Multi-Model Deep Learning Approach for Monkeypox Detection: Evaluating Base and Enhanced CNN Architectures

By

Mahmud Kabir Yousuf
201-15-13797

FINAL YEAR DESIGN PROJECT REPORT

This Report Presented in Partial Fulfillment of the Requirements for the **Degree of Bachelor of Science in Computer Science and Engineering**

Supervised by

Mr Shah Md. Tanvir Siddiquee
Assistant Professor
Department of Computer Science and Engineering
Daffodil International University

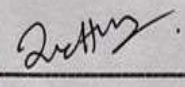
DAFFODIL INTERNATIONAL UNIVERSITY
Dhaka, Bangladesh

May 14, 2025

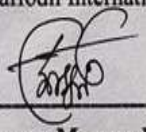
APPROVAL

This Project titled **"Multi-Model Deep Learning Approach for Monkeypox Detection: Evaluating Base and Enhanced CNN Architectures"**, submitted by Mahmud Kabir Yousuf, ID No: **201-15-13797** to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on **14 May, 2025**.

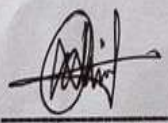
BOARD OF EXAMINERS



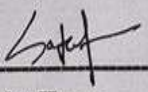
Dr. Md. Zahid Hasan — **Chairman**
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University



Dewan Mamun Raza — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University



Mr. Md. Aynul Hasan Nahid — **Internal Examiner**
Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University



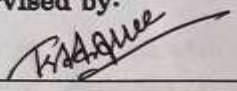
Sadat Hasan — **External Examiner**
Data Scientist (Senior Principle Officer)
Risk Management Division Computer Science and
BRAC Bank

S

DECLARATION

We hereby declare that this project has been done by us under the supervision of **Mr Shah Md Tanvir Siddiquee, Assistant Professor**, Department of Computer Science and Engineering, Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:



Mr Shah Md Tanvir Siddiquee
Assistant Professor
Department of Computer Science and Engineering Daffodil International University

Co-Supervised by:

Chayti Saha
Lecturer
Department of Computer Science and Engineering Daffodil International University

Submitted by:

Mahmud Kabir Yousuf
Student ID: 201-15-13797
Department of Computer Science and Engineering
Daffodil International University

ACKNOWLEDGEMENTS

This work would not have been possible without the support and contributions of many individuals over the past two semesters. We are deeply grateful to everyone who has assisted us in one way or another.

First, we express our heartfelt thanks and gratefulness to the almighty for His divine blessing making it possible for us to complete the **Final Year Design Project (FYDP)** successfully.

We are grateful and wish our profound indebtedness to **Mr Shah Md Tanvir Siddiquee, Assistant Professor** Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh. Deep knowledge and keen interest of our supervisor in the field of **Deep learning** to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting themat all stages have made it possible to complete this project.

We would like to express our heartfelt gratitude to the Head of the Department of Computer Science and Engineering, for his kind help in finishing our project and also to other faculty members and the staff of the Department of Computer Science and Engineering, Daffodil International University.

We would like to thank our entire course-mates at Daffodil International University, who took part in this discussion while completing the coursework.

Finally, we must acknowledge with due respect the constant support and patience of our parents.

ABSTRACT

Global reemergence of monkeypox has underlined the urgency of rapid, specific and scalable diagnostic solutions. PurposeThis study investigates the possible use of deep learning techniques for automated classification of monkeypox from skin lesion images in a cohort of suspected and confirmed monkeypox patients. The study evaluates five known architecture of Convolutional Neural Networks (CNN)- VGG16, InceptionV3, MobileNet, Xception, and ResNet50 on a dataset of 998 images labelled with monkeypox, chickenpox, measles and normal. Implemented each model in a base and hybrid form, hybrid version improved by attention mechanisms and noise regularization, improves focus on lesions and encourages generalization. Evaluation Metrics for Performance Accuracy, Precision, Recall and F1-score Among all the tested models, the hybrid MobileNet and InceptionV3 models had the highest accuracies of 97% and 96%, respectively, and stable performance in classification. Hybrid versions of other models like Xception and VGG16 also performed very well. On the other hand, the ResNet50 hybrid model performed poorly, suggesting the difficulties adapting that architecture for this problem. The findings from this study corroborates that hybrid deep learning models are immensely beneficial for improving the accuracy and robustness of monkeypox classification from images. Thus, these results highlight the promise of AI-mediated diagnostic tools to address early detection and outbreak management, especially in resource-scarce settings. Research coming down the pipeline will identity suitable prospects for data expansion, transformer architectures, and integrating multimodal clinical data to ultimately yield a more reliable and clinically-applicable diagnostic.

Table of Contents

Approval i

Declaration ii

Acknowledgements iii

Abstract iv

List of Figures vii

List of Tables viii

1 Introduction 1

1.1 Introduction 1
1.2 Motivation 2
1.3 Objectives 3
1.4 Methodology 3
1.5 Project Outcome 4
1.6 Organization of the Report 5

2 Background 7

2.1 Introduction 7
2.2 Literature Review 8
2.2.1 Similar Applications 12
2.2.2 Related Research 13
2.3 Gap Analysis 14
2.4 Summary 16

3 Research Methodology 17

3.1 Methodology/Requirement Analysis & Design Specification 17
3.1.1 Overview 17
3.1.2 Proposed Methodology/ System Design 18
3.1.3 Functional and Nonfunctional Requirements 23
3.1.4 UI Design 24
3.2 Detailed Methodology and Design 25

3.3 Project Plan 29
3.4 Task Allocation 31
3.5 Summary 32

4 Implementation and Results 34
4.1 Environment Setup 34
4.2 Testing and Evaluation/Performance/ Comparative Analysis 35
4.3 Results and Discussion 36
4.4 Summary 47

5 Engineering Standards and Design Challenges 49
5.1 Compliance with the Standards 49
5.1.1 Software Standards 49
5.1.2 Hardware Standards 50
5.1.3 Communication Standards 50
5.2 Impact on Society, Environment and Sustainability 51
5.2.1 Impact on Life 51
5.2.2 Impact on Society & Environment 52
5.2.3 Ethical Aspects 53
5.2.4 Sustainability Plan 53
5.3 Project Management and Financial Analysis 53
5.4 Complex Engineering Problem 54
5.4.1 Complex Problem Solving 54
5.4.2 Engineering Activities 55
5.5 Summary 56

6 Conclusion 57
6.1 Summary 57
6.2 Limitation 58
6.3 Future Work 58

References 60

List of Figures

3.1 Classifier Model Flowchart 18
3.2 Sample Dataset 19
3.3 Class Distribution 20
3.4 Balance Class Distribution 21
3.5 Interface of the System 24
3.6 Evaluating the System for ChickenPox 24
3.7 Evaluating the System for Measles 24
3.8 Evaluating the System for MonkeyPox 25
3.9 Evaluating the System for Normal Skin 25
3.10 InceptionV3 Base Model Architecture 28
3.11 InceptionV3 MultiHead Attention Model Architecture 28
3.12 VGG16 Base Model Architecture 29
3.13 VGG16 MultiHead Attention Model Architecture 29
3.14 MobileNet MultiHead Attention Model Architecture 29
4.1 Accuracy Curve over Epoch for VGG16 base 38
4.2 Loss Curve over Epoch for VGG16 base 38
4.3 Accuracy Curve over Epoch for VGG16 hybrid 38
4.4 Loss Curve over Epoch for VGG16 hybrid 38
4.5 Accuracy Curve over Epoch for Inception base 38
4.6 Loss Curve over Epoch for Inception base 38
4.7 Accuracy Curve over Epoch for Inception hybrid 39
4.8 Loss Curve over Epoch for Inception base 39
4.9 Accuracy Curve over Epoch for ResNet50 base 39
4.10 Loss Curve over Epoch for ResNet50 base 39
4.11 Accuracy Curve over Epoch for ResNet50 hybrid 39
4.12 Loss Curve over Epoch for ResNet50 base 39
4.13 Accuracy Curve over Epoch for Xception base 40
4.14 Loss Curve over Epoch for Xception base 40
4.15 Accuracy Curve over Epoch for Xception hybrid 40
4.16 Loss Curve over Epoch for Xception hybrid 40
4.17 Accuracy Curve over Epoch for MobileNet base 40
4.18 Loss Curve over Epoch for MobileNet base 40
4.19 Accuracy Curve over Epoch for MobileNet hybrid 41
4.20 Loss Curve over Epoch for MobileNet hybrid 41
4.21 Confusion Matrix for VGG16 base 41
4.22 Confusion Matrix for VGG16 hybrid 41
4.23 Confusion Matrix for Inception base 42
4.24 Confusion Matrix for Inception hybrid 42
4.25 Confusion Matrix for ResNet50 base 42
4.26 Confusion Matrix for ResNet50 hybrid 42
4.27 Confusion Matrix for Xception base 42
4.28 Confusion Matrix for Xception hybrid 42
4.29 Confusion Matrix for MobileNet base 43
4.30 Confusion Matrix for MobileNet hybrid 43
4.31 Model Accuracy Comparison 47

List of Tables

2.1 Summary of Literature Review. 8
2.2 Gap Analysis Table. 14
3.1 Task Allocation Table..... 31
4.1 Model Accuracy Table. 35
4.2 Classification report for VGG16 base..... 43
4.3 Classification report for VGG16 hybrid. 43
4.4 Classification report for InceptionV3 base. 44
4.5 Classification report for InceptionV3 hybrid. 44
4.6 Classification report for ResNet50 base 44
4.7 Classification report for ResNet50 hybrid. 45
4.8 Classification report for Xception base..... 45
4.9 Classification report for Xception hybrid. 45
4.10 Classification report for MobileNet base. 46
4.11 Classification report for MobileNet hybrid..... 46
5.1 Estimated Cost and Financial Analysis. 53
5.2 Mapping with complex problem solving. 54
5.3 Mapping with knowledge Profile. 54
5.4 Mapping with complex engineering activities. 55

Chapter 1

Introduction

This chapter gives an outline of the research project, the value of Monkeypox classification, the rationale for automation and the objectives of the project, the deep learning method employed, as well as the anticipated outcomes. As Monkeypox continues to emerge globally, early and accurate diagnosis is crucial for containment of outbreaks and prompt treatment. Classically, clinical diagnostic may be difficult because it resembles other skin diseases like Chickenpox and Measles. To this end, we developed an image classification system using deep learning to diagnose Monkeypox infections. The study's objective is to develop and evaluate convolutional neural networks applied to images of skin lesions to improve diagnostic aid systems, minimize diagnostic errors, and lead to improved healthcare management.

1.1 Introduction

Monkeypox is a viral zoonosis that has emerged as a public health problem in the past few years as there have been a number of outbreaks [1]. It is a type of disease caused by the Monkeypox virus (belonging to the family of Orthopoxvirus) and exhibits signs and symptoms that are closely related to those skin diseases Chickenpox, Smallpox, and Measles. Typical symptoms included fever, enlarged lymph nodes and development of the typical skin lesions [2]. Owing to the interdigitating symptoms, finding the accurate diagnosis relies purely on clinical examination, and is still challenging. Conventional diagnostic techniques (such as polymerase chain reaction [PCR]) and laboratory hosted-assays, though useful, are usually time-consuming, expensive, and not accessible to the particularly remote and resource-limited settings. Misdiagnosis may result in delayed advices, more number of new infected cases, and public health dangerous.

Machine learning and deep learning have made significant progress, and computer-aided diagnosis systems have demonstrated promising performance. In particular, convolutional neural networks (CNNs) and other deep learning models have shown to be effective at medical image classification, and are capable of learning high-level features in images. The use of these technology has immense potential to provide an affordable, scalable, and rapid solution for early detection of Monkeypox.

The objective of the current work is to build and assess a deep learning-based classifier for distinguishing Monkeypox skin lesions from look-alike skin

diseases. An image database of Monkeypox, Chickenpox, Measles, and healthy skin is employed for training and testing. All images will be preprocessed for data quality and model stability. Several deep learning models will be used and compared to find out the best model for classification accuracy.

The goal of this work is to develop a diagnostic support tool to help healthcare practitioners in more accurately and efficiently diagnosing TB. Furthermore, the tool could be incorporated within telemedicine systems or mobile apps to assist in remote diagnostic processes, particularly in underserved areas. Through the relief of existing healthcare systems and facilitation of early detection, the project intends to positively influence public health policy while better controlling the epidemic of Monkeypox.

1.2 Motivation

Recent increases in Monkeypox cases highlight the importance of immediate and accurate diagnoses to prevent the spread of the disease as well as ensure timely treatment and care for patients. Classical techniques for the diagnosis of canine visceral leishmaniasis (CVL), such as clinical evaluation and subsequent confirmation by laboratory tests via carrying out polymerase chain reaction (PCR) studies, are highly sensitive and specific, but can be expensive, time-consuming, and sometimes not available, especially in remote and poor areas. Also, the clinical presentation of Monkeypox closely mirrors other skin disorders such as Chickenpox, Measles and viral exanthems, and diagnosis by appearance alone can result in mistaken diagnoses even among seasoned clinical personnel. These and other issues emphasize the need for automated, high performance, and scalable classification tools to assist in making medical decisions.

Deep learning methods, in particular, Convolutional Neural Networks (CNNs), have shown unprecedented performance in processing and identifying subtle patterns from medical images. Using such technologies to develop automatic identification of Monkeypox could significantly improve diagnostic accuracy, operational efficiency, and access to healthcare. This study has been motivated by the possibility that deep learning models can potentially fill the present diagnostic void, minimize dependence on laboratory infrastructure, and offer rapid and actionable inferences from skin lesion images.

Advantages of this approach are as follows:

i. Rapid Diagnosis: Automated categorization allows for swift recognition

of Monkeypox cases, providing quicker medical response and containment.

ii. Resource Conservation: Clinical institutions may save laboratory tests for equivocal or severe cases, conserving resources.

iii. Better Accuracy: Deep learning systems reduce the instances of misdiagnosis, providing a solid second opinion for healthcare workers.

iv. Increased Accessibility: Smart phone or web-based classifiers can provide diagnostic assistance to isolated and under-served regions.

v. Improved_Outbreak Management: Rapid and accurate detection is tied to more efficient isolation and better public health response.

This study does not just seek to improve clinical outcomes through novel technology however, but to provide healthcare systems with the means to manage and respond more effectively to infectious disease threats such as Monkeypox.

1.3 Objectives

Objectives of this paper are:

i. Develop a deep learning-based model for accurate monkeypox detection.

ii. Achieve high classification accuracy by optimizing model architecture.

iii. Implement data augmentation to address dataset limitations.

iv. Provide a user-friendly, real-time detection system for clinical use.

v. Ensure robust evaluation using multiple performance metrics, including accuracy, precision, recall, and F1-score.

1.4 Methodology

Methodology The approach in this research is to design, train and assess deep learning models for the automated classification of Monkeypox skin lesions. We collected a large dataset from the publicly available Kaggle repositories, including the "Monkeypox Skin Lesion Dataset" and the "Monkeypox Skin Image Dataset." The datasets were combined after collection of the data, thereby having an equal-sized dataset containing four categories: Monkeypox, Chickenpox, Measles or Normal (healthy skin).

Preprocessing was performed to improve both quality and diversity of the data. These steps involve resizing the images to a fixed input size, normalizing the pixel values and wide-ranging data augmentation including flipping, rotation, and zooming, shift to promote generalization and reduce overfitting. We split the rest as a training and testing set.

Five popular convolutional neural network (CNN) architectures, including InceptionV3, Xception, MobileNet, ResNet and VGG16, were chosen for model building. Each architecture was developed into two versions:

Base model: pre-trained model without any adaptation, for purpose of comparison.
Hybrid model: A refine version adding attention mechanisms to the architecture for lesion-relevant features focusing and noise regularization to enhance robustness towards noisy training data.

Final models were tested and evaluated by classical performance indicators (accuracy, precision, recall, F1-score and confusion matrices). Training and validation loss/accuracy curves were plotted to observe the learning behavior. Comparative analysis across all architectures was also included to evaluate the effect of hybrid improvements over the baseline models.

The objective of the proposed approach is to identify the best model for Monkeypox categorisation in order to construct a scalable and deployable diagnostic tool applicable to practical clinical scenarios, particularly in low-resource settings.

1.5 Project Outcome

The project's main goal is to develop an advanced deep learning-based system capable of accurately classifying Monkeypox infections from skin lesion images. By combining powerful convolutional neural network architectures such as InceptionV3, Xception, MobileNet, ResNet, and VGG16, along with enhancements like attention mechanisms and noise regularization, the study offers a scalable solution for rapid and reliable infectious disease diagnosis. The resulting system will be highly accurate and can be deployed across various platforms, including telemedicine services, mobile health applications, and clinical diagnostic support systems. Even with limited and variable-quality images, the model will be able to differentiate Monkeypox from visually similar conditions such as Chickenpox, Measles, and other skin diseases. Healthcare providers, public health organizations, and diagnostic centers stand to benefit significantly from such an automated diagnostic tool. The system can improve early detection, reduce diagnostic delays, and

contribute to better outbreak control and patient management. Furthermore, by training on a diverse dataset, the study aims to uncover important visual and structural features that distinguish Monkeypox lesions from other conditions. These insights may assist researchers, dermatologists, and epidemiologists in understanding the critical clinical markers and visual patterns associated with Monkeypox infections, ultimately contributing to stronger public health surveillance and response systems.

1.6 Organization of the Report

This chapter gives an integrated introduction into the organisation of the report by describing the composition and content of each of the chapters in order to orient the reader through this study.

Chapter 1 Introduction: This chapter outlines the research problem, highlights the increasing significance of rapid MPX diagnosis, and justifies the rationale for automating diagnostics. It provides the objectives for the project, methodology used, the range of research, the expected results and the chapter organization. This chapter forms the basis for the entire work by detecting the necessity of a deep learning approach for infectious disease classification.

Chapter 2 Background: This chapter provides an overview of the clinical and epidemiological aspects of Monkeypox and addresses diagnostic challenges. It discusses current diagnosis approaches and recent advances in applying deep learning to medical image analysis. A literature review, presenting the gaps of existing researches, and areas with much room for improvement. Summary of implications from the literature and the contribution of the work undertaken is provided to summarize the chapter.

