022). Ethical and regulatory considerations ranging from patient data privacy to liability in the event of misdiagnosis, also remain critical hurdles. Addressing these concerns will be essential for ensuring that AI complements rather than replaces clinical judgment, particularly in sensitive medical contexts.

2.4 Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) represent the cornerstone of modern computer vision and have become the dominant architecture for image classification tasks, including dermatological image analysis. Inspired by the organization of the animal visual cortex, CNNs are designed to automatically learn hierarchical feature representations directly from raw pixel data, eliminating the need for manual feature engineering (LeCun et al., 2015).

A typical CNN architecture comprises several key components:

1. **Convolutional Layers** – These layers apply a series of learnable filters to the input image, producing feature maps that capture spatial hierarchies such as edges, textures, and shapes. Early layers detect simple patterns, while deeper layers learn more complex and abstract features relevant to classification.

2. **Pooling Layers** – Pooling operations (e.g., max pooling, average pooling) reduce the spatial dimensions of feature maps, thereby decreasing computational load and controlling overfitting while retaining the most salient information.

3. **Activation Functions** – Non-linear activation functions, such as ReLU (Rectified Linear Unit), introduce non-linearity into the network, enabling it to learn complex decision boundaries.

4. **Fully Connected Layers** – These layers interpret the high-level features extracted by convolutional and pooling layers, ultimately producing class probability scores through a softmax or sigmoid activation function.

The strength of CNNs in dermatology lies in their ability to extract discriminative features from lesion images, even in the presence of significant variability in lesion size, shape, color, and background noise. This is particularly important for differentiating between lesions with subtle morphological differences, such as differentiating an atypical mole from early-stage melanoma.

In the context of skin lesion classification, CNNs have been successfully applied to both binary classification tasks (e.g., benign vs malignant) and multi-class classification covering a range of lesion types (Kassem et al., 2021). Their performance is further enhanced through techniques such as **data augmentation** (rotations, flips, scaling, color jittering) and **regularization** (dropout, weight decay), which improve generalization and robustness.

However, training CNNs from scratch requires large, balanced datasets, an obstacle in medical imaging where data is often scarce and unevenly distributed across classes. This limitation has driven widespread adoption of **transfer learning**, wherein CNN architectures pre-trained on large general-purpose datasets (such as ImageNet) are fine-tuned on specific medical datasets. Models like ResNet, Inception, and EfficientNet have been extensively leveraged in this manner, yielding strong results even with limited medical image data.

For this project, CNNs form the foundational architecture of the lesion detection and classification system, enabling automated feature extraction and classification in a manner that is both scalable and adaptable to diverse deployment environments. Their proven success in similar tasks makes them a natural choice for addressing the diagnostic challenges associated with skin lesions in Cameroon and other low-resource settings.

2.5 Transfer Learning and ResNet Models

Training deep learning models from scratch typically requires vast datasets containing millions of annotated images. In medical imaging, however, such large-scale datasets are rare due to privacy concerns, the cost of expert annotations, and the limited prevalence of certain conditions (Gouda et al., 2022). This scarcity makes **transfer learning** a practical and effective approach for developing high-performing models with limited domain-specific data.

Transfer learning involves leveraging a model pre-trained on a large, general-purpose dataset such as ImageNet, which contains over 14 million images across 1,000 classes and adapting it to a new but related task (LeCun et al., 2015). In this paradigm, the lower layers of the pre-trained model, which capture generic features like edges, textures, and shapes, are retained, while the higher layers, which learn task-specific features, are fine-tuned on the target medical dataset. This approach reduces training time, lowers computational requirements, and mitigates the risk of overfitting.

Among the various architectures used for transfer learning, **Residual Networks (ResNet)** have become particularly prominent due to their ability to train very deep networks without succumbing to the vanishing gradient problem (He et al., 2016). ResNet's innovation lies in its *residual blocks*, which introduce shortcut connections that allow gradients to flow more easily through the network during backpropagation. This architecture enables the training of models with hundreds of layers while maintaining stability and accuracy.

ResNet-18, ResNet-34, and ResNet-50 are among the most widely used variants in medical imaging applications, each offering a trade-off between computational complexity and representational power. For skin lesion classification, ResNet architectures have consistently demonstrated strong performance, particularly when fine-tuned with domain-specific data and augmented through preprocessing techniques such as rotation, scaling, and color normalization (Kassem et al., 2021).

In the context of this study, ResNet-18 was selected for its balance between accuracy and computational efficiency, making it suitable for deployment on both research environments and resource-constrained devices like smartphones. Transfer learning with ResNet-18 allows the model to benefit from robust, pre-learned visual features while adapting to the nuances of dermatological imagery from diverse sources, including those representing darker skin tones and non-cancerous lesions that

are often underrepresented in global datasets.

2.6 Dataset Challenges in Medical Imaging

The success of AI systems in medical imaging is intrinsically tied to the quality, diversity, and representativeness of the datasets used for training and evaluation. In dermatology, this presents several significant challenges.

1. Data Scarcity and Class Imbalance – High-quality, annotated dermatological datasets are limited, particularly for rare conditions and underrepresented demographics. Publicly available datasets such as those on Kaggle or DermNetNZ often contain disproportionate numbers of images for certain lesion types, leading to class imbalance. Models trained on such datasets risk becoming biased toward majority classes, resulting in reduced accuracy for minority categories (Gouda et al., 2022).

