PREMIER UNIVERSITY, CHITTAGONG
Department of Computer Science and Engineering

Final AI Report

On

Skin Cancer Type Detection

SUBMITTED BY

Name: Md Nishadul Islam Chy Shezan
ID: 0222220005101014

Name: Md Sakib
ID: 0222220005101019

Name: Fariha Rashid Noha
ID: 0222220005101035

Name: Rimjhim Dey
ID: 0222220005101039

Under the Supervision of

Avisheak Das
Lecturer
Department of Computer Science and Engineering
Premier University, Chittagong

6 April, 2025

Author's Declaration of Originality

We hereby declare that the project work entitled **"Skin Cancer Type Detection"** submitted to Premier University, Chittagong is a record of original work carried out by us under the guidance of Mr. Avisheak Das, Lecturer, Department of Computer Science and Engineering, Premier University, Chittagong.

This work is submitted in fulfillment of the requirements for the degree of Bachelor of Science in Computer Science and Engineering. We also affirm that the results of this project have not been submitted to any other university.

Md Nishadul Islam Chy Shezan
ID: 0222220005101014

Md Sakib
ID: 0222220005101019

Fariha Rashid Noha
ID: 0222220005101035

Rimjhim Dey
ID: 0222220005101039

Table of Contents

List of Figures

List of Tables

Abstract

Skin cancer is one of the most commonly diagnosed cancers globally, and its early detection is critical for effective treatment and patient survival. However, traditional diagnostic methods that rely on visual inspection and dermatoscopic analysis can be time-consuming and prone to human error. This project aims to develop an automated skin cancer classification system using deep learning techniques to assist in faster and more accurate diagnoses.

We utilized a publicly available dataset from Kaggle, consisting of dermatoscopic images of various types of skin lesions. The project involved training and evaluating multiple deep learning models, including DenseNet-121, ResNet-50, EfficientNet-B0, and Vision Transformer (ViT), through transfer learning. These models were assessed using standard performance metrics such as accuracy, precision, recall, and F1-score. Among them, DenseNet-121 achieved the highest performance, reaching an accuracy of 88.6%.

This work highlights the potential of artificial intelligence in medical image analysis and demonstrates that deep learning can significantly improve the classification of skin cancer types. By supporting dermatologists with reliable diagnostic tools, such systems can reduce the risk of misdiagnosis and contribute to better clinical outcomes.

Keywords: deep learning, transfer learning, convolutional neural network (CNN), DenseNet-121, skin cancer classification

CHAPTER 1

INTRODUCTION

1.1 Problem Statement

Skin cancer is one of the most common and life-threatening types of cancer worldwide, with millions of new cases diagnosed annually. Early detection is critical for effective treatment and improved survival rates. However, traditional diagnostic methods, such as visual inspection by dermatologists, can be subjective, time-consuming, and dependent on the availability of experienced medical professionals. This creates a significant need for innovative, automated approaches to enhance diagnostic accuracy and accessibility.

With the rapid advancement of artificial intelligence and deep learning technologies, there is a strong potential to revolutionize skin cancer diagnosis. Automated models can provide fast, consistent, and highly accurate identification of skin cancer types from medical images, reducing the risk of human error and enabling early intervention. This project addresses these challenges by developing a state-of-the-art AI-driven diagnostic system.

In this project, we utilize the HAM10000 dataset [1], which contains a diverse and extensive collection of skin lesion images, to train a deep learning model based on the DenseNet-121 architecture. Our model achieves an impressive accuracy of 88.6% in classifying various types of skin cancer, demonstrating its potential as a powerful diagnostic tool. Furthermore, we enhance our study by creating a custom dataset of 200 additional images, further testing and validating the model's performance in varied conditions.

Beyond high classification accuracy, our model offers an intuitive and informative graphical representation of the predictions. These visualizations provide medical professionals with clearer insights into model decisions, facilitating better clinical interpretations, and aiding in more informed decision-making. This approach represents a significant step toward improving diagnostic efficiency and patient outcomes.