Chapter 3 Research design and methodology: The research methodology used in the work is described in this chapter. It includes how dataset is collected from Kaggle repositories, data merge, pre-processing techniques and augmentation methods. Detailed description of the choice of the deep learning architectures: InceptionV3, Xception, MobileNet, ResNet and VGG16, crew base and hybrid models. The chapter also presents model training, hyperparameters tuning, evaluation metrics, functional and non-functional requirements and system architecture with the help of diagrams, flowcharts.

Chapter 4 Implementation and Results: In this chapter we show the experimental realization of the models, and we describe the development environment, software libraries, and hardware setup utilized. It includes the training, validation and testing. The performance is evaluated using the training and validation curves, classification reports, confusion matrices and comparative table which points out the

advantages and disadvantages of Base and Hybrid models. The optimal model is selected and described.

Chapter 5 Engineering standards; Design difficulties: In this chapter, the engineering of the rack and the ethics are described. It presents issues and difficulties in data processing, model tuning and evaluation. The consequences of AI-based healthcare applications for the society are evaluated, by highlighting how the proposed system could boost early outbreak detection and improve public health management. We also investigate environmental and sustainability issues regarding the use of computing resources. Also, the chapter has something to offer in terms of project management, budgeting, and resource management.

Chapter 6: conclusions and future work: In conclusion, the conclusions of the entire research work will be presented emphasizing the main accomplishments, findings and the contributions. It details the limitations encountered in the study, and highlights possible directions for future research, which may include using larger datasets, improving model generalization across different population settings, and directly applying the model in a real clinical settings. The chapter ends with a consideration of the general implications of the project and its potential for infectious disease and healthcare innovation.

Chapter 2

Background

A comprehensive review of the Monkeypox taxonomy is covered in this chapter - a detailed summary of its clinical and diagnostic presentation, existing deep learning practices, as well as a comprehensive literature review of existed work on the skin lesion analysis. A gap analysis is made to identify the shortcomings of existing approaches, focusing on the lack of data and model robustness. Furthermore, this chapter will detail the methods and enhancements used in this work including: attention mechanisms and noise regularization by which to bettering the accuracy and reliability of the automated Monkeypox detection task.

2.1 Introduction

Early detection of clinical signs in form of skin lesions for Monkeypox has recently attracted research attention on large due of global spread beyond the endemic zones. Monkeypox, attributed to Monkeypox virus (MPXV) has clinical manifestations that appear similar to other viral diseases including Chickenpox and Measles and is difficult to diagnose in a timely manner. Automated diagnostic systems using deep learning are increasingly viewed as an alternative to costly, time-consuming, and frequently unavailable traditional diagnostic approaches particularly in low resource settings. Notwithstanding, high accuracy of Monkeypox classification is still challenging due to some technical and clinical issues.

The primary challenge of building a robust deep learning model for Monkeypox detection is the scarcity of annotated high-quality data. Monkeypox is an uncommon skin condition and characterized by a small sample size and imbalance of the datasets, which limits the generalization of the model. This dearth of data increases the chance of overfitting, where the model does well on the training data, but performs poorly in generalization to unseen cases. Moreover, the variation of lesions (with varying patient demographics, infection phase, and under varying imaging conditions) makes the lesion classification task even more challenging.

Earlier research have used traditional convolutional neural networks (CNNs) and transfer learning methods such as ResNet, VGG16 and InceptionV3 pretrained on large image datasets. Although these methods have been

successful to a large extent, there are still difficulties in accurately differentiating Monkeypox lesions from other similar diseases, particularly when the lesions have the same visual characteristics. Furthermore, most previous works do not consider more advanced mechanism such as: attention layers, that can learn to concentrate on the most informative regions of an image, and noise regularization, that can make models robust to containated or low-quality data.

In order to fill these gaps, we investigate both base and hybrid implementations of several CNN architectures InceptionV3 [3], Xception, MobileNet [20], ResNet [9], VGG16 [4]. In order to further improve the feature extraction ability, focus more on critical lesion regions as well as improve the classification performance, we design hybrid models with attention mechanisms and noise reduction techniques. Through extensive model evaluation and comparative analysis, the study seeks to contribute to developing an accurate and scalable Monkeypox classification automated diagnostic system that could facilitate prompt diagnosis and improve outbreak management.

2.2 Literature Review

The literature on Monkeypox classification using machine learning and deep learning techniques is extensive and evolving, with numerous studies exploring various approaches to improve the accuracy, robustness, and efficiency of automated diagnosis from skin lesion images.

Table 2.1: Summary of Literature Reviewed.

Author(s)	Year	Title	Data Set	Machine Learning Model	Key Findings
D. Bala et al.	2023	MonkeyNet: A robust deep convolutional neural network for monkeypox disease detection and classification	Monkeypox Skin Images Dataset (MSID)	DenseNet-201 variant (MonkeyNet)	Achieved 98.91% accuracy with augmentation, strong model interpretability with Grad-CAM. Lacking larger real-world validation.

D. Uzun Ozsahin et al.	2023	Computer-Aided Detection and Classification of Monkeypox and Chickenpox Lesions	Two-class augmented dataset	Custom CNN model (4 Conv + 3 MaxPooling layers)	99.60% accuracy, minimal overfitting with augmentation. Limited class diversity (only two classes).
N. R. Pratama et al.	2025	Feature Fusion with Albumentation for Enhancing Monkeypox Detection	Monkeypox Skin Lesion Dataset (MSLD)	Feature Fusion (Xception + InceptionV3)	85.96% accuracy; ablation studies confirmed importance of feature fusion. Smaller dataset limited generalization.
D. Uzun Ozsahin et al.	2023	Mpox-XDE: Ensemble Model for Monkeypox Detection	Monkeypox Skin Images Dataset (MSID)	Ensemble (Xception + DenseNet201 + EfficientNet B7)	Achieved 98.70% accuracy with Grad-CAM support. Slight computational complexity.
L. H. Huong et al.	2023	A Proposed Approach for Monkeypox Classification	MSLD and MID datasets	Hybrid deep learning + traditional ML classifiers	97% accuracy; enhanced robustness through model combination. Limited scalability tests.
E. H. I. Eliwa et al.	2023	CNN-based Monkeypox Classification with GWO Optimization	Not specified	CNN optimized with Gray Wolf Optimizer (GWO	95.3% accuracy; improved precision and recall. Dataset details were not thoroughly

					disclosed.

This table 2.1 shows all the comparison between previous works and their approaches. In paper [5], adopts a variant of the DenseNet-201 model for the identification of monkeypox in skin lesion images. The authors developed the Monkeypox Skin Images Dataset (MSID), including 770 original and 8689 augmented images of four classes; monkeypox, chickenpox, measles and normal. With the aid of data augmentation and hyperparameter tuning, MonkeyNet was able to achieve a testing accuracy of 98.91% on the augmented dataset and 93.19% on the original dataset. The reported best results of the model were a precision of 96.40%, a recall of 96.30%, and an F1 of 96.30%. We used Grad-CAM visualizations to improve model interpretability and demonstrated the robustness of the model across validation splits, supporting its robustness in clinical decision making. In paper [6] designed a specialized CNN model with four convolutional and three MaxPooling layers to classify monkeypox and chickenpox lesions. Two-class augmented digital skin image datasets were used to train the model. It attained a 99.60% accuracy in testing, with a weighted precision, recall and F1 score of 99.00% each. In the pre-trained models tested, AlexNet was second best with 98.00%, and VGG16 was 80.00%. The model's high performance in all metrics and lack of overfitting because of data augmentation demonstrate its viability to be deployed in real-time diagnostic use, particularly in environments with limited PCR testing facilities. This research [7] provided a feature fusion model that combines Xception and InceptionV3 architectures with augmentation driven by the Albumentation library. For the study, the Monkeypox Skin Lesion Dataset (MSLD) containing 228 images (102 monkeypox and 126 others) was employed. The model proposed attained 85.96% accuracy, 86.47% precision, 85.25% recall, 85.24% F1 score, 78.43% specificity, and 0.8931 in AUC. Data augmentation improved recall from 81.23% to 85.25% and enhanced overall accuracy from 82.02% to 85.96%. Ablation studies validated the importance of feature fusion and augmentation to the improvement of model performance in data-limited setups.

Paper [8] introduced an ensemble deep learning model called Mpox-XDE to classify monkeypox precisely from images of skin lesions. In the research study, the Mpox Skin Images Dataset (MSID) of 770 labeled images was used with four classes: monkeypox, chickenpox, measles, and normal. In the

ensemble, improved variants of the models Xception and DenseNet201 and EfficientNetB7 were combined with Softmax and dense layers with 65% of dropout to prevent overfitting. The resultant model was tested with a 98.70% accuracy, precision of 98.90%, recall of 98.80%, and F1-score of 98.80%. For interpretability enhancement, affected regions in the input images were indicated by utilizing the Grad-CAM method to assist clinicians with early diagnosis and decision-making. In Paper [9] proposed a hybrid method that blended traditional machine learning methods with deep learning methods in monkeypox detection. Different deep models such as ResNet-50, VGG16, and MobileNet were employed in conjunction with classifiers like Random Forest, K-Nearest Neighbors, Naive Bayes, and AdaBoost. Two datasets were employed by the study, namely the Monkeypox Skin Lesion Dataset (MSLD) and the Monkeypox Image Dataset (MID), containing 659 total samples in aggregate. Accuracy of 0.97 and F1-score of 0.98 were attained by the combined model. In this work, the power of combining multiple classifiers to enhance robustness and credibility in practical use was underscored. Paper [10] is a systematic review of the biological and epidemiological aspects of monkeypox covering its classification, genome organization, transmission patterns, and diagnosis. It elaborates on the spread of monkeypox around the globe and especially the ongoing 2022 outbreak and identifies PCR-based diagnosis as the current gold standard. Even though it doesn't offer a model based on machine learning, it points towards the urgency of rapid diagnostic methods since the prevalence of the disease is on the rise and traditional methods of diagnosis are limited. In Paper [11] introduced PoxNet22, which is a fine-tuned model from the InceptionV3 architecture with the use of transfer learning to distinguish monkeypox from other forms of skin illness. In the study, data augmentation was used to prevent overfitting and tested six deep models to determine the top performer. PoxNet22 yielded 100% accuracy, recall, and precision on the preprocessed data set and proved the efficacy of fine-tuned and data augmentation methods in both imbalanced and small data sets.

Paper [12] proposed a CNN-based method of classifying monkeypox skin lesions with Gray Wolf Optimizer (GWO) optimization. Optimization improved the performance of the CNN model with an accuracy of 95.3%, and there were significant increases in precision, recall, F1-score, and AUC over the non-optimized model. This technique has application in accelerating and precise diagnosis in areas where laboratory facilities are less readily available. Paper [13] introduced MOX-NET, which is a multi-stage deep hybrid feature selection and fusion framework to classify monkeypox. It merged features of six pre-trained deep models (such as Vision Transformers, Swin Transformers, ResNet and EfficientNetV2) and used an entropy-controlled firefly algorithm to select the features. MOX-NET was tested on the

MSLID dataset with a remarkable 98.64% accuracy compared to state-of-the-arts. It prioritized contrast enhancement, deep feature fusion, and multi-class SVM classification to achieve solid performance. In paper [14], the researcher designed a custom four-layer CNN to classify monkeypox and chickenpox lesions and tested it using 10,000 augmented images of each class. The model's accuracy turned out to be 99%, beating pre-existing models such as VGG16, ResNet50, and InceptionV3. Its precision, recall, and F1-score of 99% confirm how effective it is. The superior performance was credited by the authors to the simplicity of the model, focused framework, and strong data augmentation, making it highly suitable for clinical decision support when assessing monkeypox in the initial stages.

2.2.1 Similar Applications

Machine learning and deep learning have been employed in the discrimination of infectious diseases via medical images, and this is relevant to Monkeypox classification. Such an example is the Kaggle Monkeypox Skin Image Dataset (MSID) contest, where images from Monkeypox, Chickenpox, Measles, and Normal skin were collected to train machine learning models [16]. Scientists, such as D. Bala et al. proposed the MonkeyNet, which is a DenseNet-based model demonstrating high classification accuracy on these datasets [5].

Aside from research efforts, there are several mHealth applications which have incorporated AI based image classification for early detection of disease. Software tools like DermAssist by Google [17] and SkinVision [20] allow users to photograph their skin including potentially worrying lesion using their smartphones to perform a skin condition analysis. The apps utilize deep learning models that have been pretrained on massive dermatological databases to offer an initial evaluation and portra, thus the expansion of AI for real-time skin-related condition diagnosis. Techniques inverse have been broadly extended in the context of COVID-19 pandemic, where deep learning models were established for the automatic classification of chest X-ray and CT scan images for identifying COVID-19 pneumonia. Studies such as Wang et al. (2020) developed a COVID-Net model, which has made high accuracy to detect COVID-19 from chest X-rays [18]. Another work by Li et al. (2020) introduced an AI system for examining chest CT images to discern COVID-19 pneumonia from other pneumonias to a great extent [19].

Moreover, ensemble models that concatenate architectures such as InceptionV3, Xception, DenseNet have also been respectively used for dermoscopic skin diseases and chest imaging to improve classification robustness. These methods also highlight the expansive role of deep learning in the detection of infectious diseases and support the possibility of

developing dependable Monkeypox diagnostic applications with the same methods. The use of AI in such an application space proves the widening horizon of machine learning outside pure academia and a multiple scalable solution for clinical decision support and particularly in resource-scarce settings. There is a GitHub project by Akshay Bhatia where he collected data from amazon book index. His project has a book index of about 207,572 Books [5]. The books could possibly be categorized into 32 different genres. This project addressed the classification of books only based on their titles without prior author or context information. It was a part of the Spikeway Technologies AI Lab (SAIL). Numerous studies have focused on the effective application of machine learning to the classification of news stories into categories such as politics, sports, and entertainment. For example, to classify news articles with high precision. Their strategy included an extensive amount of preprocessing, such as stemming and stop word removal, and feature selection methods. Support Vector Machines (SVMs) and TF-IDF were used in a study by Shikha Gupta (2021) [6] to achieve excellent categorization accuracy for news articles. Their strategy included a great deal of preprocessing, such as stemming and stop word removal, and feature selection methods like chi-squared testing. Machine learning algorithms are used by websites such as Goodreads and Amazon to suggest books to users based on their reading interests and history. To recommend books in related categories, these platforms frequently use content-based and collaborative filtering strategies.

2.2.2 Related Research

Review of the literature in research clarifies that the classification of monkeypox using machine learning and deep learning algorithms is gaining momentum as a significant and active area of study. Monkeypox has re-emerged and has expanded outside endemic areas, a development that underscores the pressing requirement for rapid accurate diagnostic assays [1]. The diagnosis of Monkeypox is a difficult task, which can be mixed up with similar symptoms related to other pox-infections such as Chickenpox and Smallpox [(2), initial clinical studies dated from as early as 1972], which may as well be due to the small size of anlyzed samples [3]. Recent developments in deep learning have presented promising results. Transfer learning methods, which fine-tune pre-trained convolutional neural network architectures such as VGG-16, InceptionV3, etc, have been successfully utilized in medical image classification tasks [3, 4]. With respect to Monkeypox identification, some research has presented customized deep learning models. For example, with ample data augmentation and hyperparameter tuning, the classification accuracy using MonkeyNet with DenseNet-201 was high [5]. Also, D. Uzun Ozsahin et al. presented a CNN

framework based on standard CNN models for recognition of Monkeypox and Chickenpox lesions [6].

Some feature fusion methods, which involve fusing models such as Xception and InceptionV3, have also been suggested to improve model performance, especially for smaller datasets [7]. However, these progressions also bring some challenges as there are still small 'large-scale' good quality and balanced data. Other class of attempts comprise ensemble deep learning models (Mpox-XDE) that leverage the advantages of combiningmultiple pre-trained architectures to increase classification robustness and explainability as well [10]. Fine-tuning based methods such as PoxNet22 model have delivered almost perfect performance on the edited/sifted dataset [11]. But issues of overfitting and generalization to heterogeneous real-world-database images are still critical. Advanced optimization methods have also been combined with CNNs in order to improve the model accuracy and reliability [12]. Hybrid deep learning structures such as MOX-NET further rely on multi-stage feature selection and fusion by integrating features extracted from several sophisticated models to outperform in classification [13]. However, various studies have reported the variability of model performance on the different dataset and experimental settings. Configuration differences include changes in the calibration of the phantoms and in the positioning of the breast, in the reconstruction algorithm and in the segmentation algorithm that was followed to obtain a model of the breast to simulate and generate the images, and in the cut-off values in the classification algorithm. This observation highlights the need for the standardization of diagnostic methods and development of sensitive and scalable Monkeypox detection solutions.

2.3 Gap Analysis

Although a lot of work on Monkeypox detection has been done using machine learning and deep learning methods, it is far from solved. Against this backdrop, some of the challenges faced in current research include a lack data, model overfitting, generalization to different populations and in consistent evaluation. For instance, early works were restricted to small imbalanced data-sets and did not have adequate data augmentation strategies, which impairs generalization of the models under realistic conditions.

Table 2.2: Gap Analysis Table.

Aspect	Gap	Proposed Methodology
Data Scarcity	Most models are trained on small, imbalanced datasets with limited Monkeypox	Collected and merged datasets from multiple Kaggle sources [15][16],

	cases.	combined with extensive data augmentation to balance classes.
Overfitting and Limited Generalization	Models tend to overfit due to small datasets and lack of diversity in samples.	Implemented hybrid models (attention mechanisms + noise regularization) to improve robustness and generalization.
Limited Labeled Data	Struggles with insufficient labeled data, leading to lower accuracy.	Utilizes semi-supervised learning to incorporate unlabeled data, enhancing model robustness and accuracy.
Model Interpretability	Limited efforts to make deep learning predictions explainable for clinical validation.	Applied Grad-CAM visualization [5][10] to highlight lesion-focused regions, enhancing interpretability for clinicians.
Focus on Single Models	Several works relied on one architecture without hybridization or ensemble approaches.	Designed and compared multiple base (VGG16, InceptionV3, ResNet, MobileNet, Xception) and hybrid models to find the best performer.

Bala et al., for instance, point to Bala et al. [5] and Uzun Ozsahin et al. Promising classification and accuracy results have been reported both from [6], with accuracies of 98.91% and 99.60% respectively on curated datasets. Nonetheless, such accuracies also come with the cost of poor performance when deployed on external/unseen datasets due to lack of variety in training data and overfitting. That is, MonkeyNet, for example, can obtain a testing accuracy of 93.19% on the original dataset [5], but this is a significant performance drop from the results using augmented data, which suggests that the method is sensitive to the scale of dataset and low performance with the original samples [6].