2. Limited Skin Tone Representation – Many benchmark datasets are heavily skewed toward lighter skin tones, reflecting their origins in high-income, predominantly Caucasian populations (Gouda et al., 2022). This lack of diversity can lead to systematic biases, where models perform well on lighter skin but poorly on darker tones—a critical limitation when deploying AI tools in African contexts.

3. Variability in Image Acquisition – Images in dermatology can be captured under a wide range of conditions, including variations in lighting, background, focus, and resolution. Differences in equipment from professional dermatoscopes to smartphone cameras, introduce further heterogeneity, making model generalization more difficult (Zafar et al., 2023). Robust preprocessing pipelines, including color correction, normalization, and augmentation, are essential to mitigate these effects.

4. Annotation Quality and Consistency – Accurate labeling of lesion types often requires expert dermatological input. In public datasets, annotations may be inconsistent or based solely on visual inspection rather than biopsy-confirmed diagnoses. This introduces noise into the training data, potentially lowering model performance (Hees & Naafs, 2005).

5. Ethical and Privacy Concerns – Medical image datasets must comply with strict privacy reg-

ulations to protect patient identities. This can limit the sharing of comprehensive datasets, particularly those including metadata such as patient age, sex, and medical history. De-identification processes, while necessary, may also remove potentially valuable contextual information.

For this project, these challenges were addressed through **multi-source dataset compilation**, selecting high-quality images from multiple Kaggle repositories and the DermNetNZ image database. Images were chosen based on clarity, annotation reliability, and diversity in lesion presentation. Preprocessing steps, including augmentation techniques, were applied to improve model robustness. While this approach does not entirely eliminate dataset-related limitations, it provides a strong foundation for training a model capable of handling real-world variability in skin lesion imagery.

2.6.1 Review of Existing Systems

The development of automated skin lesion detection systems has evolved significantly over the past three decades, transitioning from rule-based image processing algorithms to advanced deep learning frameworks capable of multi-class classification. Early computer-aided diagnosis (CAD) systems, developed in the late 1990s and early 2000s, primarily relied on handcrafted features such as color histograms, shape descriptors, and texture measures, combined with classical classifiers like Support Vector Machines (SVMs) and k-Nearest Neighbors (k-NN). While these systems achieved moderate success in controlled environments, their reliance on manually engineered features made them highly sensitive to variations in lighting, resolution, and lesion morphology (Gouda et al., 2022).

Over the past decade, deep learning, particularly Convolutional Neural Networks (CNNs) has revolutionized dermatological image analysis. CNN-based models eliminate the need for manual feature extraction by learning hierarchical feature representations directly from pixel data (LeCun et al., 2015). Notable milestones include the work of Esteva et al., who demonstrated dermatologist-level classification of skin cancer using a CNN trained on over 129,000 clinical images (Esteva et al., 2017), and the introduction of transfer learning approaches that significantly reduce the data requirements for effective model training (He et al., 2016).

Several prominent publicly available systems illustrate the current state of the art. For example, the **ISIC (International Skin Imaging Collaboration)** challenge platforms have driven substan-

tial progress by providing standardized datasets and benchmarking opportunities for skin lesion classification and segmentation tasks. Solutions from top-performing teams often integrate deep CNN architectures such as ResNet, Inception, and EfficientNet, combined with ensemble learning and advanced data augmentation techniques.

In commercial and practical deployment contexts, mobile applications such as *SkinVision* and *Miiskin* leverage AI to provide lesion risk assessments directly to users via smartphone cameras. While these tools have expanded public access to preliminary screening, they often focus on binary classification (e.g., suspicious vs. non-suspicious) and may not comprehensively cover the range of benign and malignant lesion types relevant in diverse populations, particularly in low-resource settings like Cameroon.

Within academic research, several studies have explored multi-class lesion classification. Kassem et al. proposed a CNN-based framework achieving high accuracy across seven lesion classes in the HAM10000 dataset (Kassem et al., 2021), while Ogudo et al. adapted a transfer learning approach for African skin tones, underscoring the importance of dataset diversity (Ogudo et al., 2023). These studies highlight the growing focus on inclusivity and generalizability in AI dermatology systems.

Despite these advancements, limitations persist. Many existing systems are trained on datasets that lack representation of darker skin tones, have a narrow lesion type coverage, or fail to address deployment constraints such as offline functionality and low computational resources. Furthermore, while commercial solutions often reach end-users quickly, they may not undergo the same level of clinical validation or peer-reviewed scrutiny as academic research outputs.

2.7 Research Gap

While the progress in AI-powered dermatological diagnostics over the last decade has been remarkable, several gaps remain, particularly in the context of African healthcare systems.

1. Dataset Representation – Most high-performing models are trained on datasets dominated by images from lighter skin tones. This limits their diagnostic accuracy when applied to African populations, where lesion appearance can differ significantly due to higher melanin content and unique environmental exposure patterns (Gouda et al., 2022). The lack of comprehensive datasets

covering both malignant and non-malignant lesions in darker skin is a critical barrier.