1.2 Motivation

The motivation behind this project stems from the alarming increase in the incidence of global skin cancer and the urgent need for an early and accurate diagnosis. In many regions, access to specialized dermatological care is limited, leading to delayed diagnoses and poorer prognoses. By developing an AI-driven diagnostic model, we aim to bridge this gap and empower healthcare providers with advanced tools for precise and timely detection.

The success of machine learning models in medical imaging has already shown transformative potential. By applying these techniques to skin cancer detection, we can reduce diagnostic subjectivity, improve efficiency, and provide consistent data-driven evaluations. The high accuracy of our model and the ability to visually represent predictions address critical gaps in current diagnostic practices, ultimately improving patient care and outcomes.

1.3 Contributions

In this project, we make the following impactful contributions:

- Developed a robust and efficient deep learning model based on the DenseNet-121 architecture, achieving 88.6

- Constructed a novice dataset of 200 dermoscopic images, which we manually labeled and validated, further confirming the model's generalizability and enhancing its practical applicability.

- Demonstrated that the model maintained its performance on the custom dataset, supporting the reliability of its predictions across different image distributions.

- Designed clear and informative graphical representations of the model predictions, enabling intuitive analysis and helping medical professionals in better decision-making.

- Provided a comprehensive performance evaluation that showcased the potential of the model for real-world clinical deployment and its ability to support an early and accurate diagnosis.

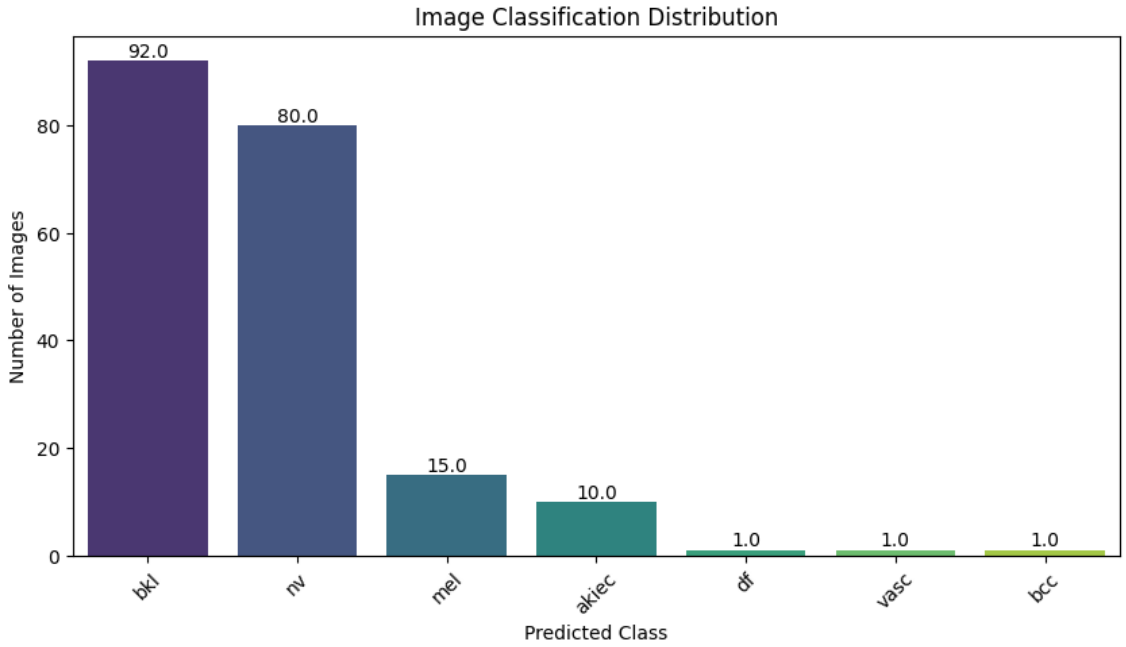

Figure 1.1. Image Classification Distribution of New Test dataset

1.4 Report Outline

In this section, we outline the structure of this report, detailing the key chapters and their contributions to the overall study.