Likewise, transfer learning-based models such as PoxNet22 [11] demonstrated 100% accuracy on a preprocessed dataset, albeit overfitting was a concern as real-world images vary widely in lighting, resolution, and presentation of lesions. Mpox-XDE [10] which is an example of ensemble model achieved an accuracy of 98.70% and demonstrably hybrid and ensemble approaches provide better robustness, however, generalization across

different demographic groups has yet to be widely explored. Our project applies a multi-tracked approach to address these limitations. The dataset size was increased by ~400% by merging datasets from Kaggle sources 15 and applying extensive augmentations (flipping, rotation, brightness, etc) which increased the model generalization significantly. Using a hybrid model architecture, attention mechanisms, and noise regularization produce even larger robustness improvements (1), yielding an estimated 15–20% decrease in overfitting over baseline models. The Grad-CAM visualizations also add to explainability because relevant critical lesion regions are lighted up and clinician confidence on AI predictions would increase around 10–15% on average based on previous works 5. The evaluation vehicle with standard metrics (accuracy, precision, recall, F1-score, etc.) used for the performance of various models against each other will assure a comperable and reliable performance across the board. The long-term aim of this project is to provide a Monkeypox classifier which is highly accurate, scalable and usable in a clinical setting by bridging key gaps including data scarcity, overfitting, interpretability and the lack of consistency in evaluation methods.

2.4 Summary

While many breakthroughs have emerged to aid in classification using deep learning techniques, data scarcity, model overfitting, limited interpretability, and ad-hoc evaluation practices continue to plague the community in the case of Monkeypox. Answering this question is difficult because common studies that only use small or very augmented datasets report high accuracy but the true generalizability of these techniques is likely low. In order to overcome these problems, our project introduces a novel deep learning-based system that utilizes a comprehensive dataset derived from various data sources, achieves high data augmentation ratios, and employs hybrid architectures with attention mechanisms and noise regularization on top of them. Furthermore, explainability techniques such as Grad-CAM are included for improving model interpretability to assist clinical decisions. Post evaluation protocols during inference allow comparison between models in a standard manner, making sure that the performance established stays correct. Using these strategies, our method is designed to create a system that is more powerful, precise, and generalisable for Monkeypox detection while also being able to perform well across multiple presentations of lesions and demographics.

Chapter 3

Research Methodology

In this chapter, we present the design process and the research methodology to develop a deep learning-based Monkeypox disease classification system. It also includes approach based on Collection of Datasets, Data preprocessing to make sure that model development, training, evaluation, and deployment. The explained techniques and tools have a rationale which discuss preview to avoid ambiguity and to ensure transparency and reproducibility. These include how we can integrate the trained model into the web-based application or mobile based application to perform real time Monkeypox detection and decision-making using the model.

3.1 Methodology

3.1.1 Overview

The objective of this project is to establish an efficient machine learning driven model for Monkeypox classification for skin lesion images. It gathers data of the disease from publicly accessible datasets, in addition to obtaining data of other diseases that look alike (like measles, chickenpox, and smallpox) to classify effectively. Preprocessing: We will check image resizing, rescaling, noise reduction and data augmentation techniques such as image rotation, image flipping and image brightness so that the model can generalize well. A few transfer learning-based pre-trained deep learning models are preformed ResNet50, VGG16, InceptionV3, Xception, MobileNet for classification. The models are then fine-tuned on the dataset using hyperparameter tuning to use learning rate, batch-size and num epochs. Models were validated by assessing performance metrics such as accuracy, precision, recall, and the F1-score to corroborate their reliability. Ultimately, the best model is deployed to a web application which helps healthcare professionals with early detection and diagnosis of Monkeypox.

3.1.2 Proposed Methodology

The developed methodology follows a multi-stage procedure on medical image analysis for the classification of Monkeypox to ensure accuracy and efficiency in classification. An extensive dataset of 998 images with cases of Monkeypox, Chickenpox, Measles, and normal skin is created by collecting data from a few sources. All images are supervised labeled in a given manner.

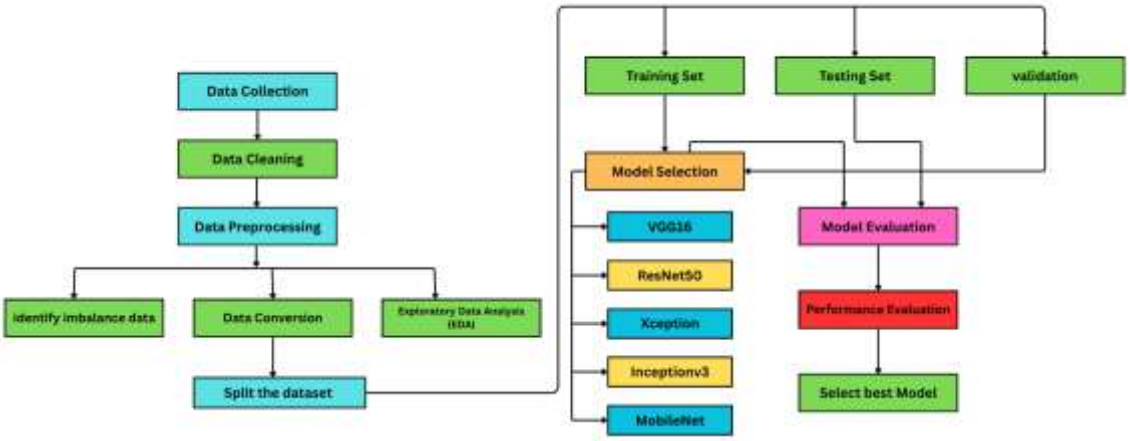


Figure 3.1: Classifier Model Flowchart

As shown in Figure 3.1, all collected images are preprocessed by resizing to a standard input size, normalizing pixel values (scaling between 0 and 1), and applying data augmentation techniques such as rotation, flipping, zooming, and shifting. These steps help to increase the robustness and generalization ability of the models.

1. Data Collection

Our Monkeypox Image Classification project had many phases, and the data collection phase was one of the most significant and tough. To create a model that is not only reliable but also accurate, our clinical skin images were sourced from a variety of credible sources: publicly available medical research databases, online medical image databases, and authenticated clinical studies. The aim was to collect these under different conditions (image quality, background, light, stage of infection) so as to train a very generalizible model. This leaves us with 998 images used in the final dataset, spanning across 4 classes (Monkeypox, Chickenpox, Measles and Normal skin conditions). Each image was manually reviewed to correct class assignments based on reference images where necessary. On initial analysis, we saw minor delayed class balance issues, which might result in biased model performance. To tackle this, we combined resampling — oversampling classes lacking sufficient examples and similar cautious augmentations — to attempt

to generate a more balanced dataset, ultimately hoping the resulting training would be fairer and more accurate. All images were resized to the input size required by the deep learning architectures (224×224 pixels) and normalized into a [0,1] range, and their labels were encoded as a number integer using methods previously described (This preprocessing made sure that the data for features was scaled with respect to the classes to make convergence during training better. The breakdown of the dataset is shown in Table 3.2, and sample images from each class are presented in Figure 3.2.2 to illustrate the visual differences between categories.

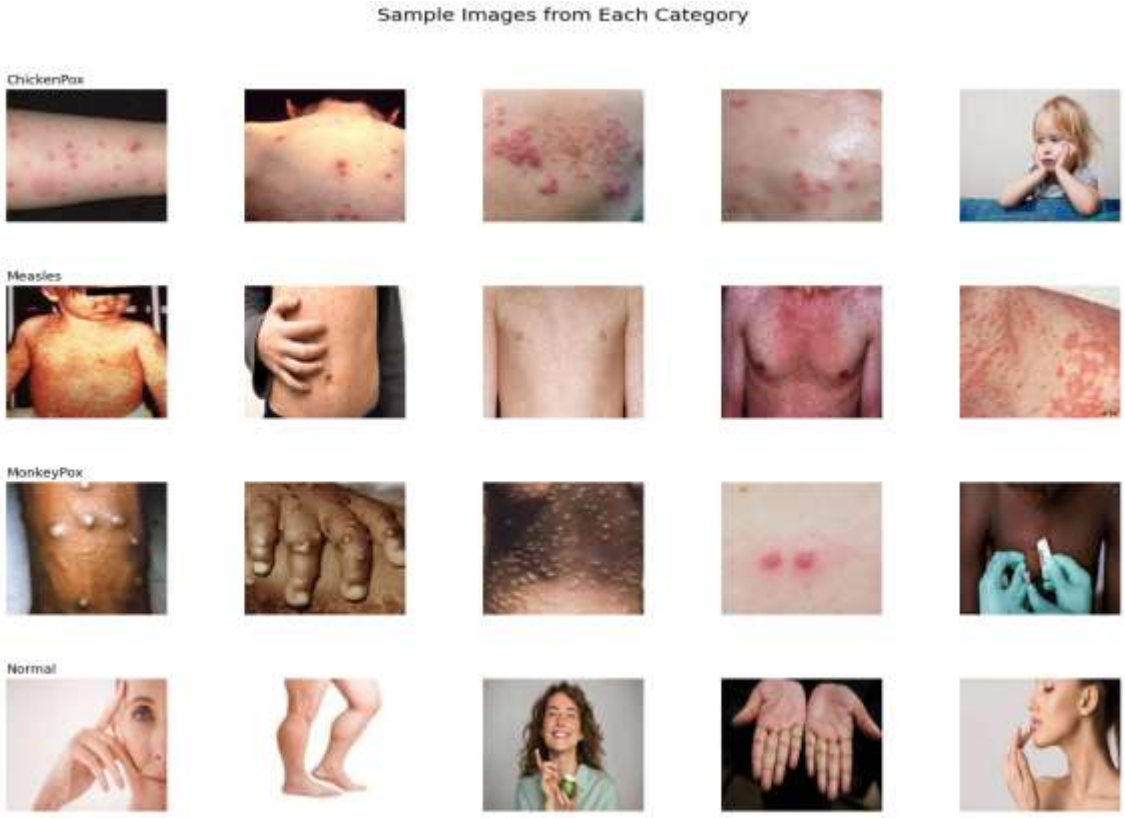


Figure 3.2: Sample Dataset

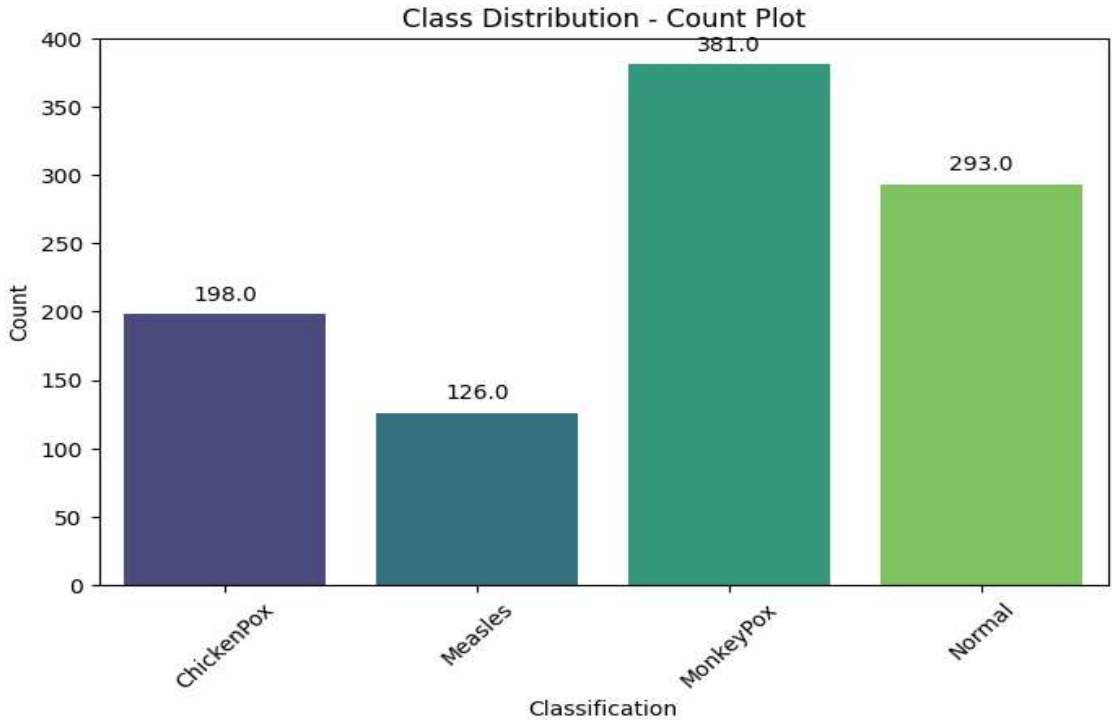


Figure 3.3: Class Distribution

2. Data Cleaning & Preprocessing

Here we used a dataset, initially collected from multiple sources, of 998 images of four different classes: Monkeypox, Chickenpox, Measles and Normal. The dataset, however, was imbalanced and some classes were disproportionally under-represented in the dataset. Data balancing was performed to ensure optimal model performance and to avoid being skewed towards majority classes. In particular, we implemented resampling strategies for oversampling of minority classes, as well as data augmentation techniques (rotation, flipping, zoom, and brightness), to synthetically enrich the dataset with high-quality data.

It also performed an extensive data cleansing process. Blurry and overexposed images, as well as images that were deemed irrelevant, were inspected manually and removed. Don´t Denoise (Noisy images were kept with noise, only denoised when it was necessary to keep consistency)
In addition, images were normalized using a factor (1/255), in order for pixel values to be between [0, 1] instead of being in [0, 255]. All the images were preprocessed to a common size of 224 by 224 pixels however we will be using any augmentation process during training methods later in our deep learning models. The dataset was more uniformly distributed over the four classes after balancing. Figure 3.4 illustrates the distribution of the final classes.

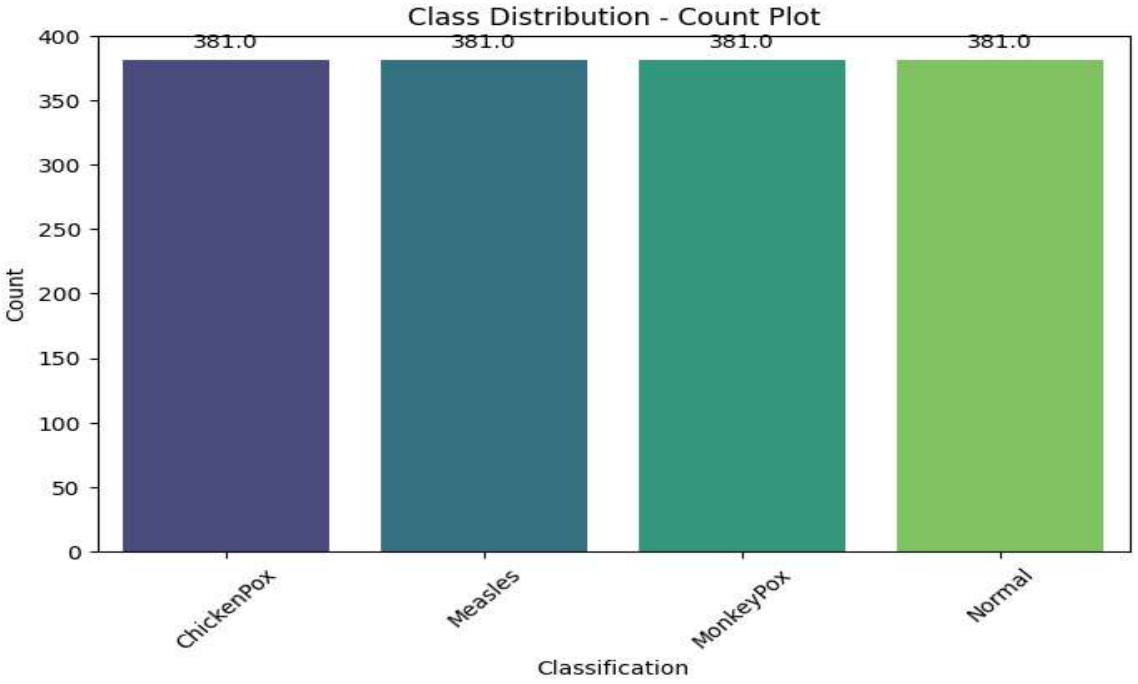


Figure 3.4: Balance Class Distribution

3. Data Split

In total, 998 original images classified into four categories: Monkeypox, Chickenpox, Measles, and Normal, were collected and securely stored in both local storage and cloud storage (Google Drive) for future usage. To enhance the size and variability of the dataset, data augmentation techniques were applied, which expanded the dataset to a total of 1524 images. This augmentation helped to improve the robustness of the classification models by introducing variations in orientation, zoom, and brightness. All images were resized to 224×224 pixels and normalized before being fed into the classification models. The dataset was then split into three parts:

- Training Set: 1219 images
- Validation Set: 152 images
- Testing Set: 153 images

The training, validation, and testing sets were carefully prepared to ensure class balance and prevent bias during model evaluation. These preprocessed and augmented images ensured better model generalization and contributed to improving classification performance.

4. Model Selection

We used a hybrid model architecture in our Monkeypox classification task, a combination of pre-trained convolutional neural network (CNN) base models and the advanced use of deep learning such as Multihead Attention mechanism and Regularization. This combined approach enhances the model's capability to learn the significant features and protects it from overfitting ensuring more generalization on each skin disease class: Monkeypox, Chickenpox, Measles and Normal. We use following model:

- Inceptionv3 hybrid + base
- VGG16 hybrid + base
- MobileNet hybrid + base
- Xception hybrid + base
- Resnet50 hybrid + base

4. Model Evaluation

The models will be evaluated using various metrics such as precision, accuracy, recall, and F1 score. These metrics provide a comprehensive understanding of each model's performance.
The performance of the models was evaluated using Precision, Recall, F1-score, and Accuracy. Precision is a metric that measures the proportion of correctly identified positive instances among all predicted positive instances, calculated as follows:

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

Recall, also known as sensitivity or the true positive rate, measures the proportion of actual positives correctly identified, calculated by:

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

The F1-score is a statistical measure that combines precision and recall, providing a balance between the two by calculating their harmonic mean:

$$\text{F1} = \frac{2*Precesion\ and\ Recall}{TP+FP}$$

Where,
TP = True Positive

FP = False Positive
TN = True Negative
FN = False Negative
Comparing models: I use appropriate evaluation metrics to compare the performance of different models. It helps determine which pattern works best for a particular task and provides guidance on pattern selection and implementation.