2. Multi-Class Coverage – Many commercial and academic systems prioritize melanoma detection due to its high mortality rate in fair-skinned populations. However, in African contexts, non-cancerous lesions, such as fungal infections, eczema, and pigmentary disorders, constitute the majority of dermatological cases and therefore require equal diagnostic attention.

3. Deployment in Low-Resource Settings – Existing systems often assume reliable internet connectivity and high-performance computing infrastructure, making them unsuitable for rural or underserved regions. Offline-capable, computationally efficient models are needed to bridge this gap.

4. Clinical Integration and Trust – Limited collaboration between AI developers and healthcare providers in Africa has slowed the integration of AI tools into clinical workflows. Moreover, clinicians may be hesitant to adopt AI systems without clear interpretability, validation on local datasets, and regulatory approval.

5. Comprehensive Diagnostic Workflow – While classification accuracy is a major focus of existing research, other aspects of the diagnostic process such as lesion segmentation, tracking over time, and integration with patient health records, remain underexplored in resource-constrained contexts.

This study addresses these gaps by developing a **multi-class skin lesion classification system** optimized for deployment in Cameroon. The approach combines a curated dataset from multiple open-access sources (Kaggle repositories and DermNetNZ), robust preprocessing and augmentation, transfer learning using ResNet-18, and deployment pathways that accommodate both research and mobile application environments. By doing so, it aims to improve diagnostic accuracy across a broader range of lesions, enhance representation for darker skin tones, and enable accessibility in low-resource healthcare settings.

2.8 Summary

This chapter has provided a comprehensive review of the literature relevant to the development of AI-powered skin lesion detection and classification systems. The discussion began with an overview of skin lesions, distinguishing between non-cancerous and cancerous categories, and em-

phasizing their prevalence and diagnostic challenges both globally and within the African context. The role of AI in dermatology was examined, with a particular focus on CNN architectures, transfer learning strategies, and the ResNet model family, which has proven highly effective in medical imaging tasks.

Dataset challenges in medical imaging were explored in detail, highlighting issues such as class imbalance, limited representation of darker skin tones, variability in image acquisition, and annotation quality. The review of existing systems illustrated the evolution of lesion detection technologies from early handcrafted-feature approaches to modern deep learning-based solutions, while also identifying their limitations in terms of inclusivity, deployment feasibility, and comprehensive lesion coverage.

Finally, the research gap analysis underscored the need for multi-class, skin-tone-inclusive, resource-efficient diagnostic systems that can be deployed effectively in low-resource settings such as Cameroon. The following chapter will present the methodology adopted for this study, detailing the dataset compilation, preprocessing, model training, evaluation, and deployment strategies used to develop the proposed skin lesion classification system.

3 Chapter 3: Methodology

3.1 Introduction

The methodology adopted for this research project provides a systematic framework for developing, training, and evaluating a skin lesion detection and classification system. The primary objective of the methodology is to ensure that each stage of the project, from data acquisition to model deployment is structured, reproducible, and grounded in established machine learning best practices. The complexity of dermatological image analysis, combined with the variability of skin lesion types, demands a rigorous approach to both data handling and algorithmic design. In this study, the methodology not only addresses the technical aspects of convolutional neural network (CNN) implementation but also incorporates considerations for application development, user interaction, and deployment.

Skin lesion classification presents unique challenges. Unlike standard image classification tasks, dermatological images often exhibit high intra-class variability and subtle inter-class differences. Conditions such as **melanoma** and **nevus** can appear visually similar under certain imaging conditions, while diseases like **chickenpox**, **measles**, and **monkeypox** may share overlapping lesion morphologies. This complexity is further compounded by variations in image acquisition methods, lighting conditions, skin tones, and anatomical locations. A robust methodology must therefore be capable of addressing these challenges through appropriate data preprocessing, feature extraction, and model optimization techniques.

The development of this system follows a structured lifecycle that draws upon the **CRoss Industry Standard Process for Data Mining (CRISP-DM)** methodology, widely recognized for its flexibility and applicability to data science projects. The CRISP-DM model consists of six iterative phases: **Business Understanding**, **Data Understanding**, **Data Preparation**, **Modeling**, **Evaluation**, and **Deployment**, each of which is adapted to the specific context of this project. This structure facilitates not only the creation of a high-performing CNN-based classifier but also ensures that the final product can be integrated into practical diagnostic workflows through both a Command-Line Interface (CLI) and a mobile application.

A key differentiator of this project, compared to many prior works (Khan et al., 2021; Ogudo et al., 2023; Zafar et al., 2023), is the scope of the dataset. While datasets such as HAM10000 primarily focus on malignant and benign skin tumors, the dataset in this study encompasses a broader range of dermatological conditions, including viral, bacterial, fungal, autoimmune, and parasitic skin diseases. This comprehensive dataset, covering **33 distinct categories** such as **Acne**, **Actinic Keratosis**, **Ringworm**, **Leprosy**, **Basal Cell Carcinoma**, and **Vitiligo**, allows for a more inclusive diagnostic support system capable of handling diverse clinical scenarios.