The report begins with the problem statement, establishing the context, challenges, and objectives of the project. The motivation section underscores the significance of early skin cancer detection and the potential impact of AI-driven diagnostics. The methodology chapter elaborates on dataset preparation, model architecture, and the training process. The results section presents comprehensive performance metrics and graphical outputs, highlighting the model's effectiveness. Finally, the report concludes with an in-depth discussion on the broader implications, limitations, and future directions of this research.

CHAPTER 2

LITERATURE REVIEW

2.1 Literature Review

Deep learning has emerged as a powerful tool in medical image analysis, particularly in the domain of dermatology where accurate skin lesion classification is critical. Numerous studies have shown that convolutional neural networks (CNNs) can match or even surpass human-level performance in specific diagnostic tasks.

Esteva et al. [2] demonstrated the viability of CNNs for skin cancer detection by training a model on over 120,000 clinical images, achieving dermatologist-level accuracy. Their work laid the foundation for automated diagnostic support using deep learning. Han et al. [**?**] extended this by building a multi-class classification system capable of distinguishing between seven types of skin lesions using a CNN-based approach.

A key milestone in the field was the release of the HAM10000 dataset by Tschandl et al. [3], which provided a standardized and diverse dataset of dermatoscopic images. It has since been widely adopted for benchmarking skin lesion classification tasks.

To address limitations such as small dataset sizes and training time, transfer learning has become a common strategy. Models pre-trained on ImageNet — such as ResNet, DenseNet, EfficientNet, and Vision Transformers (ViT) — have been fine-tuned for medical imaging tasks with notable success. Recent work, including that by Das et al. [**?**], explored ensemble-based techniques to further improve classification reliability.

Building upon this literature, our project experimented with four advanced models: ResNet-50, EfficientNet, ViT, and DenseNet-121. Among them, DenseNet-121 consistently outperformed the others in terms of accuracy and generalization on the HAM10000 dataset. This study contributes to existing research by evaluating multiple deep learning architectures and confirming the superior performance of DenseNet-121 in practical skin lesion classification.

An essential aspect of deploying deep learning models in the medical domain is their interpretability. Clinicians need to trust the decisions made by AI models, especially in high-stakes applications like cancer diagnosis. In response to this, several studies have focused on integrating explainable AI (XAI) tools into diagnostic pipelines. One of the most widely used techniques is Gradient-weighted Class Activation Mapping (Grad-CAM), which provides visual explanations by highlighting the regions in the input image that were most influential in the model's decision.

Grad-CAM overlays heatmaps on dermatoscopic images, helping experts verify whether the model is focusing on relevant lesion areas. This improves the transparency of the diagnostic process and enables human-in-the-loop validation. Studies such as those by Selvaraju et al. [4] have demonstrated the efficacy of Grad-CAM in enhancing model interpretability without compromising performance.

CHAPTER 3

METHODOLOGY

3.1 Dataset Description

The dataset used for this study is the HAM10000 dataset, which contains 12,827 images of skin lesions across seven distinct classes of skin cancer. The dataset includes the following attributes:

Table 3.1. Dataset Attributes Description

Attribute	Description
lesion_id	Unique identifier for each lesion.
image_id	Unique identifier for each image.
dx	Diagnosis of the lesion (e.g., akiec, bcc, mel, etc.).
dx_type	Type of diagnostic procedure used (e.g., histopathology, clinical, consensus, etc.).
age	Age of the patient.
sex	Gender of the patient.
localization	Location of the lesion on the body.