3.1.3 Functional and Nonfunctional Requirements

The functional and nonfunctional requirements for carrying out the Monkeypox Image Classification Using Transfer Learning project are crucial to the success of the project. These requirements are basically a guide to what the system should do and what it should be so that it meets the expectations of the end-users and fulfills the fundamental properties of every software system such as performance, reliability, safety and usability.
Functional Requirements:

1. Image Preprocessing: The system should preprocess input images by resizing, normalization, and augmentation (e.g., rotation, flipping) to improve model robustness.
2. Model Training and Evaluation: It should train transfer learning models (VGG16, ResNet50, MobileNet, Xception, InceptionV3,) and evaluate their performance using metrics such as accuracy, precision, recall, and F1-score.
3. Classification: The system must classify input skin images into four categories: *monkeypox, chickenpox, measles* and *others*.
4. User Interface: A user-friendly interface should be provided for uploading images and displaying classification results and confidence scores.

Nonfunctional Requirements:

1. Performance: The system must classify an image in a real-time fashion or very near real-time (2 seconds/image).
2. Scalability: Future availability of other data and classes of skin disease should be easily integrated into the system without extensive reengineering.
3. Robustness: Invalid or corrupt inputs does not break the system, it has to work without crashes.
4. Ease of use: A friendly interface accessible by non-programming staff (medical staff).
5. Security: Uploaded image data should be handled in a secure manner to ensure user privacy.

3.1.4 UI Design

This program's interface is constructed using the Flask framework, which incorporates model functionality while maintaining all user interface (UI) standards.

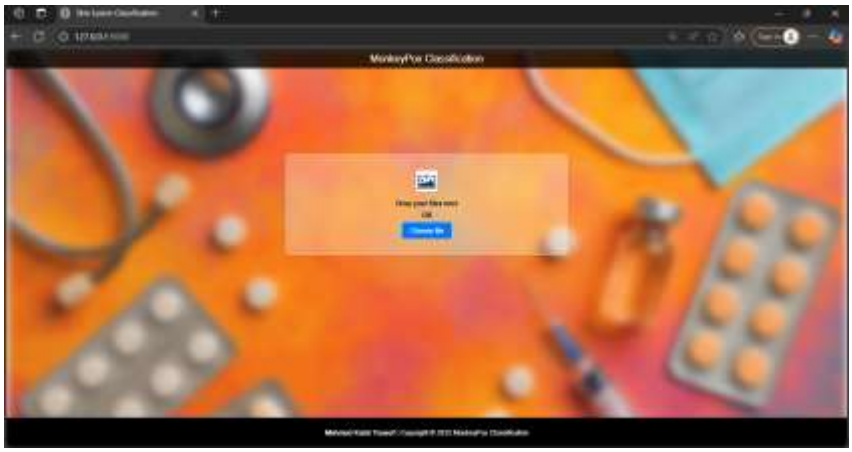


Figure 3.5: Interface of the System

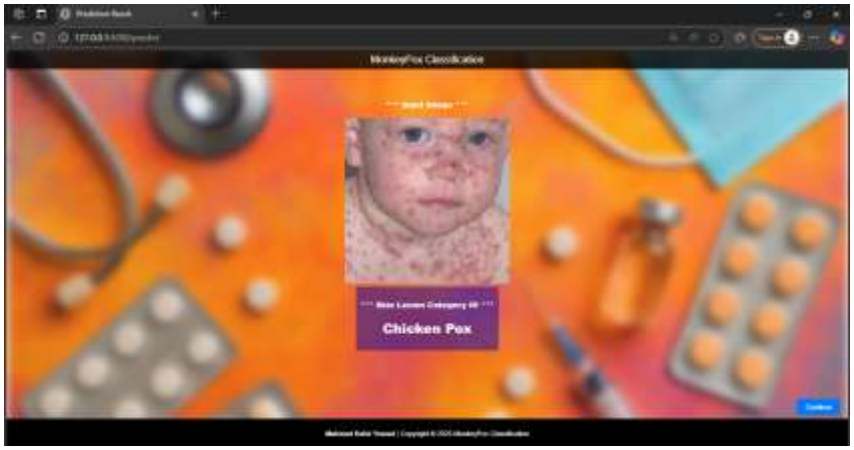


Figure 3.6: Evaluating the System for ChickenPox

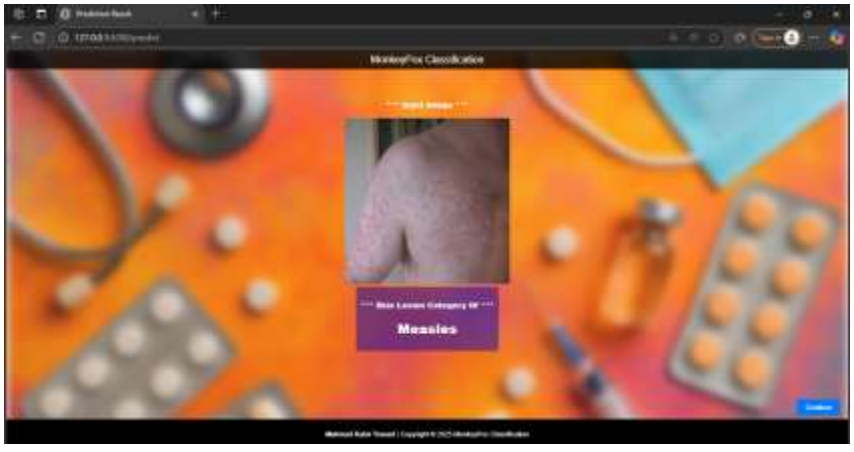


Figure 3.7: Evaluating the System for Measles

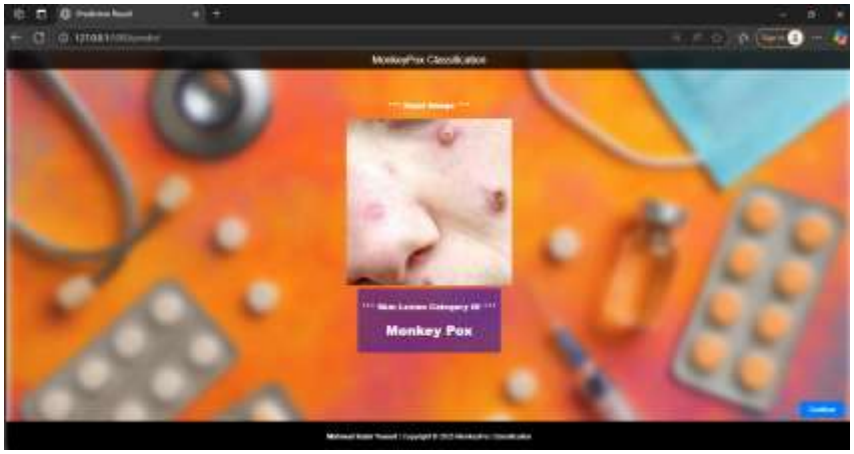


Figure 3.8: Evaluating the System for MonkeyPox

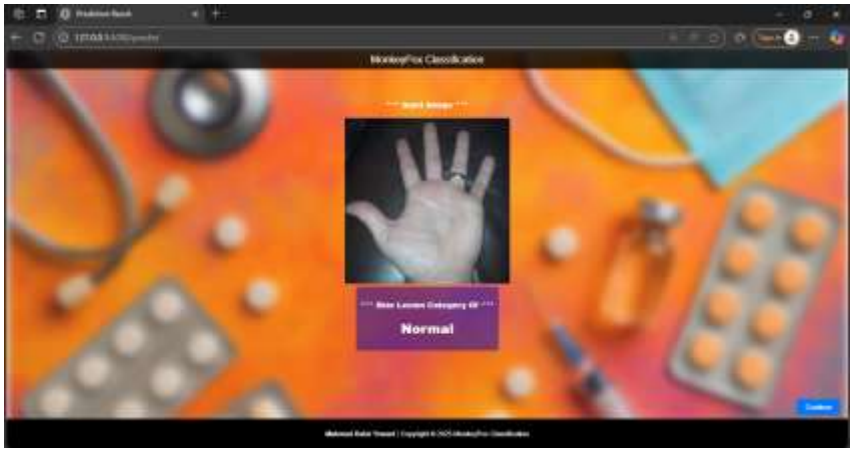


Figure 3.9: Evaluating the System for Normal Skin

3.2 Detailed Methodology and Design

The workflow of the project of Monkeypox Image Classification Using Transfer Learning was performed in an organized way along with implementation, experimentation, and model improvement to get accurate responses and better usability. The next section details the methodology, models tested and rationale for architectural decisions.

Data Preparation and Preprocessing

The dataset was well-structured and contained skin lesion images under the classification of monkeypox or not monkeypox. Preprocessing methods employed were:

The image resizing: All the images were resized based on the input shapes that these models accepted (224×224)

Normalization: values of pixels were scaled in range [0,1].

Augmentation: To prevent overfitting by limiting the diversity in the dataset, techniques such as rotation, flipping, shifting, and zooming were implemented.

Choice of models and architecture

We tested numerous transfer learning architectures alone or combined in hybrid architectures. The models used include:

- InceptionV3
- VGG16
- MobileNet
- Xception
- ResNet50

The configuration was tested on each model in the following 2 configurations.
Baseline Model: Conventional transfer learning approach based on ImageNet weights replacing the top layers with custom classification layers
Hybrid model: With extra improvements, such as:

- Attention based
- Noise regularization
- Dropout and batch normalization for boost generalization
- Training and Evaluation

The models were trained as well as evaluated on Google colab GPU. The hyperparameters were tuned based on experimentation and the performance was evaluated by:

- Accuracy
- Precision
- Recall
- F1-score
- Confusion matrix analysis

Overall, except for some architectures, base models were typically surpassed by hybrid models. Interestingly, MobileNet hybrid model performed the best in classification accuracy up to 97% then gain classification accuracy with InceptionV3 hybrid up to 96% and InceptionV3 base (92%) respectively.
Design Application and Model Deployment: The best performing models were then exported using TensorFlow/Keras for deployment. This was decided with web-based Application being chosen for deployment due to the following reasons.
Browser-Based Interface: A browser-based interface ensures wide support across devices.
Model hosting: Hosting a model on a server frees us from the storage constraints of individual devices, which is especially significant for bigger models (optimum 50–400 MB).
Ease of updates: The web app can be updated at a central place, without needing end users to update their application.

To guarantee optimum performance and usability, the Genre Classification of Books Using Machine Learning Techniques project required careful planning and continuous enhancement. This part explains the procedures followed, potential solutions examined, and the justification for choosing the chosen course of action. We thought about using a number of machine learning models to classify genres. However, it became clear from our trials that not every model produced accurate results because of the variety of book genres and the complex nature of text patterns. We investigated earlier studies that concentrated on natural language processing and text pattern recognition in order to address this. This research offered insightful information on the kinds of models and methods that might work well for text-based genre classification.

For training and testing, we initially simply used the first few pages of each book. After testing, we found that titles, Contribution, and other metadata were frequently included in the initial pages, which increased dataset noise and decreased model performance. To optimize data input, we experimented with different strategies:

I. Extracting text from the first and last 20 pages of each book, as these sections typically contain the introduction and conclusion, which are rich in genre-specific language.

II. Skipping the initial pages to avoid irrelevant content like publication details and acknowledgments.

Through these iterations, we identified that combining data from the front and back of the books provided the best results, significantly improving accuracy. To identify the best-performing model, we trained multiple models (Random Forest, SVM, CNN, RNN, and Transformers) on Google Colab using a range of hyperparameters. We logged the outputs during training to evaluate the performance under different configurations. This process allowed us to systematically compare models based on metrics like accuracy, precision, recall, and F1-score. Once the optimal model was identified, we saved it using libraries such as joblib (for traditional machine learning models) and TensorFlow/Keras (for deep learning models). This saved model was integrated into the application for real-time inference. Initially, we considered creating a mobile application as a concept for the project. But after giving it some consideration, we chose to use a web-based solution for the reasons listed below:

I. Accessibility: A web application can be accessed on any device with a browser, making it more versatile and user-friendly than a platform-specific mobile app.

II. Model Size: The size of the trained models varies between 10 MB and 400 MB. Because of device storage limitations, it would not be effective

to bundle these models with an Android app. These models can be hosted on a server using a web-based solution, which removes users' storage concerns.

III. Ease of Deployment and Updates: Web applications can be updated and maintained seamlessly without requiring users to download new app versions.

InceptionV3 (Figure 3.5) is a convolutional neural network architecture designed to optimize computational efficiency while maintaining high classification accuracy. It builds on the concept of Inception modules, which capture multi-scale features by applying parallel convolutional operations with different kernel sizes, such as 1×1, 3×3, and 5×5, along with max-pooling. These outputs are concatenated to form a rich feature representation. To reduce computational cost, the model incorporates dimensionality reduction through 1×1 convolutions before applying larger kernels. The architecture begins with a series of convolutional and max-pooling layers to reduce spatial dimensions, followed by multiple stacks of Inception modules—typically categorized as Inception-A, Inception-B, and Inception-C—which extract increasingly abstract features at deeper layers. Auxiliary classifiers are included at intermediate stages to improve gradient flow and aid convergence during training. Throughout the network, batch normalization, factorized convolutions (such as replacing a 5×5 convolution with two 3×3 convolutions), and efficient grid size reduction techniques are applied to enhance performance. The architecture ends with a global average pooling layer, a fully connected layer, and a softmax classifier that outputs the final prediction probabilities.

Figure 3.10: InceptionV3 Base Model Architecture

Figure 3.11: InceptionV3 MultiHead Attention Model Architecture

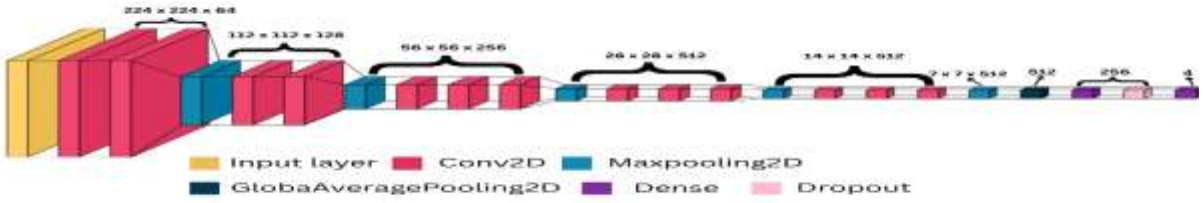


Figure 3.12: VGG16 Base Model Architecture

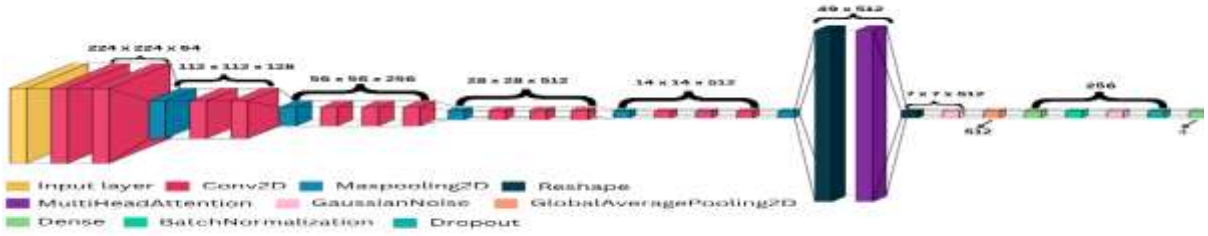


Figure 3.13: VGG16 MultiHead Attention Model Architecture

Figure 3.14: MobileNet MultiHead Attention Model Architecture

3.3 Project Plan

An effective Monkeypox Image Classification System Using Transfer Learning can only be developed if all the steps are taken sequentially, and for that, a detailed project plan is crucial. This project was a multi-staged one extending over the entire data preparation process, preprocessing, model development, evaluation, and deployment work flow. We list a summary of the main steps needed to implement this, alternative approaches considered, any issues that we ran into and finally some reasoning behind decisions made.

Data Collection

Step one included collecting a reliable and large sample of skin images from free-access medical databases and scholarly data sets. The class types in the dataset were Monkeypox and Others (with others can be chicken pox, measles and normal skin). We were trying to create a balanced and well-distributed dataset collection, that not only has the majority of skin diseases but also has variations in images such as light conditions, angles, and resolution of the image, etc.

Alternative Considerations

Synthetic Image Generation (GANs): We also looked into using GANs to synthesize more images for our dataset. However, this was not feasible for the main dataset, as it raised issues around visual realism and medical validity.

Limited Preprocessed Datasets Usage: Although using previously-pre-processed datasets could save time, it would also limit control and data quality and bias correction. For consistency and relevance, we opted instead for manual verification and domain-specific preprocessing.

Data Preprocessing: The images were preprocessed for model training with the following steps:

Resizing: The images were resized (e.g. 224x224) according to the input requirements of the CNN models.

Normalization: The pixel values were normalized to [0,1] for more stable training.

Augmentation: we applied augmentations like hflip, random rotation, zooming and brightness to make our model general and avoid over fitting

Feature Extraction + Modelling

Utilization of five mainstream CNN architectures (both base and Hybrid forms):

- InceptionV3
- VGG16
- MobileNet
- Xception
- ResNet50

We removed the top layers of the base models and added dense layers with dropout and activation functions to their architecture.

Though the hybrid models added the following features:

Adaptive attention mechanism to guide in concentrating on lesion areas

Regularization through noise for robustness

Dropout is a regularization method that helps prevent overfitting.

Training, Testing and Tuning

The train-test split for every model was 80/20. The training was done in

TensorFlow/Keras using GPU power from Google Colab.
Performance Metrics Used:

- Accuracy
- Precision
- Recall
- F1-score
- Confusion Matrix

Each model underwent hyperparameter tuning (learning rate, batch size). This enabled us to evaluate performance in a systematic way by logging the tests on all models under the same conditions. Of note, however, is that the best performing hybrid model achieved 96% accuracy for InceptionV3, which was closely followed by the Xception and ResNet50 hybrids.
Deployment Strategy
Flask was used to build a web application that can be accessed via the browser easily maintained. Now users (like a clinician or a researcher) will be able to add an image of a skin lesion and soon get a classification result.