The methodology outlined in this chapter provides the foundation for:

- Efficient preprocessing and augmentation techniques to enhance dataset diversity.
- Fine-tuning of a **ResNet-18** architecture pre-trained on ImageNet to adapt to the specific features of dermatological imagery.
- Training strategies incorporating **class weighting**, **early stopping**, and **learning rate scheduling** to improve generalization and avoid overfitting.
- Evaluation protocols employing multiple performance metrics, including **accuracy**, **precision**, **recall**, and **F1-score**.
- Integration of the trained model into an accessible and user-friendly application environment.

Ultimately, the methodological approach ensures that the system is both technically sound and practically relevant, with the potential for deployment in clinical and teledermatology contexts.

3.2 Project Framework and Approach

The **CRoss Industry Standard Process for Data Mining (CRISP-DM)** methodology serves as the guiding framework for the development of the skin lesion detection and classification system. Originally designed for data mining projects, CRISP-DM has been widely adapted in the field of machine learning and artificial intelligence due to its structured yet flexible nature (Ruparelia, 2010; Shafiq et al., 2021). The six phases of CRISP-DM, **Business Understanding**, **Data Understanding**, **Data Preparation**, **Modeling**, **Evaluation**, and **Deployment**, are not strictly linear but rather iterative, allowing for feedback loops and refinements at any stage.

3.2.1 Business Understanding

The primary goal of this project is to develop an automated skin lesion classification system that can aid healthcare professionals and patients in the early detection and differentiation of dermatological conditions. The motivation stems from several factors:

1. **Rising global incidence of skin diseases** – Both malignant and non-malignant skin conditions are increasingly prevalent worldwide, with skin cancer alone affecting millions annually (Esteva et al., 2017; Hay et al., 2006).
2. **Shortage of dermatologists in certain regions** – Many rural and under-resourced areas lack access to specialist care, creating a need for supportive diagnostic tools (Hees & Naafs, 2005).
3. **Potential for teledermatology** – Mobile and cloud-based diagnostic systems can extend dermatological services to remote areas, enabling earlier interventions.

The **business objectives** for this project include:

- Developing a CNN-based classification model capable of identifying multiple dermatological conditions with high accuracy.
- Providing a CLI for researchers and dermatology students to perform batch classification, evaluation, and model experimentation.
- Creating a mobile application for non-technical users, enabling on-device or cloud-assisted skin lesion classification.
- Designing the system to be scalable for future integration with electronic medical records (EMR) and telemedicine platforms.

The **success criteria** for the project are defined by:

- Achieving a **minimum validation accuracy of 75%** across all classes.
- Demonstrating robustness in real-world conditions with varying lighting, angles, and skin tones.
- Ensuring inference times suitable for near real-time use in both CLI and mobile environments.

3.2.2 Data Understanding

The dataset is a curated collection of high-resolution dermatological images representing **33 distinct categories** of skin conditions. Unlike traditional datasets such as HAM10000 (Khan et al., 2021), which primarily focus on tumor detection, this dataset spans multiple disease types including:

- **Viral infections** (e.g., Chickenpox, Monkeypox, Measles, Shingles)
- **Bacterial infections** (e.g., Impetigo, Leprosy, Whitlow)
- **Fungal infections** (e.g., Athlete's foot, Ringworm, Nail fungus)
- **Parasitic infestations** (e.g., Scabies, Tungiasis, Larva Migrans)
- **Inflammatory and autoimmune conditions** (e.g., Eczema, Psoriasis, Autoimmune diseases)
- **Benign and malignant tumors** (e.g., Basal Cell Carcinoma, Squamous Cell Carcinoma, Melanoma, Seborrheic Keratosis)

The images vary in background context, lighting conditions, and anatomical site, reflecting real-world diagnostic challenges. This diversity enhances the model's generalization capability but also necessitates careful preprocessing to normalize image characteristics. Each image is labeled according to expert annotation or verified dataset metadata, ensuring label reliability.

3.2.3 Data Preparation

Data preparation is a critical stage in CRISP-DM, directly influencing model performance. For this project, data preparation involved:

- **Data cleaning** – Removing duplicate images, mislabeled samples, and corrupted files.
- **Dataset partitioning** – Splitting the dataset into training (70%), validation (15%), and testing (15%) sets to ensure robust performance evaluation.
- **Data augmentation** – Applying transformations such as random horizontal flips, small rotations, and color jitter to simulate variations in real-world conditions and mitigate overfitting (LeCun et al., 2015).

The final preprocessing pipeline, implemented using the **Torchvision Transforms API**, included:

- **RandomHorizontalFlip** – to account for left-right orientation differences.
- **RandomRotation (10°)** – to simulate variations in image capture angles.
- **ColorJitter** – to replicate differences in lighting and camera exposure.
- **Resize to (224×224)** – to match the input dimensions expected by ResNet-18.
- **Normalization** – using the mean and standard deviation values from ImageNet, facilitating optimal transfer learning.