The dataset is well-structured with 12,827 non-null entries and a total of seven columns. A sample of the metadata is shown below:

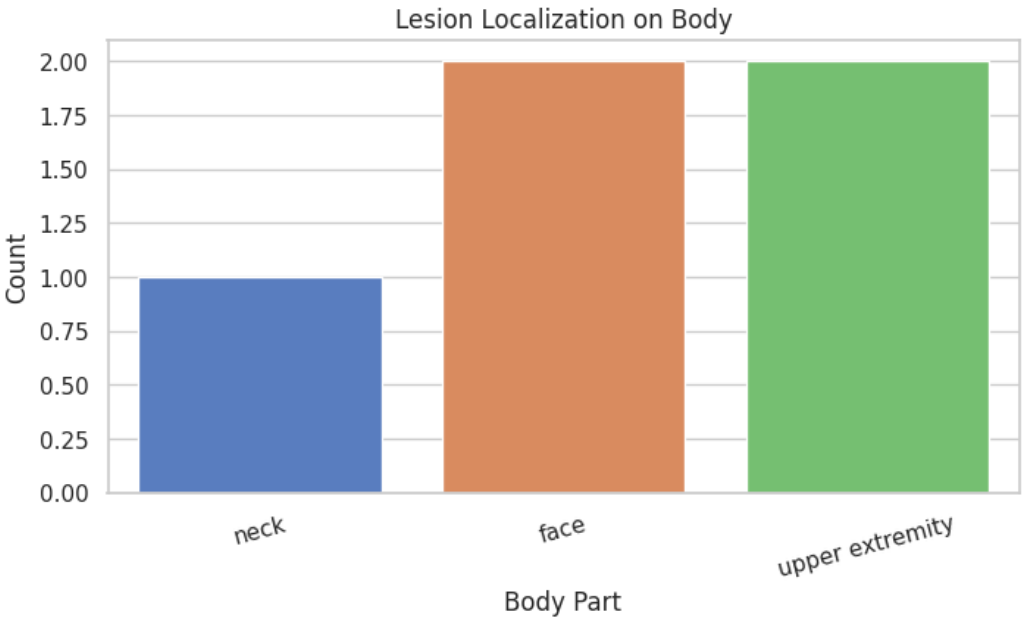

Figure 3.1. Lesion Localization on Body

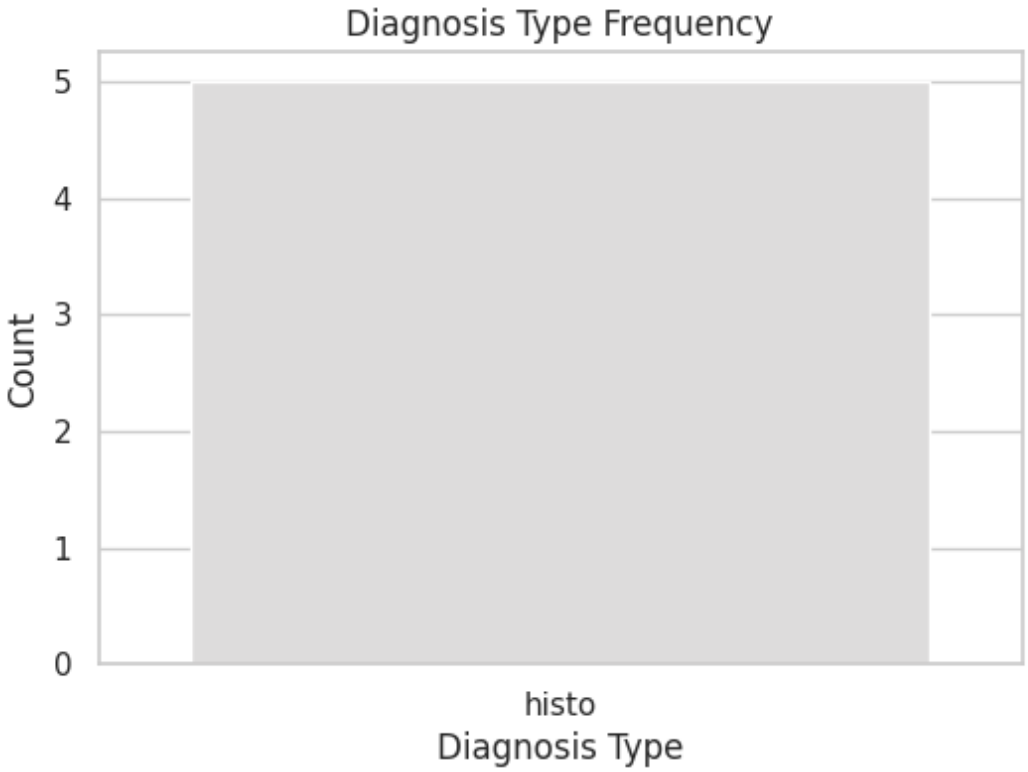

Figure 3.2. Diagnosis Type Frequency

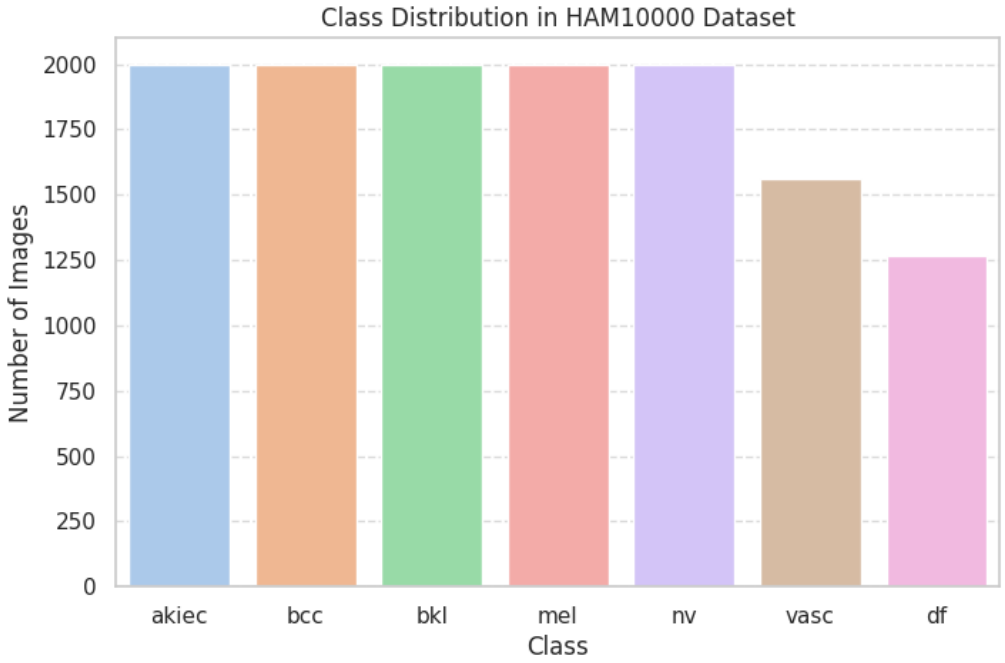

Figure 3.3. Class Distribution in HAM10000 Dataset

Table 3.2. Sample Metadata from HAM10000 Dataset

lesion_id	**image_id**	**dx**	**dx_type**	**age**	**sex**	**localization**
HAM_0002644	ISIC_0029417	akiec	histo	80.0	female	neck
HAM_0006002	ISIC_0029915	akiec	histo	50.0	female	face
HAM_0000549	ISIC_0029360	akiec	histo	70.0	male	upper extremity
HAM_0000549	ISIC_0026152	akiec	histo	70.0	male	upper extremity
HAM_0000673	ISIC_0029659	akiec	histo	70.0	female	face

The class distribution across the dataset is fairly balanced:

Table 3.3. Class Distribution

Class	**Count**
akiec	2000
bcc	2000
bkl	2000
mel	2000
nv	2000
vasc	1562
df	1265

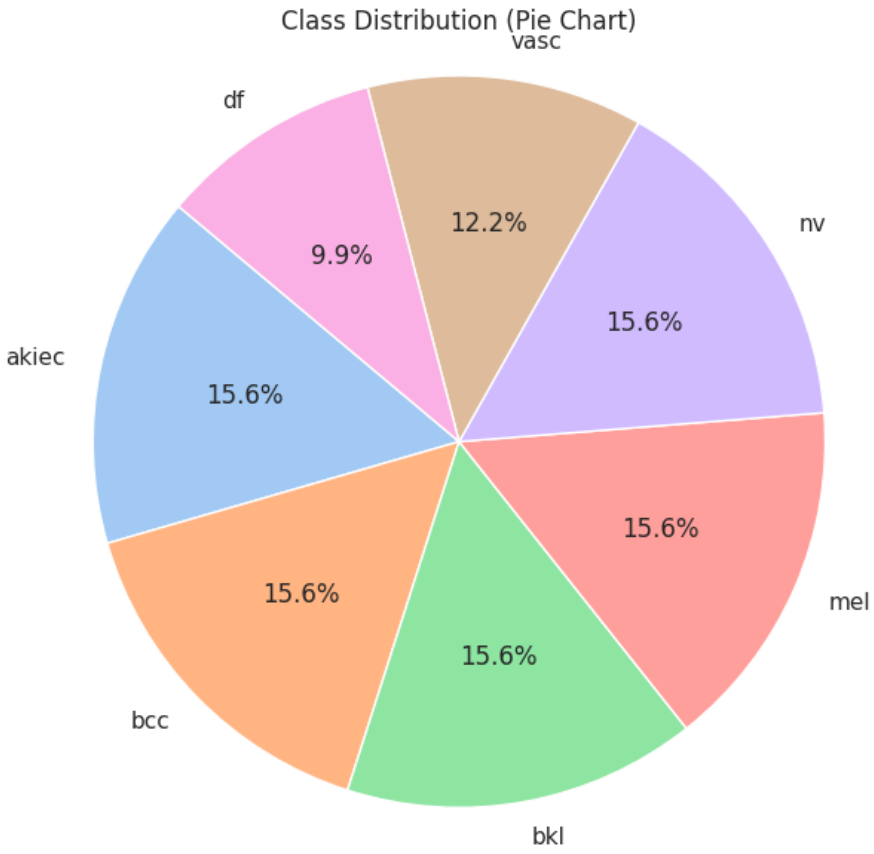

Figure 3.4. Class Distribution Pie-chart

3.2 Preprocessing

Preprocessing steps are essential to ensure data quality and optimal model performance. We applied the following steps:

3.2.1 Step 01: Data Cleaning

- Removed any duplicate images and missing data. - Standardized image sizes and formats for consistency. - Normalized pixel values to a range of 0 to 1 for faster convergence during training.

3.2.2 Step 02: Data Augmentation

- Applied random transformations like rotation, flipping, and zooming to prevent overfitting.
- Balanced class representation by oversampling underrepresented classes.

3.3 Experimented Models

To identify the most effective model for skin cancer classification, we evaluated several cutting-edge deep learning architectures. The performance of each model was assessed based on classification accuracy using both the HAM10000 dataset and our custom-labeled dataset. The following summarizes our findings:

- **DenseNet-121**: Achieved the highest classification accuracy of **88.6%**. DenseNet-121's architecture, characterized by dense connections between layers, allowed better gradient flow and reuse of features, which significantly improved learning efficiency and accuracy on both datasets.
- **ResNet-50**: Recorded an accuracy of **22%**. Despite its powerful residual learning capability, ResNet underperformed on our dataset, likely due to its deeper architecture not being well-suited to the size and variability of the training data.
- **EfficientNet-B0**: Achieved an accuracy of **21.02%**. While EfficientNet is optimized for performance and efficiency using compound scaling, the limited dataset size restricted its potential, leading to underfitting.
- **Vision Transformer (ViT)**: Reached an accuracy of **19%**. Although ViT excels in capturing long-range dependencies and global features, it typically requires much larger datasets and more training epochs. In our experiments, its performance suffered due to data limitations and higher computational demands.