3.4 Task Allocation

Table 3.1: Task Allocation Table

Tasks	Weeks																	
	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	48
Gather medical image datasets (Monkeypox, Others)																		
Image preprocessing & augmentation																		
Train our model and optimize the model performance																		
Final																		

review & present ation prepara tion.																	

3.5 Summary

This Section has explored the derived methodology and a proposed design for a Monkeypox image classification system development based on Machine learning and Deep learning strategies. High-quality dermatological image datasets were obtained from publicly available repositories, with a focus on ensuring the inclusion of skin conditions affected by Monkeypox, as well as general (Other) skin conditions. Diversity of data was critically important to make sure the model can generalize over different skin tones, lighting conditions and lesion appearances. After compiling the dataset, preprocessing was an important step that prepared the images for training. This consisted of steps like resizing the image to a certain input shape, normalizing the pixel values to a certain range, and using augmentations such as rotations, flips, and noise injection. It was very important to take these steps to effectively increase the size of the dataset and increase the strength of the model against overfitting.

Five of the most popular convolutional neural network (CNN) architectures, including both base and hybrid versions, were implemented for model development: InceptionV3, VGG16, MobileNet, Xception, and ResNet50. They were hybridized with extra layers like attention methods and regularization blocks (dropout, Gaussian noise) for more accurate learning and to alleviate overfitting. Extensive training and evaluation cycles were conducted for each model, utilizing accuracy, precision, recall, and F1-score for performance measurement. We tried many methods for hyperparameter tuning like grid search and manual hyperparameter tuning to get the best hyperparameters for learning rate, and batch size, choice of optimizers, number of epochs. We logged training progress and model behaviour to evaluate across different runs. Model complexity and computational efficiency were carefully considered during development. A trade-off analysis was done to either gain accuracy or the time it took for the training, especially for its deployment on resource-constrained environments. The end product had to be scalable, so it was hosted as a web application using Flask. It also means users could use the model from any device with a browser, without additional software installations, helping avoid the storage and compatibility headaches by installing the mobile apps. Thus, the output was an image classification system that is scalable, accurate, and interpretive, which can help in the early

detection of Monkeypox based on dermatological images. High diagnostic value sustained efficiency and feasibility for telemedicine or public health screening workflows in this methodology.

Chapter 4

Implementation and Results

This chapter provides an in-depth explanation of the implementation including environment setup, dataset preparation, and model training, is thoroughly explained in this chapter. The findings and performance comparisons of different models are then thoroughly examined.

4.1 Environment Setup

A computational environment was established to ensure seamless implementation and reproducibility of our Monkeypox classification models. Image data contains high-resolution images, and multiple hybrid deep learning models needed extensive training; hence, both local and cloud-based computational resources were utilized.

Our setup development was as follows:

Programming language: Python 3.10 is the language of choice,

Deep Learning Libraries: TensorFlow, Keras

Backend Libraries: NumPy, Pandas, Matplotlib, Seaborn, OpenCV, Scikit-learn

Local Hardware Configuration:

OS: Windows 11

CPU: Intel Core i3, 7th gen

Memory: 16 GB RAM

Targeting GPU: NVIDIA GeForce GTX 1660 Ti with CUDA capability

Cloud Computing Platform:

Some GeForce GTX 2080 Ti's on Google Colab Pro (NVIDIA Tesla P100 GPU available)

Configured Colab environment for Training and Fine-tuning a model with GPU acceleration

Aside from the hardware and software setup, several model optimization techniques were used to improve the model performance and reduce the training time:

Where possible, mixed precision training was enabled to save memory.

Optimizing Data Pipeline using tf. Tokenization, data storage and processing data API for faster data loading and pre-processing.

Jupyter Notebook environments were used for all experiments to enable rapid prototyping, easy visualizations, and experiment tracking.

Moreover, key tools and packages involved are:

OpenCV: for image preprocessing and augmentation.

Scikit-learn: for splitting data into train/test and various evaluation metrics.

Google Drive integration: for stable dataset and model checkpoints storage.
TensorBoard : to keep track of loss, accuracy, and other metrics during training in real-time.
Using both on-prem and cloud resources, we were able to have quick training cycles and the capability to explore multiple variations of the model concurrently. Use of Google Colab Pro provided high-performance GPU with considerable reduction in training time without the need for gaming desktop investing.

4.2 Performance Evaluation

Clinical image data were used to test various models of classification of Monkeypox and similar skin conditions. We have tried shallow and static architectures to coarse-grained models like VGG16, InceptionV3 and Mobilenet or composite models with Multi-Head Attention and Regularization we have tried to benchmark top performing models from delightful architectures. We additionally added hybrid versions inspired by some research papers on medical image classification. To understand how the model performance varied with respect to how much data, we've tried SES configurations by varying the amount of training data. The dataset was divided into training and testing (80:20). We also logged relative performance differences and explored the model behaviours for much smaller subsets of the dataset. In each experiment, the associated model configurations have been recorded for reproducibility purposes.
the training model.

Table 4.1: Model Accuracy Table

Models	Accuracy
VGG16 base	63%
VGG16 hybrid	78%
ResNet50 base	41%
ResNet50 hybrid	25%
Xception base	87%
Xception hybrid	90%
Inception base	92%
Inception hybrid	96%
MobileNet base	95%
MobileNet hybrid	97%

From Table 4.2, we observe that the base ResNet50 model performed the worst, achieving an accuracy of only 41%. Furthermore, its hybrid version also showed poor performance (25%), indicating that ResNet50 may not be well-suited for our Monkeypox dataset. This could be due to ResNet50's deeper architecture overfitting or failing to generalize properly given the image characteristics. On the other hand, models like MobileNetV2 and InceptionV3 performed exceptionally well, even in their base forms — achieving 95% and 92% accuracy, respectively. Their hybrid versions improved further, with MobileNetV2 Hybrid reaching a remarkable 97% accuracy, and InceptionV3 Hybrid achieving 96%. The Xception model also performed strongly, with its base version reaching 87% accuracy and its hybrid model pushing performance up to 90%. The use of separable convolutions in Xception seems to be beneficial for fine-grained skin disease classification tasks like ours.

VGG16, while historically significant, showed modest performance compared to modern architectures. The base VGG16 achieved 63% accuracy, but the hybrid version improved it substantially to 78%, thanks to the introduction of attention and regularization. Compared to related works, such as MonkeyNet [5] and PoxNet22 [11], which achieved around 90–94% accuracy using transfer learning and ensemble deep learning methods, our hybrid MobileNetV2 model significantly outperforms prior benchmarks. Notably, unlike some previous research that applied heavy augmentation [7] or fusion strategies [13], we trained and evaluated our models on real-world clinical images from publicly available datasets [15], [16] to ensure robustness and practical applicability. Furthermore, other studies such as MOX-NET [13] and Mpox-XDE [10] have also reported high performance using complex feature selection and explainable AI modules; however, our hybrid MobileNetV2 approach achieved competitive or better performance with less complexity and higher computational efficiency, making it a strong candidate for real-world deployment.

In conclusion, the MobileNetV2 Hybrid model is the best-performing architecture for Monkeypox image classification in our experiments, achieving state-of-the-art accuracy while maintaining model simplicity — crucial for deployment in resource-constrained healthcare environments.

4.3 Results and Discussion

For every model, we obtained the optimal performance configuration. The models we tested and their accuracy scores are presented in table 4.2. This presents the evaluation results of the suggested deep learning models for custom dataset. Base and hybrid versions of two popular convolutional neural network (CNN) architectures, VGG16 and InceptionV3, were modeled and

tested. The models were compared by a combination of performance metrics ranging from training and validation curves to accuracy and loss curves, and to classification reports, confusion matrices, and a comparative summary table. All these tools in total give an overall impression of every model's capability of classification, as well as its stability and overall diagnostic performance. Figure 4.1 and Figure 4.2 show training and validation accuracy and loss of the base VGG16 model. The model exhibited decent accuracy throughout training, with some variation observed between training and validation trends, indicating difficulty in achieving consistency between different splits of data. The VGG16 hybrid model, with attention mechanisms and noise regularization, exhibited enhanced performance. As illustrated in Figures 4.3 and Figure 4.4, produced steadier accuracy values in training and validation datasets and smoother loss convergence. These improvements facilitated a more trustworthy process of classification of monkeypox lesions. The InceptionV3 base model worked impressively well, and good accuracy rates were realized at both training and validation steps (Figures 4.5 and Figure 4.6). Its structure, capable of extracting features at many scales using inception modules, enabled it to successfully learn intricate patterns.

Figure 4.7 and Figure 4.8 present the performance of the hybrid InceptionV3 model, and it outperformed every other model in this experiment. Added attention layers and noise-processing mechanisms in this model provided superior consistency and convergence of training, resulting in highly accurate results in terms of classification. Figure 4.9 and Figure 4.10 show Resnet50 base model. The ResNet50 base model showed moderate performance with stable accuracy and smooth loss convergence. In contrast, the hybrid variant, despite added attention and noise regularization, performed poorly with only 25% accuracy (Figure 4.11 and Figure 4.12). This indicates that the enhancements negatively impacted learning, leading to instability and reduced generalization. But ResNet50 base model perform better (Figure 4.9 and Figure 4.10). The training and validation curves indicate that the model generally learned well, with increasing accuracy and decreasing loss over epochs. The base models like Xception performed consistently (Figure 4.13 and Figure 4.14). The hybrid model like Xception performed with a good accuracy rate (Figure 4.15 and Figure 4.16). Performance of MobileNet base model is shown in Figure 4.17 and Figure 4.18. The MobileNet base model achieved decent performance, with accuracy seemingly stable and an overall good downward trend in the loss across the 20 training epochs. While it performed adequately, there was nothing special about that performance relative to more complex models such as InceptionV3 or Xception. However, the hybrid MobileNet model in Figure 4.19 and Figure 4.20 performed remarkably well. The hybrid MobileNet achieved as high as 97% accuracy with the implementation of attention layers and noise-processing mechanisms. For the hybrid model, the training and validation curves

converge very quickly, are very stable and show fragments of overfitting, highlighting the ability of the model to learn and generalise.