3.2.4 Modeling

The **Modeling** phase centered on adapting a pre-trained **ResNet-18** architecture (He et al., 2016) for multi-class skin lesion classification. The transfer learning approach leveraged ImageNet-trained weights, allowing the network to benefit from generalized feature extraction capabilities while focusing fine-tuning on dermatological features. Specific modifications included:

- Unfreezing the **layer4** and **fully connected (fc)** layers for fine-tuning, while freezing earlier layers to retain learned low-level features.
- Replacing the final classification layer with a fully connected layer matching the number of classes (**33**).
- Applying **class-weighted cross-entropy loss** to address class imbalance.
- Utilizing the **Adam optimizer** with a learning rate of 0.001 and **L2 weight decay** to enhance generalization.
- Implementing a **ReduceLROnPlateau** learning rate scheduler to adaptively lower the learning rate upon performance plateaus.

3.2.5 Evaluation

Evaluation metrics for the system include:

- **Accuracy** – for overall classification performance.

- **Precision, Recall, and F1-score** – for per-class performance evaluation.
- **Confusion matrix** – to visualize misclassification patterns.
- **ROC and AUC curves** – .
- **Inference time measurements** – for both CLI and mobile deployment environments.

3.2.6 Deployment

The **Deployment** phase of CRISP-DM for this project covers integration into:

1. **CLI tool** – enabling bulk image classification, metric computation, and dataset testing.
2. **Mobile application** – developed in **React Native**, allowing on-device image capture and cloud-assisted classification.
3. **Cloud hosting** –

Deployment will also involve developing user interfaces for both technical and non-technical audiences, ensuring accessibility while maintaining performance.

3.2.7 CRISP-DM Iteration

CRISP-DM emphasizes iteration. During development, multiple cycles of model training, hyperparameter tuning, and evaluation were conducted. Feedback from early experimental results informed adjustments to data augmentation parameters, class weighting strategies, and model layer unfreezing policies. The iterative nature of this process ensured that each phase benefited from empirical insights, ultimately improving model robustness.

3.3 Dataset Description and Preprocessing

The dataset used in this study was curated from multiple open-source dermatological image repositories and validated clinical sources. It comprises high-quality images covering **33 unique skin condition categories**, spanning viral, bacterial, fungal, parasitic, autoimmune, inflammatory, and neoplastic conditions. Unlike conventional dermatological datasets such as **HAM10000**, which

focus predominantly on melanocytic lesions (Khan et al., 2021; Tschandl et al., 2018), this dataset extends coverage to a broader spectrum of dermatological manifestations, enabling a more versatile diagnostic tool.

3.3.1 Dataset Composition

The dataset categories are as follows:

- **Viral Infections:** Chickenpox, Monkeypox, Measles, Shingles, Herpes, Cowpox.
- **Bacterial Infections:** Impetigo, Leprosy, Whitlow.
- **Fungal Infections:** Athlete's Foot, Ringworm, Nail Fungus.
- **Parasitic Infestations:** Scabies, Tungiasis, Larva Migrans.
- **Inflammatory and Autoimmune Conditions:** Eczema, Psoriasis, Autoimmune Disease, Rash Dermatitis, Vitiligo.
- **Benign and Malignant Tumors:** Basal Cell Carcinoma (BCC), Squamous Cell Carcinoma (SCC), Actinic Keratosis, Melanoma, Nevus, Seborrheic Keratosis, Benign Tumor, Moles.
- **Other Dermatological Conditions:** Acne, Neurofibromatosis, Vascular Lesion, Healthy skin.

This broad coverage is particularly valuable for **primary care and teledermatology**, where non-specialist practitioners encounter diverse skin pathologies (Adegun & Viriri, 2020; Zafar et al., 2023).

3.3.2 Data Sources

The dataset was compiled from:

- **Public dermatology datasets** — e.g., HAM10000, ISIC Archive (for melanoma, nevus, keratosis, and carcinoma images).
- **Open-access image repositories** — curated collections for conditions such as chickenpox, monkeypox, and scabies.

- **Institutional contributions** — small-scale, ethically approved clinical image sets for rare conditions such as tungiasis and larva migrans.

All images were reviewed to ensure:

- Minimum resolution of 224×224 pixels before resizing.
- Clear lesion visibility with minimal occlusions.
- Correct disease labeling by source datasets.

3.3.3 Dataset Distribution and Class Imbalance

The dataset is **imbalanced**:
Classes like *Acne* and *Eczema* have large sample counts, while rare conditions such as *Tungiasis* or *Larva Migrans* have far fewer. Such imbalance can bias models toward majority classes, leading to reduced recall for rare categories (Buda et al., 2018; Kassem et al., 2021).

Mitigation strategies applied:

- **Class weighting** — Computed via `sklearn.utils.class_weight` and incorporated into the cross-entropy loss.
- **Data augmentation** — Uniformly applied at runtime, with possible further targeted augmentation for minority classes in future iterations.

3.3.4 3.3.4 Data Partitioning

Images were split into:

- **Training set:** 70%
- **Validation set:** 15%
- **Test set:** 15%

Splitting was **stratified per class** to preserve class proportions across sets, ensuring fair evaluation.

3.3.5 Exploratory Data Analysis (EDA)

Before training, **EDA** was conducted on the processed dataset to assess its suitability:

1. **Class distribution analysis** — Generated bar plots to visualize image counts per class for each dataset split, confirming imbalance patterns.
2. **Sample inspection** — Random samples per class were displayed to verify label correctness and assess intra-class variability.
3. **Image dimension profiling** — Recorded resolutions and aspect ratios from random subsets to guide resizing and padding strategies.