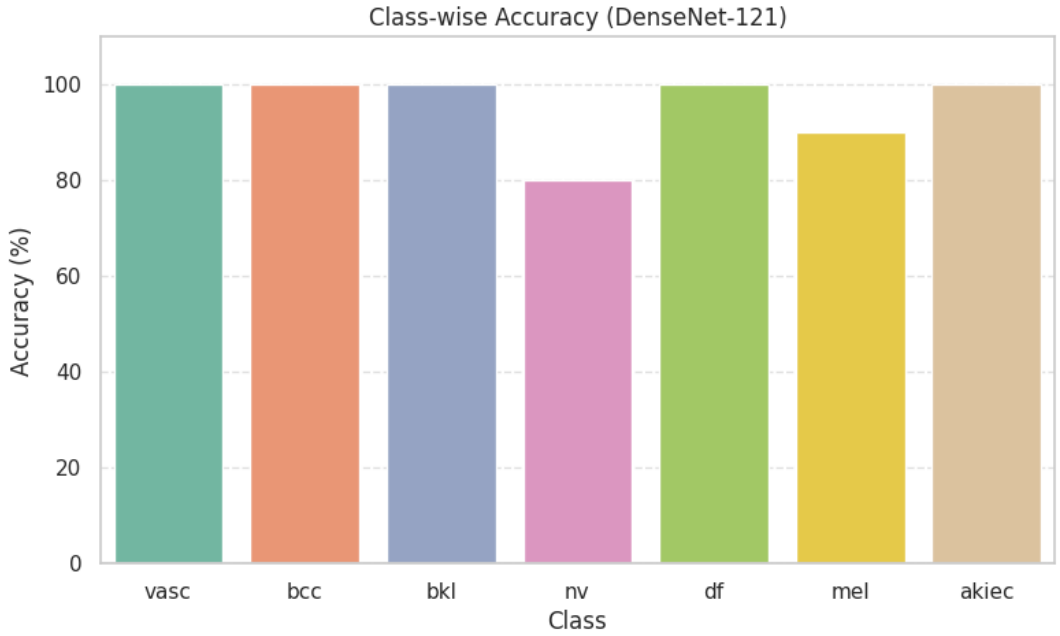

Figure 3.5. Class Wise Accuracy using DenseNet-121

Based on these experiments, DenseNet-121 was selected for final deployment due to its superior performance and training stability.

3.4 Class-wise Accuracy

The following table summarizes the accuracy achieved per class using the DenseNet-121 model. The results demonstrate strong classification ability across most categories, especially for common and well-represented classes in the dataset.

Table 3.4. Class-wise Accuracy using DenseNet-121

Class	**Accuracy**
Vascular lesions (vasc)	100.00%
Basal cell carcinoma (bcc)	100.00%
Benign keratosis-like lesions (bkl)	100.00%
Melanocytic nevi (nv)	80.00%
Dermatofibroma (df)	100.00%
Melanoma (mel)	90.00%
Actinic keratoses (akiec)	100.00%

CHAPTER 4

EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Performance Evaluation Metrics

We assessed the performance of each model using standard evaluation metrics: accuracy, precision, recall, and F1-score. These metrics helped us evaluate not just overall correctness but also how well the model performed across different types of skin lesions.

4.2 Hyperparameter Settings

The models were trained and tested using Google Colab in Python. We used the Adam optimizer as imported from `torch.optim`, and the following hyperparameters were applied consistently during training:

- Learning Rate: 0.001
- Batch Size: 32
- Epochs: 50
- Optimizer: Adam

These settings were determined after multiple trials, balancing model accuracy with training time.

Figure 4.1. Accuracy vs Training time for the implemented models

4.3 Comparison among Implemented Models

We experimented with four deep learning models — DenseNet-121, ResNet-50, EfficientNet-B0, and Vision Transformer (ViT). The table below summarizes their classification accuracy and approximate training time:

Table 4.1. Model Comparison

Model	Accuracy	Training Time
DenseNet-121	88.6%	2 hours
ResNet-50	22.0%	1.5 hours
EfficientNet-B0	21.02%	2 hours
Vision Transformer (ViT)	19.0%	2.5 hours

DenseNet-121 significantly outperformed the other architectures in terms of accuracy. Based on this result, we chose DenseNet-121 for our final skin cancer classification implementation.