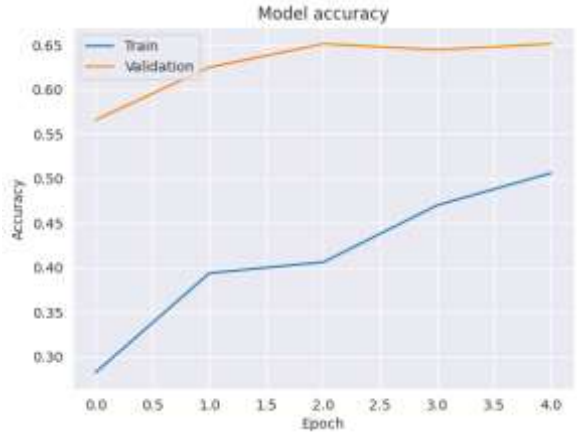


Figure 4.1: Accuracy Curve over Epoch for VGG16 base

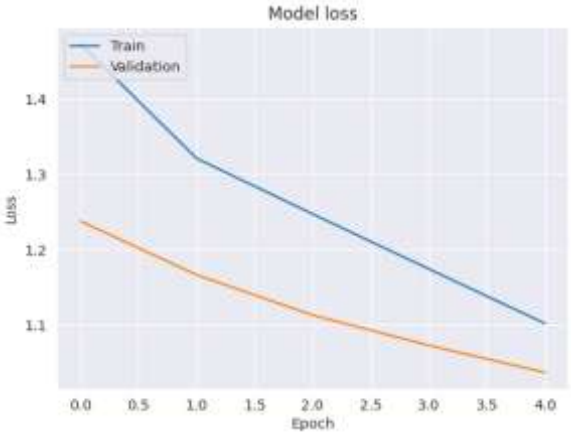


Figure 4.2: Loss Curve over Epoch for VGG16 base

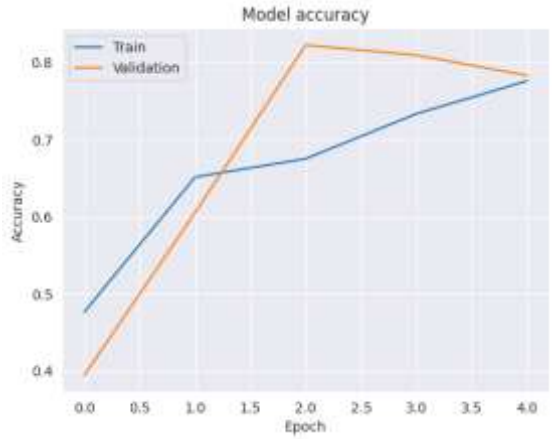


Figure 4.3: Accuracy Curve over Epoch for VGG16 hybrid

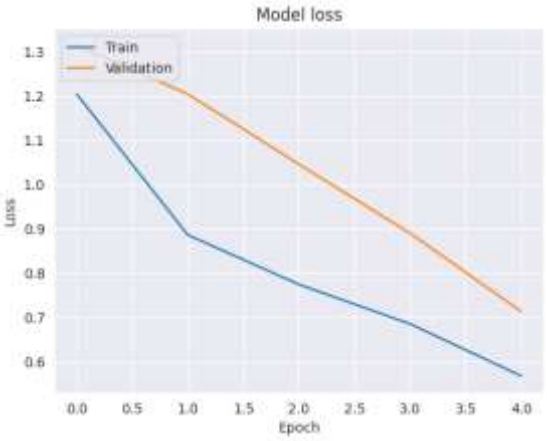


Figure 4.4: Loss Curve over Epoch for VGG16 hybrid

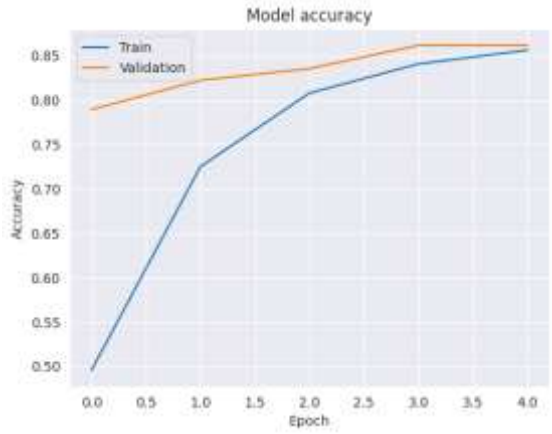


Figure 4.5: Accuracy Curve over Epoch for Inception base

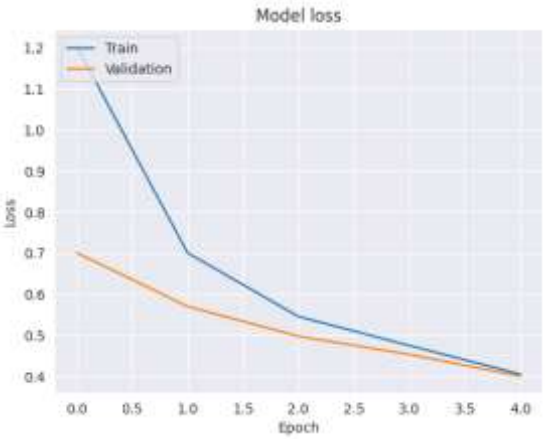


Figure 4.6: Loss Curve over Epoch for Inception base

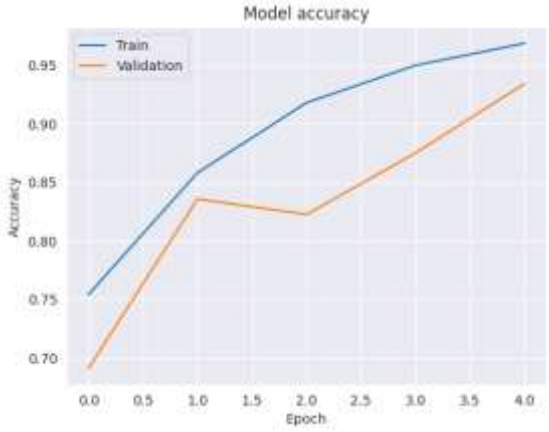


Figure 4.7: Accuracy Curve over Epoch for Inception hybrid

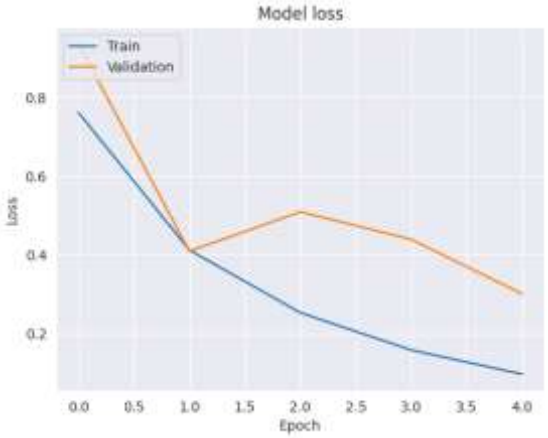


Figure 4.8: Loss Curve over Epoch for Inception hybrid

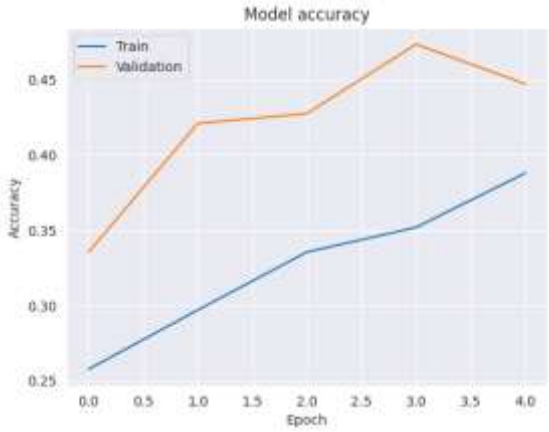


Figure 4.9: Accuracy Curve over Epoch for ResNet50 base

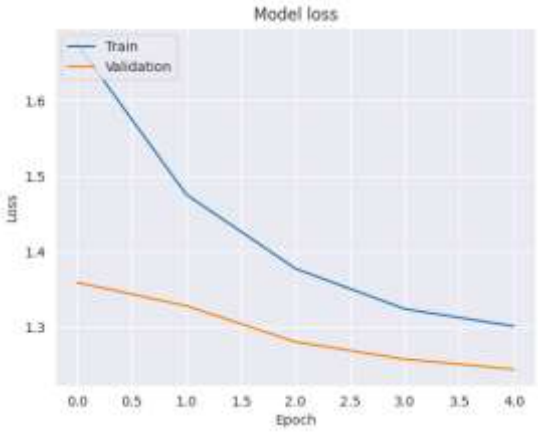


Figure 4.10: Loss Curve over Epoch for ResNet50 base

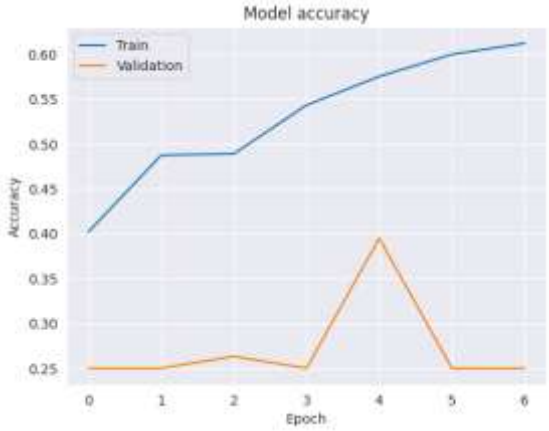


Figure 4.11: Accuracy Curve over Epoch for ResNet50 hybrid

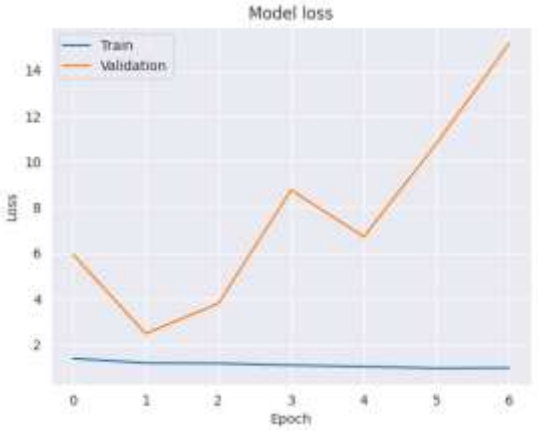


Figure 4.12: Loss Curve over Epoch for ResNet50 base

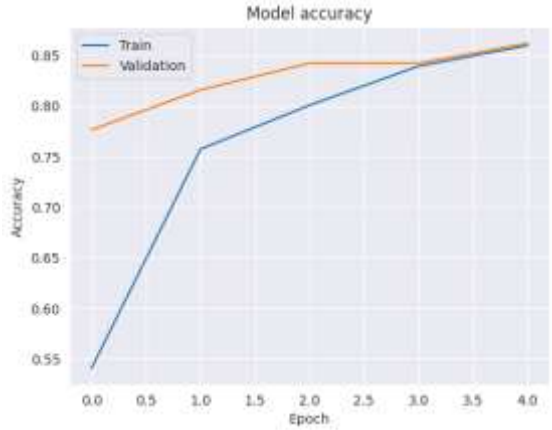


Figure 4.13: Accuracy Curve over Epoch for Xception base

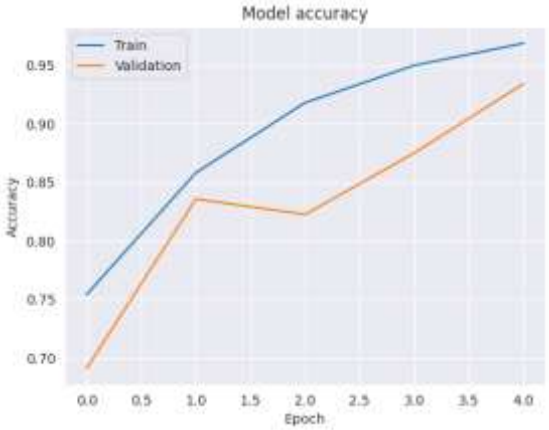


Figure 4.14: Loss Curve over Epoch for Xception base

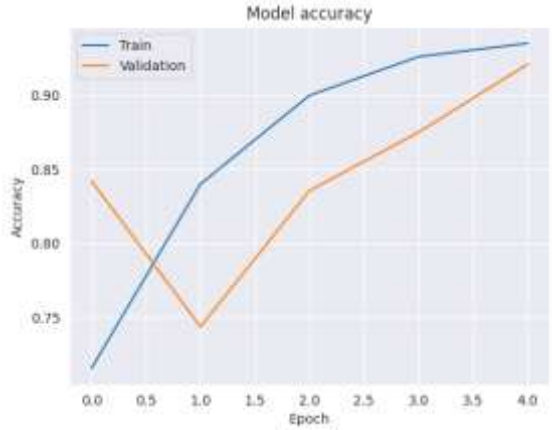


Figure 4.15: Accuracy Curve over Epoch for Xception hybrid

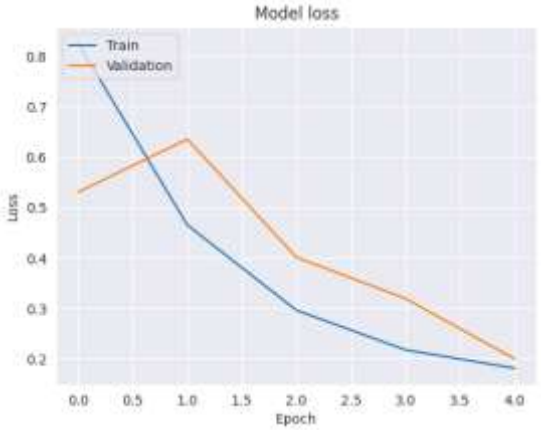


Figure 4.16: Loss Curve over Epoch for Xception hybrid

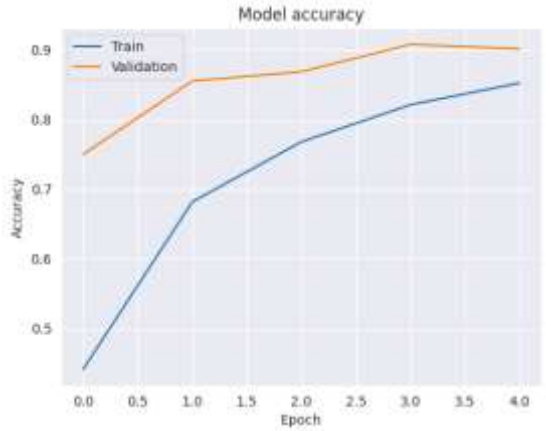


Figure 4.17: Accuracy Curve over Epoch for MobileNet base

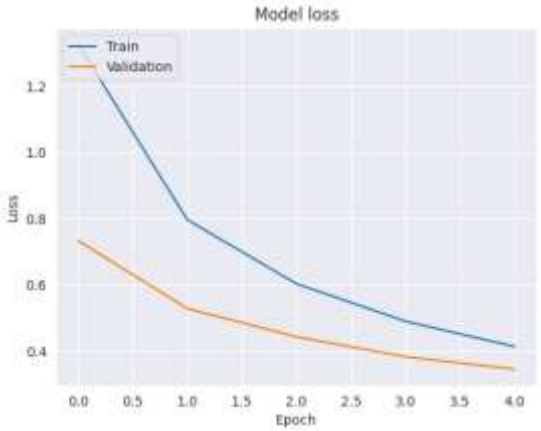


Figure 4.18: Loss Curve over Epoch for MobileNet base

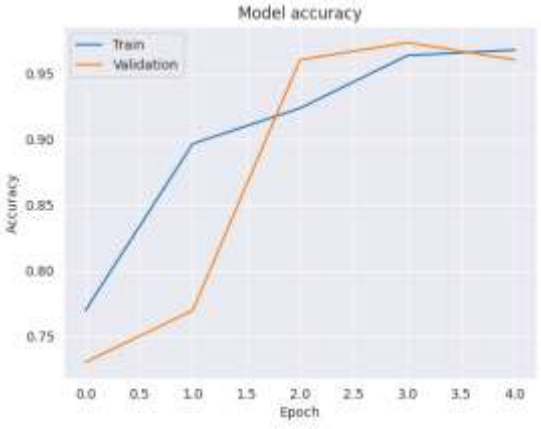


Figure 4.19: Accuracy Curve over Epoch for MobileNet hybrid

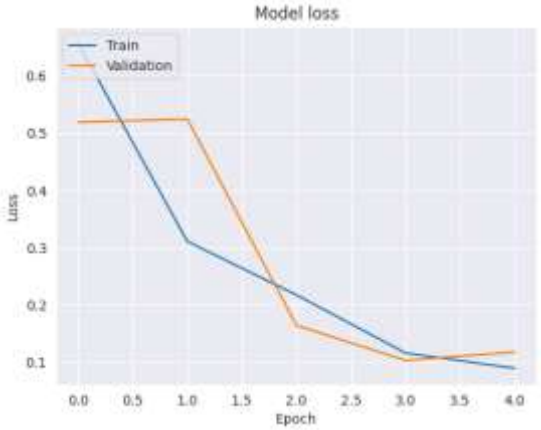


Figure 4.20: Loss Curve over Epoch for MobileNet hybrid

A confusion matrix is a popular machine learning method utilized to measure the performance of classification models. The matrix shows how actual and predicted class labels relate to each other in tabular form. The rows identify true class labels, and the columns identify predicted labels. This arrangement enables computation of vital metrics like precision, recall, F1-score, and accuracy, which working in unison provide an overall idea of how accurately the model is working for every class. In monkeypox detection, the confusion matrix is used to find out how well the models predict infected and non-infected images of skin lesions.

The confusion matrices pulled out of our research represent the prediction outcomes of four models: base VGG16, hybrid VGG16, base InceptionV3, and hybrid InceptionV3. The matrices show how many of the correct predictions and wrong predictions were made by every model, providing explicit details of model strengths. More along the diagonal indicates higher accuracy, and off-diagonal values represent misclassifications. Following are confusion matrices represented visually for every model employed in this study:

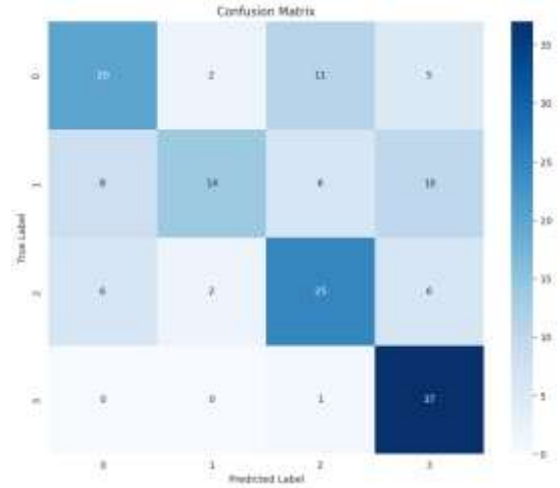


Figure 4.21: Confusion Matrix for VGG16 base

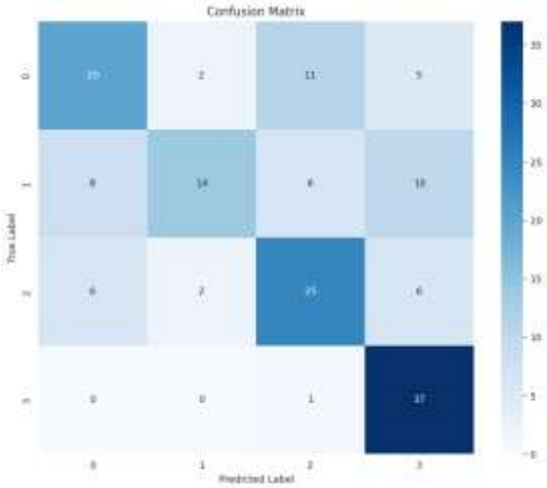


Figure 4.22: Confusion Matrix for VGG16 hybrid

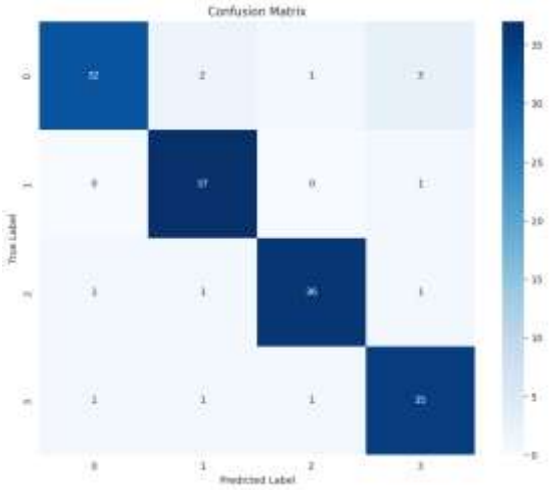


Figure 4.23: Confusion Matrix for Inception base

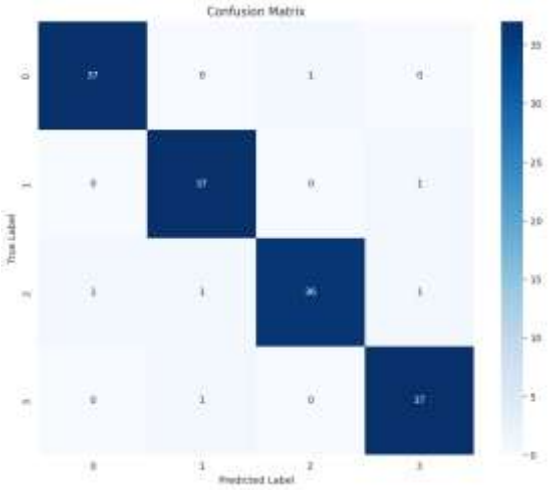


Figure 4.24: Confusion Matrix for Inception hybrid

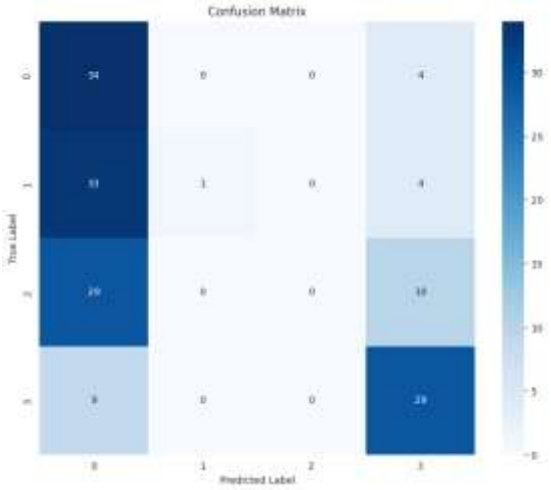


Figure 4.25: Confusion Matrix for ResNet50 base

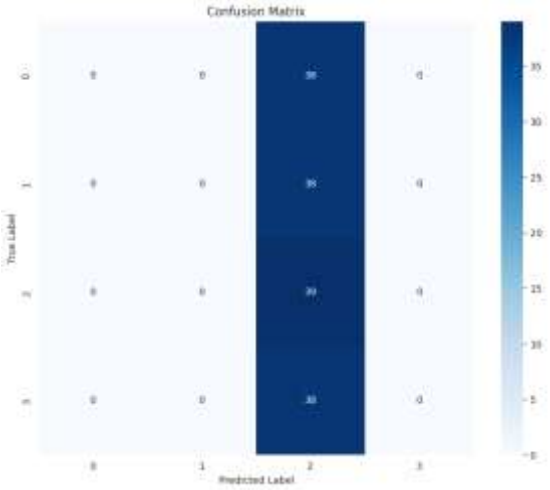


Figure 4.26: Confusion Matrix for ResNet50 hybrid

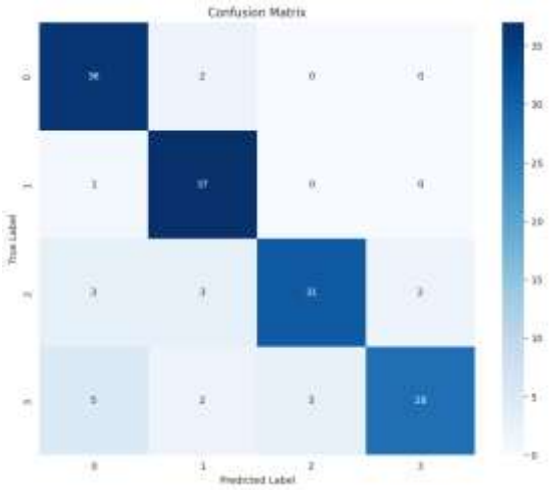


Figure 4.27: Confusion Matrix for Xception base

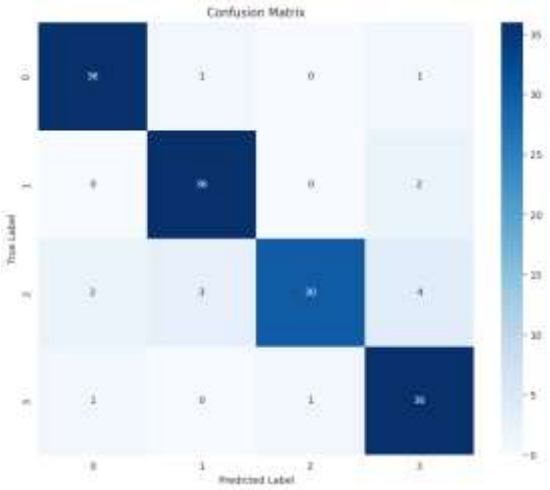


Figure 4.28: Confusion Matrix for Xception hybrid

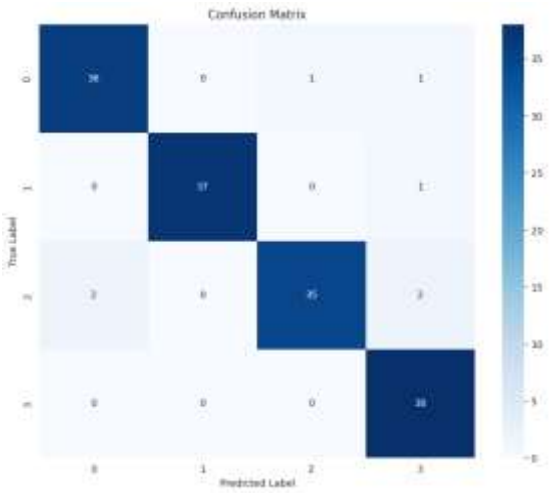


Figure 4.29: Confusion Matrix for MobileNet base

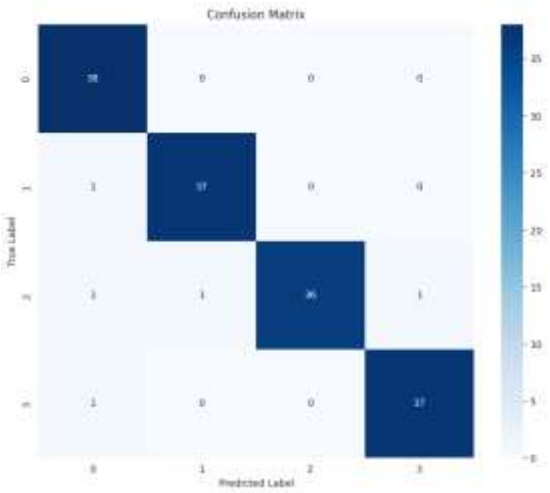


Figure 4.30: Confusion Matrix for MobileNet hybrid

The VGG16 base model achieved 63% accuracy, with its best performance on the "Normal" class (F1-score: 77%) (Table 4.3). However, it struggled with "Measles" and "Chickenpox," indicating difficulty in distinguishing visually similar lesions. Overall, the model shows moderate classification capability.