3.3.6 Preprocessing Pipeline

A reproducible preprocessing pipeline was implemented in Python, combining **offline standardization** and **online augmentation**:

3.3.6.1 Offline Processing

- **Image Conversion** — Ensured all images were in RGB mode to match the pretrained ResNet input format.
- **Resizing with Padding** — Used **Lanczos resampling** to resize while preserving aspect ratio, padding with a neutral grey background (128, 128, 128) to reach 224×224 pixels.
- **Format Standardization** — Saved all outputs as high-quality JPEG files for consistency.
- **Directory Structuring** — Images organized into `train`, `val`, and `test` directories by class.

3.3.6.2 Online Augmentation (During Training)

Transformation	Purpose	Parameters
Random Horizontal Flip	Simulate left/right lesion orientation	p = 0.5

Transformation	Purpose	Parameters
Random Rotation	Account for patient positioning differences	±10°
Color Jitter	Simulate lighting variations	brightness=0.2, contrast=0.2, saturation=0.2
Normalization	Match ImageNet statistics	mean=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.225]

These augmentations improve generalization and reduce overfitting (LeCun et al., 2015; Mahbod et al., 2019).

3.3.7 Impact of Preprocessing and Augmentation

The chosen preprocessing and augmentation steps:

- Improved robustness to variations in lighting, camera quality, and orientation.
- Reduced the risk of overfitting by diversifying the training data.
- Preserved lesion geometry, which is essential for dermatology-specific feature learning.

3.4 Model Architecture

The classification model is based on a **fine-tuned ResNet-18** (He et al., 2016), selected for its favorable trade-off between accuracy and computational cost. This architecture is well-suited for mobile and edge deployment due to its small parameter count and competitive accuracy in medical image tasks (Esteva et al., 2017).

3.4.1 Baseline Architecture

ResNet-18 comprises:

- An initial convolutional layer (7×7 kernel, stride 2) followed by batch normalization and ReLU activation.
- Four sequential residual blocks (layers 1–4) with skip connections to facilitate gradient flow.
- Global average pooling, followed by a fully connected (fc) layer for classification.

3.4.2 Transfer Learning Adaptations

To adapt ResNet-18 for multi-class lesion classification:

- **Weights** initialized from ImageNet-pretrained ResNet-18 (`ResNet18_Weights.DEFAULT`).
- **Frozen layers:** Layers 1–3 kept frozen to retain generic visual features.
- **Unfrozen layers:** Layer 4 and the fc layer trained for dermatology-specific adaptation.
- **Modified final layer:** Replaced `fc` with `nn.Linear(512, 33)` for 33-class output.

This selective fine-tuning strategy combines pretrained low- and mid-level feature extraction with high-level feature specialization (Yosinski et al., 2014).

3.4.3 Loss Function and Class Balancing

A **weighted cross-entropy loss** was employed, where each class weight was inversely proportional to its frequency in the training set. This approach ensured that minority classes contributed more to gradient updates, improving detection rates for rare conditions.

3.4.4 Optimization and Regularization

Training configuration:

- **Optimizer:** Adam (`lr=0.001`, `weight_decay=1e-4`).
- **Scheduler:** ReduceLROnPlateau (factor=0.5, patience=2).
- **Early Stopping:** Stop if validation loss fails to improve for 5 epochs.
- **Batch Size:** 32.

Regularization via L2 weight decay and early stopping reduced overfitting, while dynamic learning rate adjustment improved convergence stability.

3.4.5 Rationale for ResNet-18 Selection

ResNet-18 was chosen because:

1. It delivers high accuracy with relatively low inference time.
2. Prior studies confirm ResNet's effectiveness for skin lesion classification (Albahar, 2019; Harangi, 2018).
3. Its moderate depth supports deployment on resource-constrained devices without heavy hardware requirements.

3.5 Training Strategy

The training phase was carefully designed to achieve a balance between **computational efficiency**, **classification accuracy**, and **generalization capability** to unseen images. This section describes the training configuration, rationale for hyperparameter selection, and mechanisms adopted to handle the inherent challenges of medical image classification.

3.5.1 Hardware and Frameworks

All model training was conducted on a GPU-enabled workstation equipped with CUDA support. This significantly reduced training time compared to CPU-based execution, enabling faster experimentation and hyperparameter tuning. The model was implemented in **PyTorch**, chosen for its flexibility and strong community support in research-oriented deep learning tasks. Supporting libraries included:

- **torchvision** for pretrained model loading and image transformation pipelines.
- **scikit-learn** for computing class weights and auxiliary evaluation metrics.
- **tqdm** for real-time progress tracking.
- **TensorBoard** for monitoring loss, accuracy, and learning rate trends over epochs.

3.5.2 Data Loading and Batching

The processed dataset (Section 3.3) was loaded via PyTorch's `DataLoader` class with the following configurations:

- **Batch Size:** 32 images per iteration — a compromise between computational efficiency and GPU memory constraints.
- **Shuffling:** Enabled for the training set to ensure diverse sample exposure per batch, breaking any inherent ordering in the data.
- **Parallelism:** Four worker threads (`num_workers=4`) with a prefetch factor of four, minimizing idle GPU time due to data loading delays.