CHAPTER 5

CONCLUSION AND FUTURE WORKS

5.1 Performance Evaluation Metrics

We utilized standard evaluation metrics—accuracy, precision, recall, and F1-score—to assess the performance of each model. These metrics offer a comprehensive analysis of classification ability, especially in distinguishing between multiple types of skin lesions. They ensure that the model performs well not just on overall accuracy but also on correctly identifying positive cases and minimizing false alarms.

5.2 Hyperparameter Settings

All models were trained using Python on Google Colab, and optimized by tuning the following hyperparameters:

- **Learning Rate:** 0.001 — enabled efficient and stable convergence.
- **Batch Size:** 32 — provided a balance between computation efficiency and model performance.
- **Epochs:** 50 — allowed adequate training without overfitting.
- **Optimizer:** Adam — chosen for its adaptive learning rate and strong performance in sparse-gradient scenarios.

5.3 Conclusion

The results from our experiments clearly indicate that DenseNet-121 significantly outperforms other architectures, achieving an accuracy of 88.6%. In contrast, other models like ResNet-50, EfficientNet-B0, and Vision Transformer (ViT) showed relatively poor accuracy—below 23%—making them less suitable for this classification task with the current dataset.

This outcome highlights the effectiveness of using advanced deep learning architectures like DenseNet-121 in medical image analysis. By leveraging transfer learning and a carefully curated dataset, the model was able to classify different types of skin cancer lesions with high reliability. This work sets a foundation for future research focused on enhancing diagnostic tools using AI.

5.4 Impact Assessment and Responsible Practices

5.4.1 Societal and Cultural Impact

Our proposed model has the potential to assist in the early detection of skin cancer, contributing to reduced mortality and alleviating healthcare burdens. However, to ensure successful adoption, it is vital to address cultural barriers, digital literacy, and healthcare accessibility, particularly in underserved regions.

5.4.2 Health and Safety Considerations

While the model demonstrates high accuracy, it must be used as a decision-support tool alongside professional medical evaluation. Continuous validation using diverse datasets is essential to prevent misdiagnosis and ensure consistent performance across demographic variations.

5.4.3 Legal and Regulatory Compliance

Strict adherence to privacy regulations such as HIPAA and GDPR is required when using medical datasets. This includes proper anonymization, informed consent, and secure handling of patient data. Ethical and legal deployment mandates transparent practices and accountability.

5.4.4 Professional and Ethical Responsibility

It is crucial to ensure that the model performs equitably across age groups, skin tones, and genders. Addressing dataset biases and involving medical experts in the model's lifecycle will help mitigate risks. Ethical deployment also involves disclosing limitations and uncertainties to users and stakeholders.

5.4.5 Sustainability (Technical, Economic, Environmental)

From a technical perspective, using a relatively lightweight architecture like DenseNet-121 ensures manageable resource requirements. Economically, deploying such a tool in low-resource areas could reduce diagnostic costs. Environmentally, optimized training procedures help minimize energy consumption, supporting green AI initiatives.

5.5 Future Work

Future research will focus on:

- Expanding the dataset to include underrepresented skin tones and rare lesion types.
- Enhancing explainability using tools like Grad-CAM to build trust among clinicians and users.
- Collaborating with dermatologists for clinical validation and feedback.
- Developing a deployable and user-friendly interface suitable for integration into telemedicine platforms or mobile diagnostic tools.

REFERENCES

[1] K.-R. Müller, P. Tschandl, and N. Codella, "Ham10000: Human against machine with 10000 training images," https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000, 2018, accessed: 2025-05-03.

[2] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, 2017. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/28117445/

[3] P. Tschandl, C. Rosendahl, and H. Kittler, "The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Scientific Data*, vol. 5, p. 180161, 2018. [Online]. Available: https://www.nature.com/articles/sdata2018161

[4] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE international conference on computer vision (ICCV)*, 2017, pp. 618–626.