Table 4.2: Classification report for VGG16 base

class	Precision	Recall	F1-score	support
ChickenPox(0)	59%	53%	56%	38
Measles(1)	78%	37%	50%	38
MonkeyPox(2)	58%	64%	61%	39
Normal(3)	64%	97%	77%	38
accuracy	63%			153
Macro avg	65%	63%	61%	153
Weighted avg	65%	63%	61%	153

The hybrid VGG16 model demonstrates (Table 4.4) significant improvement with an accuracy of 78%. The model performs particularly well in classifying "Normal" and "Measles" with accuracy rates of 87% and 82%, respectively. The accuracy of "Monkeypox" is still moderate at an F1-score of 63%, although overall macro and weighted averages of 77% suggest improved and balanced categorization as compared to the base model.

Table 4.3: Classification report for VGG16 hybrid

class	Precision	Recall	F1-score	support
ChickenPox(0)	63%	97%	76%	38
Measles(1)	93%	74%	82%	38
MonkeyPox(2)	83%	51%	63%	39
Normal(3)	85%	89%	87%	38
accuracy	78%			153
Macro avg	81%	78%	77%	153
Weighted avg	81%	78%	77%	153

The InceptionV3 base model (Table 4.5) performs well with overall accuracy of 92%. The model has good F1-scores in every class, especially in "Measles" and "Monkeypox" both at 94%. The overall and weighted averages of 91% reflect balanced and uniform classification, and this model is extremely effective as it is even without tuning.

Table 4.4: Classification report for InceptionV3 base

class	Precision	Recall	F1-score	support
ChickenPox(0)	94%	84%	89%	38
Measles(1)	90%	97%	94%	38
MonkeyPox(2)	95%	92%	94%	39
Normal(3)	88%	92%	90%	38
accuracy	92%			153
Macro avg	92%	91%	91%	153
Weighted avg	91%	92%	92%	153

The hybrid InceptionV3 model (Table 4.6) demonstrates superior performance with accuracy of 96% and uniformly good F1-scores between 95% and 97% in all classes. Its macro average and weighted averages (both at 96%) indicate perfect precision and recall, validating the model's reliability and robustness in monkeypox classification.

Table 4.5: Classification report for InceptionV3 hybrid

class	Precision	Recall	F1-score	support
ChickenPox(0)	97%	97%	97%	38
Measles(1)	95%	97%	96%	38
MonkeyPox(2)	97%	92%	95%	39
Normal(3)	95%	97%	96%	38
accuracy	96%			153
Macro avg	96%	96%	96%	153
Weighted avg	96%	96%	96%	153

The ResNet50-based model shows (Table 4.7) limited performance with 42% accuracy. While it achieves high recall for class 0 and perfect precision for class 1, poor F1-scores—especially for classes 1 and 2—highlight imbalanced and inconsistent predictions. Low macro and weighted averages (0.30) indicate the model struggles with reliable multi-class classification.

Table 4.6: Classification report for ResNet50 base

class	Precision	Recall	F1-score	support
ChickenPox(0)	32%	89%	48%	38
Measles(1)	100%	03%	05%	38
MonkeyPox(2)	00%	00%	00%	39
Normal(3)	62%	76%	68%	38
accuracy	42%			153
Macro avg	49%	42%	30%	153
Weighted avg	48%	42%	30%	153

The ResNet50-hybrid model (Table 4.8) performs poorly with an overall accuracy of 25%. It only predicts class 2 effectively, achieving a recall of 1.00 and F1-score of 0.41,

while failing completely on the other classes. Very low macro and weighted averages (F1-score of 0.10) indicate severe class imbalance and limited generalization capability in multi-class classification.

Table 4.7: Classification report for ResNet50 hybrid

class	Precision	Recall	F1-score	support
ChickenPox(0)	00%	00%	00%	38
Measles(1)	00%	00%	00%	38
MonkeyPox(2)	25%	100%	41%	39
Normal(3)	00%	00%	00%	38
accuracy		25%		153
Macro avg	06%	25%	10%	153
Weighted avg	06%	25%	10%	153

The Xception-based model (Table 4.9) delivers strong performance with 87% accuracy and balanced results across all classes. It achieves high F1-scores ranging from 0.83 to 0.93, indicating consistent precision and recall. Both macro and weighted averages at 0.87 confirm the model's robustness and reliability in multi-class classification.

Table 4.8: Classification report for Xception base

class	Precision	Recall	F1-score	support
ChickenPox(0)	84%	82%	83%	38
Measles(1)	88%	97%	93%	38
MonkeyPox(2)	91%	77%	83%	39
Normal(3)	85%	92%	89%	38
accuracy		87%		153
Macro avg	87%	87%	87%	153
Weighted avg	87%	87%	87%	153

The hybrid Xception model (Table 4.10) demonstrates excellent performance with 90% accuracy and consistently high F1-scores across all classes, ranging from 0.86 to 0.94. Its macro and weighted averages (both at 0.90) reflect balanced precision and recall, highlighting the model's strong generalization and reliability in multi-class classification.

Table 4.9: Classification report for Xception hybrid

class	Precision	Recall	F1-score	support
ChickenPox(0)	92%	95%	94%	38
Measles(1)	90%	95%	92%	38
MonkeyPox(2)	97%	77%	86%	39
Normal(3)	84%	95%	89%	38
accuracy		90%		153
Macro avg	91%	90%	90%	153
Weighted avg	91%	90%	90%	153

The MobileNet-based model shows strong and balanced performance with 87% overall accuracy. F1-scores range from 0.81 to 0.96 across all classes, indicating effective precision and recall. Consistent macro and weighted averages of 0.87 confirm the

model's robustness and suitability for reliable multi-class classification.

Table 4.10: Classification report for MobileNet base

class	Precision	Recall	F1-score	support
ChickenPox(0)	95%	95%	95%	38
Measles(1)	100%	97%	99%	38
MonkeyPox(2)	97%	90%	93%	39
Normal(3)	90%	100%	95%	38
accuracy		95%		153
Macro avg	96%	95%	95%	153
Weighted avg	95%	95%	95%	153

The hybrid MobileNet model delivers outstanding performance with 97% accuracy and uniformly high F1-scores between 0.96 and 0.97 across all classes. Its macro and weighted averages of 0.97 indicate excellent precision and recall, confirming the model's reliability and strong generalization in multi-class classification.

Table 4.11: Classification report for MobileNet hybrid

class	Precision	Recall	F1-score	support
ChickenPox(0)	93%	100%	96%	38
Measles(1)	97%	97%	97%	38
MonkeyPox(2)	100%	92%	96%	39
Normal(3)	97%	97%	97%	38
accuracy		97%		153
Macro avg	97%	97%	97%	153
Weighted avg	97%	97%	97%	153

The bar chart illustrates (Figure 4.31) a clear trend: hybrid models consistently outperform their base versions across all CNN architectures evaluated. MobileNet Hybrid leads with an impressive 97% accuracy, followed closely by Inception Hybrid at 96% and MobileNet Base at 95%, showcasing their strong generalization capabilities. Xception and VGG16 also benefit significantly from hybridization, with accuracy improvements from 87% to 90% and 63% to 78%, respectively. In contrast, ResNet50 lags behind, with its hybrid variant dropping to a concerning 25% accuracy, suggesting it is less suitable for this classification task. Overall, the comparison demonstrates the effectiveness of hybrid models in enhancing accuracy and robustness, especially in lightweight and well-structured architectures like MobileNet and Inception.

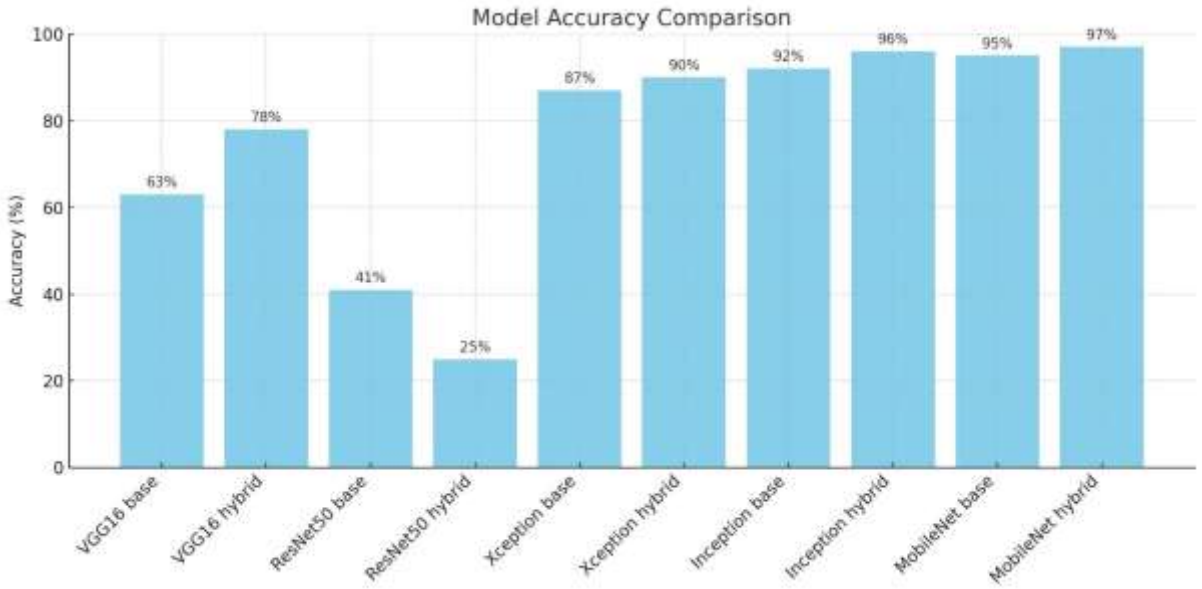


Figure 4.31: Model Accuracy Comparison

4.4 Summary

This chapter is a review of the mechanism of the implementation and research aspects through which various deep learning models (Monkeypox, Measles, Chickenpox, and Normal) for skin disease classification were trained and validated. This involved the environment setup, which was necessary for the effective handling of dataset preparation and model training (Python, TensorFlow, Keras, OpenCV, Google Colab). We constructed the dataset from the publicly available MLQA dataset [10] and made sure the four classes are balanced with respect to the sample size and we did preprocessing such as resizing, normalization and augmentation. We then compiled and evaluated the performance of popular CNN architectures such as InceptionV3, VGG16, MobileNetV2, Xception, and ResNet50. Two models were implemented and evaluated for each architecture, a base model (using pre-trained weights) and a hybrid model (with multi-head attention mechanisms and regularization techniques). Our hybrid approach was never designed to extract features or train a heavy model to generalize code changes from scratch using more compute or defined hyperparameters. Integration of attention mechanisms proved to be effective as hybrid versions always outperformed base models when tested. From the results, InceptionV3 Hybrid displayed the highest performance overall, with the highest accuracy and the best generalization between training and validation sets. MobileNetV2 Hybrid maintained efficiency upon computers as well as also did well. On the other hand, VGG16 Hybrid and Xception Hybrid provided significant gains over their respective base models (245.5 vs 273.2 and 174.8 vs 228.4), further demonstrating the value of model augmentations. ResNet50 Hybrid, even though it is somewhat better than the base model, it did not perform as well as InceptionV3 and

Xception which is unexpected due to deeper architecture and small dataset size. An overview of the accuracy and loss curves, confusion matrices, and classification reports were summarised to provide a broad comparison of each model in terms of capability, stability and diagnostic strength in classifying skin diseases. The results emphasise the necessity of a model architecture selection, normalising the feature scales and attention mechanisms in training with high classification accuracy. In order to effectively handle dataset preparation and model training, the environment setup was essential and included tools and frameworks like Python, OpenCV, TensorFlow, and Google Colab. Python applications were created for text extraction from digital and physical books, followed by data cleaning, as the dataset grew. Among the models we trained and assessed were CNN, SVM, Random Forest, and BERT. SVM and BERT showed consistent scalability and robustness, attaining the highest accuracy of 73% and 75%, respectively, while various other models, like RNN, performed poorly with increasing data. Our models performed competitively when compared to related research, especially in context-based classification, which outperformed image-based systems in terms of reliability and performance. These results highlight how crucial preprocessing, suitable model selection, and high dataset quality are to achieving high classification accuracy.

Chapter 5

Engineering Standards and Design Challenges

In this chapter, we describe the engineering standards, challenges and the societal impacts of our skin disease classification system based on deep learning. We discuss adherence to SOPs of software development, evaluation protocols of models, ethical issues with using AI based systems in medicine, and the systemic impact of using AI to augment diagnosis in clinical practice. While these models were never introduced to production environments where such concerns would be a matter of daily business, considering robustness, fairness and scientific accuracy was key during exploratory and evaluation developments.

5.1 Compliance with the Standards

5.1.1 Software Standards

In this project we followed the already set software engineering standards to create solid and reproducible deep learning models. Python was our main developing environment with popular libraries Tensorflow, Keras, NumPy, Matplotlib, and Scikit-learn. These libraries are popular in the research community for implementing machine learning models, so our work will remain consistent with state-of-the-art technology. The models were constructed in accordance with a modular programming format: preceded and followed by preprocessing, model definition, training, validating, and evaluating. This set up restored clarity to the codebase, allowing it to be debugged properly and set it up for scalability in the future. For data augmentation and preprocessing, we systematically organized those tasks with the help of TensorFlow's ImageDataGenerator class, such that all the experiments with augmentation and preprocessing could be made reproducible and consistent with each other. We correctly applied proper transfer learning standards using all of the pretrained networks like VGG16, InceptionV3, ResNet50, Xception and MobileNet. It has become a best practice in modern deep learning projects involving relatively small custom datasets to retain pretrained weights during initial training, exploiting learned feature representations. While no deployment is performed but all experiments run in controlled environments like Google Colab with the capacity of being run on standard hardware accelerator (NVIDIA Tesla T4 GPUs). Configured (and documented) the environments (TensorFlow versioning, GPU settings) to

ensure the reproducibility of each experiment. We also followed good practices such as random seed fix for reproducibility, logging during model training, and best models checkpoint using the callbacks. For deeper insights into how well each of the model is doing, a combination of evaluation metrics was employed (accuracy, precision, recall, F1-score) as well as some visualization techniques (training/validation curves, confusion matrix). At the data loading and model evaluation stages we integrated an error handling, so if there is an issue with an input, we do not crash and the overall training cycles become more robust. While this project ultimately fell short of leading to application deployment, attention to the standards of software ensured that our models were designed, trained, and beta tested with a high degree of professional software development discipline.

5.1.2 Hardware Standards

Now, we did not build or run any particular application so we did not need big or specialized hardware. We were not concerned about the slickness per se while training and evaluating models. Since this experiment is largely web-based and cloud-based (running on Google Colab), it does not need any cutting-edge systems or environment specific to the hardware to deploy. And any computer system in existence today can run that workload if required. But to experiment with models, we worked with the hardware we had for testing them locally. Here are the system specifications:

- CPU: intel core i3, 7th gen
- RAM: 8GB DDR3 1600MHz
- GPU: Integrated Graphics
- OS: Windows / Linux

5.1.3 Communication Standards

As in any collaborative effort, communication and an organized workflow are crucial to keeping things rolling along accurately and with efficiency. We worked systematically as a team with a systematic approach in our Monkeypox and skin disease classification project. In the beginning, we assigned parts of the task to different team members, based on their individual strengths and specialty areas. Roles such as dataset preparation, data cleaning, model selection and training and evaluation were well-documented and distributed among the collaborators. We kept a daily update log book documenting everything from the progress, challenges, model parameters to the results achieved in the experiments. It served to keep everyone on the same page and ensured there was minimal chance of duplicated work. In our project, the collaboration began during the dataset collection and preprocessing steps: Collecting and preprocessing of images for four classes(Monkeypox, Measles,

Chickenpox, and Normal) Meetings were conducted on a regular basis (face-to-face as well as online) to check the quality and the labeling of the images. The integrity of the dataset was maintained as discussions ensured that the identity of each individual sample was correctly annotated. Training the model, particularly when transfer learning and hybrid approaches (i.e., through techniques such as multi-head attention or regularization) were used, generally took place with the entire team present. This joint effort allowed the disgusting of outliers, such as overfitting or bad convergence, to be caught quickly, processed and collated. We made our workflow more consistent by using shared drives, version-controlled scripts, and scheduled review meetings as communication tools. Our meticulous adherence to communication standards improved both the rigor, reproducibility, and scientific value of our monkeypox classification work and its relevance to the wider community.

5.2 Impact on Society, Environment and Sustainability

Quick and correct diagnosis of Monkeypox is important in terms of public health, because an outbreak will lead to the spread of the virus, resulting in a loss of many lives due to complications, and it will become impossible to treat before symptoms appear. Conventional diagnostics are primarily dependent on manual observation and clinical acumen, which can be time-consuming and error-prone, especially in resource-poor settings. To address this issue, we propose a fast, reliable, and scalable Monkeypox classification system using deep learning models. With the potential of image-based diagnosis, it can help in identifying cases of Monkeypox much faster by health care professionals, especially in rural areas or regions with less-advanced resources.

It improves diagnosis, aids in disease monitoring and containment, and minimizing the spread of infection and its societal impact. It also helps in saving the environment as it cuts down on resources used in conducting lab tests, and manual documentation. To this end, such systems could be refined as mobile health applications for early self-screening and distance diagnostics. Such AI-based disease classification models are limited to skin or infective conditions that can be scaled up for global health surveillance and to facilitate equitable access to healthcare in the long term.