This setup ensured that each training iteration received a randomized, memory-optimized batch, which has been shown to improve convergence rates (Goodfellow et al., 2016).

3.5.3 Transfer Learning and Layer Freezing

Given the limited availability of dermatological images for certain rare conditions, training a deep convolutional neural network from scratch was deemed impractical. Instead, **transfer learning** was employed using a ResNet-18 architecture pretrained on the ImageNet dataset. The transfer learning strategy included:

- **Freezing layers 1–3:** Preserving generic feature extractors (edges, textures, shapes) learned from the large-scale ImageNet dataset.
- **Unfreezing layer 4 and the fully connected layer:** Allowing adaptation of higher-level feature representations to the unique textures and lesion patterns in dermatological images.
- **Classifier replacement:** Modifying the final fully connected layer to output 33 logits corresponding to the target disease categories.

Selective unfreezing reduced the risk of overfitting while allowing enough flexibility for domain adaptation (Yap et al., 2018).

3.5.4 Loss Function and Class Balancing

A major challenge identified during exploratory data analysis (EDA) was **class imbalance**, with conditions like *Eczema* and *Acne* having significantly more samples than rare diseases such as *Tungiasis* or *Cowpox*. Without correction, the model would bias predictions toward majority classes. To address this:

- **Weighted Cross-Entropy Loss** was used, with per-class weights computed using:

$$w_c = \frac{N}{n_c \cdot K}$$

where:

- N is the total number of training samples,
- n_c is the sample count for class c,
- K is the number of classes.

These weights were passed directly to PyTorch's `nn.CrossEntropyLoss`, ensuring that errors on minority classes contributed proportionally more to gradient updates. This method is well-supported in medical imaging literature for combating imbalance (Krawczyk, 2016).

3.5.5 Optimization and Learning Rate Scheduling

Optimization was handled by the **Adam** optimizer with:

- Learning rate $\alpha = 0.001$,
- Weight decay $= 1 \times 10^{-4}$ for L2 regularization.

Adam was chosen for its adaptive learning rates and robustness to sparse gradients (Adam et al., 2014). To avoid stagnation, the `ReduceLROnPlateau` scheduler was employed, reducing the learning rate by a factor of 0.5 if validation loss failed to improve for two consecutive epochs.

3.5.6 Early Stopping

To prevent overfitting and unnecessary computation, **early stopping** was implemented with a patience of 5 epochs, if no improvement in validation loss was observed during this period, training terminated. This approach aligns with best practices in deep learning model regularization (Prechelt, 1998).

3.5.7 Epochs, Monitoring, and Logging

The model was trained for a maximum of **20 epochs**. Real-time metrics were logged to TensorBoard, capturing:

- Training and validation loss curves,
- Training and validation accuracy curves,
- Learning rate schedules.

3.6 Evaluation Strategy

The evaluation methodology was designed to quantify not only the **overall classification performance** but also the **per-class diagnostic reliability**, given the medical context of the task.

3.6.1 Test Set Protocol

All reported results are based on the **test set**, which was strictly isolated from training and validation data to ensure unbiased performance estimates. This set simulates real-world deployment conditions by containing images unseen by the model at any training stage.

3.6.2 Evaluation Metrics

Given the dataset imbalance and the multi-class nature of the problem, we employed the following metrics:

- **Accuracy:**

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

While intuitive, accuracy alone may mask poor minority-class performance in imbalanced datasets.

- **Precision (per class and macro-averaged):**

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

Precision is critical in medical applications where false positives can lead to unnecessary anxiety or treatment.

- **Recall / Sensitivity:**

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

High recall ensures that most true cases are detected, reducing missed diagnoses.

- **F1-Score:**

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

Balances the trade-off between precision and recall.

- **Confusion Matrix:**

Visualizes per-class prediction performance and highlights common misclassifications (e.g., *Seborrheic Keratosis* misclassified as *Actinic Keratosis*).

- **ROC AUC (One-vs-Rest):** Measures separability between each class and all others, providing insight into decision threshold robustness.

3.6.3 Rationale for Metric Selection

These metrics were selected based on clinical relevance. In dermatology AI, high recall is often prioritized to minimize missed malignant cases, while high precision reduces unnecessary biopsies or referrals. Macro-averaging ensures all classes contribute equally to the overall performance measure, regardless of prevalence.

3.6.4 Additional Performance Considerations

Given the intended integration into a **CLI** for research and a **React Native** mobile application, additional evaluations were made on:

- **Inference latency** — to ensure acceptable response times on standard mobile hardware.
- **Model size and memory footprint** — to confirm feasibility for mobile and edge deployment.

3.6.5 Visualizations

Performance visualizations will be added to the final report:

- Loss and accuracy curves across epochs,
- Confusion matrix heatmap for error analysis,
- ROC curves for multi-class separability assessment.

These plots aid in communicating results to both technical and clinical audiences.

4 Chapter 3: Methodology

- 3.1 Introduction
- 3.2 Project Framework and Approach (e.g., CRISP-DM or Nunamaker)
- 3.3 Dataset Description and Preprocessing
- 3.4 Model Architecture (ResNet18 and Modifications)

- 3.5 Training Strategy
- 3.6 Evaluation Metrics
- 3.7 Tools and Libraries Used
- 3.8 Summary

5 Chapter 4: System Design and Implementation

- 4.1 System Overview
- 4.2 Functional Requirements
- 4.3 System Architecture Diagram
- 4.4 Data Flow and Pipeline
- 4.5 CLI and GUI Components
- 4.6 Integration and Deployment
- 4.7 Inference and Prediction Logic
- 4.8 Summary

6 Chapter 5: Testing, Results, and Discussion

- 5.1 Dataset Splits (Train, Validation, Test)
- 5.2 Accuracy, Loss, and Confusion Matrix
- 5.3 ROC, Precision, Recall, F1-score
- 5.4 Model Limitations and Observed Bias
- 5.5 Comparison with Existing Systems
- 5.6 Summary

7 Chapter 6: Conclusion and Recommendations

- 6.1 Summary of Findings
- 6.2 Conclusion
- 6.3 Contributions of the Study
- 6.4 Recommendations for Future Work
- 6.5 Final Remarks

8 Appendices

- Appendix A: Source Code Snippets
- Appendix B: Model Configurations
- Appendix C: Training Graphs
- Appendix D: Screenshots of GUI
- Appendix E: User Manual

References

Adam, K. D. B. J. et al. (2014). A method for stochastic optimization. *arXiv Preprint arXiv:1412.6980*, *1412*(6).

Adegun, A. A., & Viriri, S. (2020). FCN-based DenseNet framework for automated detection and classification of skin lesions in dermoscopy images. *IEEE Access*, *8*, 150377–150396.

Albahar, M. A. (2019). Skin lesion classification using convolutional neural network with novel regularizer. *IEEE Access*, *7*, 38306–38313.

Buda, M., Maki, A., & Mazurowski, M. A. (2018). A systematic study of the class imbalance problem in convolutional neural networks. *Neural Networks*, *106*, 249–259.

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115–118.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press. https://www.deeplearningbook.org/

Gouda, W., Sama, N. U., Al-Waakid, G., Humayun, M., & Jhanjhi, N. Z. (2022). Detection of skin cancer based on skin lesion images using deep learning. *Healthcare*, *10*, 1183.

Harangi, B. (2018). Skin lesion classification with ensembles of deep convolutional neural networks. *Journal of Biomedical Informatics*, *86*, 25–32.

Hay, R., Bendeck, S. E., Chen, S., Estrada, R., Haddix, A., McLeod, T., & Mahé, A. (2006). Skin diseases. *Disease Control Priorities in Developing Countries. 2nd Edition*.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778.

Hees, C. van, & Naafs, B. (2005). *Common skin diseases africa*.

Jinnai, S., Yamazaki, N., Hirano, Y., Sugawara, Y., Ohe, Y., & Hamamoto, R. (2020). The development of a skin cancer classification system for pigmented skin lesions using deep learning. *Biomolecules*, *10*(8), 1123.

Kassem, M. A., Hosny, K. M., Damaševičius, R., & Eltoukhy, M. M. (2021). Machine learning and deep learning methods for skin lesion classification and diagnosis: A systematic review. *Diagnostics*, *11*(8), 1390.

Khan, M. A., Muhammad, K., Sharif, M., Akram, T., & Albuquerque, V. H. C. de. (2021). Multi-class skin lesion detection and classification via teledermatology. *IEEE Journal of Biomedical and Health Informatics*, *25*(12), 4267–4275.

Krawczyk, B. (2016). Learning from imbalanced data: Open challenges and future directions. *Progress in Artificial Intelligence*, *5*(4), 221–232. https://doi.org/10.1007/s13748-016-0094-0

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444.

Mahbod, A., Schaefer, G., Wang, C., Ecker, R., & Ellinge, I. (2019). Skin lesion classification using hybrid deep neural networks. *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1229–1233.

Ogudo, K. A., Surendran, R., & Khalaf, O. I. (2023). Optimal artificial intelligence based automated skin lesion detection and classification model. *Computer Systems Science & Engineering*, *44*(1).

Prechelt, L. (1998). Early stopping—but when? *Neural Networks: Tricks of the Trade*, 55–69. https://doi.org/10.1007/3-540-49430-8_3

Ruparelia, N. B. (2010). Software development lifecycle models. *ACM SIGSOFT Software Engineering Notes*, *35*(3), 8–13.

Shafiq, S., Mashkoor, A., Mayr-Dorn, C., & Egyed, A. (2021). A literature review of using machine learning in software development life cycle stages. *IEEe Access*, *9*, 140896–140920.

Tschandl, P., Rosendahl, C., & Kittler, H. (2018). The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific Data*, *5*(1), 1–9.

Yap, J., Yolland, W., & Tschandl, P. (2018). Multimodal skin lesion classification using deep learning. *Experimental Dermatology*, *27*(11), 1261–1267.

Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks? *Advances in Neural Information Processing Systems*, *27*.

Zafar, M., Sharif, M. I., Sharif, M. I., Kadry, S., Bukhari, S. A. C., & Rauf, H. T. (2023). Skin lesion analysis and cancer detection based on machine/deep learning techniques: A comprehensive survey. *Life*, *13*(1), 146.