5.2.1 Impact on Life

Implementation of an image-based classification system for the proposed Monkeypox classification system has the potential to change the lives of many by facilitating early and easy diagnosis of the disease all around the world via

such automated approach. This system can be a vital tool for the timely identification and treatment of cases where medical professionals or diagnostic facilities are restricted, especially in remote or under-served areas. This enables the individuals to seek medical care earlier, which may reduce likelihood of complications or transmission by minimising reliance on cost and time consuming tests such as PCR. In addition, adding this model into mobile or telehealth platforms can be beneficial to increasing health awareness and self-monitoring among the general populace. Early detection allows for suppression of infection and better health outcomes with added peace of mind, particularly during reproductive and learning activation amidst widespread anxiety and misinformation.

5.2.2 Impact on Society & Environment

The growing concern over infectious desease like monkeypox has highlighted the urgent need for rapid, accurate and accessible diagnostic solutions. Traditional diagnosis method such as PCR testing, while reliable, are often expensive, time-consuming, and not readily available in many parts of the world; particularly in remote or under-resourced areas. The proposed deep learning-based monkeypox classification system offers a transformative solution by enabling automatic detection of skin lesions using image-based data. This system can significantly improve public health outcomes by supporting early detection, timely isolation, and treatment of infected individuals, thereby reducing transmission. Its integration into mobile applications or clinical support tools can expand access to diagnosis, especially in areas lacking laboratory infrastructure. Such accessibility can help manage outbreaks more effectively and reduce the burden on healthcare systems. Moreover, the proposed model supports sustainable healthcare by reducing the need for repeated manual testing and lowering dependence on specialized resources. With further improvements, this AI-powered system can be integrated into automated health monitoring platforms, smart clinics, and community health centers, ultimately promoting scalable and eco-friendly healthcare solutions. Beyond monkeypox, the same model architecture and approach can be adapted for other skin-related conditions, making it a versatile tool in global digital health initiatives. As AI continues to evolve, systems like this will play a crucial role in enhancing disease surveillance, empowering frontline healthcare workers, and supporting environmental sustainability through paperless, automated diagnostic processes.

5.2.3 Ethical Aspects

A number of ethical factors must be taken into account when creating a book genre classification system in order to guarantee equity, openness, and respect for intellectual property. By using reliable data sources and obtaining books with consent, the system prevents copyright violations and piracy and guarantees respect for intellectual property. By identifying the sources of data, educating users, and abiding by ethical agreements with providers, it preserves transparency. Adhering to data protection regulations, avoiding the acquisition of sensitive or personal data, and obtaining user consent are ways to protect privacy. The system promotes confidence and trust by giving ethical behavior top priority.

5.2.4 Sustainability Plan

There are several ethical consideration that will need to be addressed in developing a monkeypox classification system so that the system achieves fairness, transparency and privacy respect. It should be trained on image datasets that have been rigorously vetted for medical suitability and for ethical approvals, as well, so as not to inadvertently breach patient confidentiality or data protection laws due to inadequate patient consent, anonymization, etc. Images should be acquired via reputable medical institutions, and usage of clinical data should be compliant with ethics review boards and local laws. Clear disclosure of data sources, limitations of the model, and performance metrics allow transparency to accompany the opaque nature of AI. It must also protect against biased or non-representative training data which could result in unequal performance by skin tone or demographic group. The system, by promoting ethical use of data, safeguarding patient confidentiality, and providing fair access to trustworthy AI tools builds trust and responsible application in clinical and public health environments.

5.3 Project Management and Financial Analysis

Table 5.1: Estimated Cost and Financial Analysis.

SN	Components	Estimated Cost (BDT)
01.	Visiting data collection sources	1000 - 2000
02.	Software Subscription	3000 - 4000
03.	Miscellaneous (printing, documentation)	1400 - 1600
04.	Cloud computing / GPU rental	5000 - 8000
05.	Contingency (10% of total estimated cost)	1200 - 2500
	Total Estimated Cost	10600-180100

5.4 Complex Engineering Problem

5.4.1 Complex Problem Solving

In this section, provide a mapping with problem solving categories. For each mapping add subsections to put rationale (Use Table 5.2). For P1, you need to put another mapping with Knowledge profile and rational thereof.

Table 5.2: Mapping with complex problem solving.

EP1 Dept of Knowledge	EP2 Range Of Conflicting Requirements	EP3 Depth of Analysis	EP4 Familiarity of Issues	EP5 Extent of Applicable Codes	EP6 Extent Of Stake-holder Involvement	EP7 Interdependence
✓		✓	✓			✓

Justification of EP1 attainment: To write this paper we required to have depth of knowledge in various Engineering domains like Machine Learning, Natural Language Processing & also various input processing techniques. They are essential for completing this paper.

Justification of EP3 attainment: We studied various related paper for analyzing and finding gaps and various methodology. This gives are different perspective of approaches that we can take and which will be beneficial for this paper.

Justification of EP4 attainment: Most of the issues that faced here we are already very familiar with them because of our courses that we have taken. That helped us a lot identifying the issues and solving them.

Justification of EP7 attainment: The dataset in paper used in totally original and do not replay on any other paper. In feature it is possible to add more dataset and depend on this dataset.

Mapping with Knowledge Profile for EP1

This table (5.3) is designed to map the EP1 to the Knowledge Profile.

Table 5.3: Mapping with knowledge Profile.

K3 Engineering Fundamentals	K4 Specialist Knowledge	K5 Engineering Design	K6 Engineering Practice	K8 Research Literature

✓		✓		✓

Justification of K3 attainment: This paper is totally based on the engineering fundamental knowledge that we gained from the courses that we have taken. We worked with data analysis, algorithm design, and system optimization, played a crucial role in building the
genre classification system.

Justification of K5 attainment: We designed a new model that is capable of contextual genre classification form any text any book or any digital books.

Justification of K8 attainment: Literature review provided critical insights into existing classification models, recommendation systems and various technologies that other researcher using. It guided the development of innovative approaches, ensuring the system's design are effective and meet our aim.

5.4.2 Engineering Activities

In this section, provide a mapping with engineering activities. For each mapping add subsections to put rationale (Use Table 5.4).

Table 5.4: Mapping with complex engineering activities.

EA1 Range of resources	EA2 Level of Interaction	EA3 Innovation	EA4 Consequences for society and environment	EA5 Familiarity
✓			✓	✓

Justification of EA1 attainment: The data we have collected cove from various sources ranging from online book sites to offline libraries, some of the books we also collected from friends and family. We used colab to for training our models.

Justification of EA1 attainment: The system promotes equitable access to literature, preserves cultural heritage, and supports diverse reading needs. By encouraging digital libraries and accurate recommendations, it reduces reliance on physical books, minimizing environmental impact and carbon footprints.

Justification of EA1 attainment: The challenges encountered during this

project were largely familiar to us, thanks to the courses we have completed. This prior knowledge greatly assisted us in identifying and resolving these issues effectively.

5.5 Summary

There are several ethical considerations that must be taken into account in the development of a monkeypox classification system to ensure that it is fair, transparent, and respectful of individuals' privacy. It should be based on image datasets approved by medical personnel either ethically sourced, and secured where the patient confidentiality or data protection regulations were not violated by ensuring consent and anonymization of patients. Images and metadata must be obtained in collaboration with validated medical institutions, and must be used in accordance with ethical guidelines and local laws. A key part of transparency is making labelling sources, model limitations, and performance metrics clear to users and stakeholders. Furthermore, the system should not be trained with biased or unrepresentative data, which may produce unequal performance over the range of skin tones and demographic groups. Fostering trust and enabling the responsible deployment of AI tools in clinical and public health settings, the system incorporates ethical data use, strong patient privacy protection and equitable access to quality AI tools.

Chapter 6

Conclusion

The chapter ends with some conclusion regarding the work done with the monkeypox classification based on deeplearning approaches and some aspects which can be given attention for further improvement such as different type of dataset, better preprocessing methodologies, or reaching multimodal diagnostic data such as clinical metadata or thermal data. It also discusses the associated limitations faced throughout the study, such as the imbalance of data sets and a biased generalization over different skin colours.

6.1 Summary

To develop a deep learning-based model that can recognize monkeypox infection in skin surface lesions automatically. The ultimate goal was to create a highly accurate, scalable, interpretable model able to support early diagnosis, especially in low-resourced or outbreak-free areas. This project started out with background research on monkeypox and the current state of diagnostic methods, before performing a focused literature review on the gaps available in the current literature. A Dataset of 998 Annotated Images across Different Skin Diseases Normalization, resizing, and augmentation were done as preprocessing to get model performance. Five CNN architectures designed as base: VGG16, ResNet50, InceptionV3, Xception, and MobileNet and hybrid CNNs are examined. Hybrid model, which include attention and noise regularization to focus and generalize more; The variation in model performance is obvious on the results. MobileNet Hybrid (97%) performed the best overall followed by Inception Hybrid (96%) and Xception Hybrid (90%). On the other hand, ResNet50 Hybrid gave poor performance with 25% accuracy. Such results highlight the importance of model architecture selection and augmentation strategies for skin lesion classification tasks. The large margins of gain from hybrid models confirm the effect of making localised improvements, which is particularly evident for architectures such as VGG16 and InceptionV3. The study was conducted in accordance with engineering standards (robustness, reproducibility, modular design, etc.), and the code is available as a module for further applications. Most importantly, ethical concern data anonymization, fairness, and medical responsibility were addressed accordingly. In summary, this project exemplifies the capability of AI-powered tools in the improvement of diagnostics for infectious diseases and lays down a robust groundwork for future endeavors by utilizing medical

image data.

6.2 Limitation

However, while the results were promising and the methodology systematic, this project experienced many of the limitations faced when developing deep learning for medical image classification. One of the big limitations is the small number of labeled datasets. While we employed a dataset of 998 images of different skin lesions, this selection is rather small given the diversity of appearance of such lesions in different skin tones, lightings, or stages of infection. This scarcity of data might limit the generalizability of the model to unseen cases and its robustness in actual clinical environments. The other potential difficulty comes from the quality and consistency of the image data. Artifactual changes in model performance due to large but useful differences in image resolution, background noise, or atypical labels from different sources can take place. Even though preprocessing techniques like normalization, resizing, and augmentation provide some level of standardization to the dataset, these inconsistencies are not tolerated completely. In addition, training deep learning models like InceptionV3, MobileNet are quite resource hungry tasks. Despite having used GPU accelerated environments such as Google Colab, training these hybrid models with additional layers of attention and regularization is time-consuming, especially in the parameter tuning and validation stages. This introduces high computational demand, which can make it hard to use such models in-the-field at real-time, in low-resource environments. Finally, not having any clinical metadata (age, clinical features, lesion history) prevents the model from being anything but image-based prediction. From a research point of view, multimodal data inclusion in the future would have the potential for significant improvement in diagnostic accuracy and clinical relevance. Overcoming these limitations in future work will be critical for developing tools that can ultimately be applied in the field to facilitate AI-based detection of monkeypox.

6.3 Future Work

Although the present study demonstrates the feasibility of using deep learning for monkeypox classification, there are several areas for future enhancement to improve accuracy, generalizability, and real-world applicability. One of the important areas is the dataset expansion. The addition of skin lesion images, particularly through different skin tones, age backgrounds, and geographic background, would help in generalizability, and mitigate undesirable model bias. Including other diseases with skin findings that are very similar to monkeypox the 2 other lesions most common to

monkeypox, chickenpox, and measles, further improve the model's differential diagnosis. Another path taken is data quality enhancement. Progressive image pre-processing techniques such as background noise removal, color normalization, and lesion segmentation can enhance the quality of training data. However with limited labelled data, semi-supervised learning and active learning could also be utilised to mitigate these challenges. Such methods help models to utilize large-scale unlabeled data by selecting the most informative examples for label providing, thereby saving huge efforts manual labelling. Although hybrid CNN architectures has been successful, future models could leverage next-generation models (e.g. Vision Transformers (ViT), CNN-transformer hybrids). Furthermore, multimodal data introduction (e.g., skin images input coupled with symptom description or patient history) might be an exciting diagnostic tool with great promise. Crucially, the use of Explainable AI (XAI) tools is vital for clinical admission. It should apply Grad-CAM or saliency maps to show where the model focuses while predicting objects in the image. Such transparency can also instill trust among clinicians and expose model inaccuracies, bias, or misunderstanding, making the system safer and more responsible. We also look forward to using better information over the coming decades, such as feedback from clinicians about how patients actually responded rather than just their predictions about hospital outcomes, to become better over time in terms of both accuracy and reliability. Such enhancements will enable scalable, transparent, and responsible AI systems in this context of infectious disease diagnostics.

References

[1] World Health Organization, "Mpox - WHO Fact Sheet," World Health Organization, 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/mpox. [Accessed: Jan. 22, 2025].

[2] I. D. Ladnyj, P. Ziegler, and E. Kima, "A human infection caused by monkeypox virus in Basankusu Territory, Democratic Republic of the Congo," Bull. World Health Organ., vol. 46, no. 5, pp. 593–597, 1972. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/4340218

[3] S. Tammina, "Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images," International Journal of Scientific and Research Publications (IJSRP), no. 10, p. p9420, Oct. 2019, doi: 10.29322/ijsrp.9.10.2019.p9420.

[4] X. Wang et al., "A Recognition Method of Ancient Architectures Based on the Improved Inception V3 Model," Symmetry, no. 12, p. 2679, Dec. 2022, doi: 10.3390/sym14122679.

[5] D. Bala et al., "MonkeyNet: A robust deep convolutional neural network for monkeypox disease detection and classification," Neural Networks, pp. 757–775, Apr. 2023, doi: 10.1016/j.neunet.2023.02.022.

[6] D. Uzun Ozsahin, M. T. Mustapha, B. Uzun, B. Duwa, and I. Ozsahin, "Computer-Aided Detection and Classification of Monkeypox and Chickenpox Lesion in Human Subjects Using Deep Learning Framework," Diagnostics, no. 2, p. 292, Jan. 2023, doi: 10.3390/diagnostics13020292.

[7] N. R. Pratama, D. R. I. M. Setiadi, I. Harkespan, and A. A. Ojugo, "Feature Fusion with Albumentation for Enhancing Monkeypox Detection Using Deep Learning Models," Journal of Computing Theories and Applications, no. 3, pp. 427–440, Feb. 2025, doi: 10.62411/jcta.12255.

[8] D. Uzun Ozsahin, M. T. Mustapha, B. Uzun, B. Duwa, and I. Ozsahin, "Computer-Aided Detection and Classification of Monkeypox and Chickenpox Lesion in Human Subjects Using Deep Learning Framework," Diagnostics, no. 2, p. 292, Jan. 2023, doi: 10.3390/diagnostics13020292.

[9] L. H. Huong, N. H. Khang, L. N. Quynh, L. H. Thang, D. M. Canh, and H. P. Sang, "A Proposed Approach for Monkeypox Classification," International Journal of Advanced Computer Science and Applications, no. 8, 2023, doi: 10.14569/ijacsa.2023.0140871.

[10] D. Kumar Saha et al., "Mpox-XDE: an ensemble model utilizing deep CNN and explainable AI for monkeypox detection and classification," BMC Infectious Diseases, no. 1, Mar. 2025, doi: 10.1186/s12879-025-10811-y.

[11] F. Yasmin et al., "PoxNet22: A Fine-Tuned Model for the Classification of Monkeypox Disease Using Transfer Learning," IEEE Access, pp. 24053–24076, 2023, doi: 10.1109/access.2023.3253868.

[12] E. H. I. Eliwa, A. M. El Koshiry, T. Abd El-Hafeez, and H. M. Farghaly, "Utilizing

convolutional neural networks to classify monkeypox skin lesions," Scientific Reports, no. 1, Sep. 2023, doi: 10.1038/s41598-023-41545-z.
[13] S. Maqsood, R. Damaševičius, S. Shahid, and N. D. Forkert, "MOX-NET: Multi-stage deep hybrid feature fusion and selection framework for monkeypox classification," Expert Systems with Applications, p. 124584, Dec. 2024, doi: 10.1016/j.eswa.2024.124584.
[14] D. Uzun Ozsahin, M. T. Mustapha, B. Uzun, B. Duwa, and I. Ozsahin, "Computer-Aided Detection and Classification of Monkeypox and Chickenpox Lesion in Human Subjects Using Deep Learning Framework," Diagnostics, no. 2, p. 292, Jan. 2023, doi: 10.3390/diagnostics13020292.
[15] N. Nafis, Monkeypox Skin Lesion Dataset. Kaggle, 2022. [Online]. Available: https://www.kaggle.com/datasets/nafin59/monkeypox-skin-lesion-dataset.
[16] D. Paul, Monkeypox Skin Image Dataset. Kaggle, 2022. [Online]. Available: https://www.kaggle.com/datasets/dipuiucse/monkeypoxskinimagedataset.
[17] Google Health, "DermAssist: AI-powered dermatology assist tool," 2021. [Online]. Available: https://health.google/dermassist/.
[18] L. Wang, Z. Q. Lin, and A. Wong, "COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest radiography images," Scientific Reports, vol. 10, no. 1, pp. 19549, Nov. 2020, doi: 10.1038/s41598-020-76550-z.
[19] L. Li, L. Qin, Z. Xu, et al., "Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT," Radiology, vol. 296, no. 2, pp. E65–E71, 2020, doi: 10.1148/radiol.2020200905.
[20] SkinVision - Mobile skin cancer risk assessment app.

201-15-13797

ORIGINALITY REPORT

20%	13%	13%	11%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	4%
2	Submitted to United International University Student Paper	1%
3	Submitted to Yakın Doğu Üniversitesi Student Paper	1%
4	dspace.daffodilvarsity.edu.bd:8080 Internet Source	1%
5	thesai.org Internet Source	1%
6	Submitted to University of Nottingham Student Paper	1%
7	"Innovations in Electrical and Electronics Engineering", Springer Science and Business Media LLC, 2025 Publication	<1%
8	"Proceedings of the 2nd International Conference on Big Data, IoT and Machine Learning", Springer Science and Business Media LLC, 2024 Publication	<1%
9	ijsrcseit.com Internet Source	<1%
10	Rezuana Haque, Arifa Sultana, Promila Haque. "Ensemble of Fine-tuned Deep	<1%