Detecting Skin Disease Using A Hybrid Deep Learning Approach

Submitted by

Md. Mehedi Hasan

ID: 201 35-535

Department of Software Engineering

Supervised by

Dr. Marzia Ahmed

Assistant Professor

Department of Software Engineering

This thesis report is submitted in partial fulfilment of the requirements for the Bachelor of Science degree in Software Engineering.
Summer 2025

APPROVAL

APPROVAL

This thesis titled on "**Detecting Skin Disease Using A Hybrid Deep Learning Approach**", submitted by **Md. Mehedi Hasan (ID: 201-35-535)** to the Department of Software Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Software Engineering and approval as to its style and contents.

BOARD OF EXAMINERS

Fazla Elahe
______________________________ **Chairman**

Dr. Md. Fazla Elahe
Assistant Professor & Associate Head
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

13.9.25
______________________________ **Internal Examiner 1**

Dr. Marzia Ahmed
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

13.09.2025
______________________________ **Internal Examiner 2**

Dr. Shabnom Mustary
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

13.09.25
______________________________ **External Examiner**

Mohammod Abul Kashem
Professor
Department of Computer Science and Engineering
Dhaka University of Engineering & Technology, Gazipur.

DECLARATION

DECLARATION

This is to declare that this project/thesis, titled "Detecting Skin Disease Using A Hybrid Deep Learning Approach", is my original work, carried out under the supervision of Dr. Marzia Ahmed. No part of this work has been submitted elsewhere, partially or entirely, for the award of any other degree or diploma. All project-related information will remain confidential and shall not be disclosed without the formal consent of the project supervisor. Relevant previous works presented in this report have been appropriately acknowledged and cited. The plagiarism policy, as stated by the supervisor, has been maintained.

Supervised by

Marzia

Dr. Marzia Ahmed
Assistant Professor
Department of Software Engineering
Daffodil International University

Submitted by

Md. Mehedi Hasan

Md. Mehedi Hasan
ID: 201 35-535
Department of Software Engineering
Daffodil International University

LETTER OF SUBMITTAL

As a part of my thesis work titled: "Detecting Skin Disease Using A Hybrid Deep Learning Approach" I worked on an extensive research project to enhance the efficiency of device-assisted skin disease detection based on deep learning supervised learning approach. It also provided me with critical insight of the real-mile challenges of deploying Artificial Intelligence (AI) into the medical diagnostic domain.

Unfortunately, this thesis also allowed me to go through works of different deep learning architecture, especially a combined CNN with ANN. These were a few things I learnt the hard way, the HAM10000 dataset was an important fun through experience, it gave me the freedom to groove on training data, model and then, testing performances, so had to really go deep in medical image analysis and AI-based classification system.

To get the most out of my research, I worked with experts in the field that provided mentorship during the research process, I used the datasets from Kaggle with PyTorch that pushed me to grow my technical skills and analytical thinking. I also explored datasets quality weaknesses for AI models and also discussed about architecture choice importance for context-specific applications such as skin disease detection.

To summarize, the thesis work was both a learning experience and a realization for me; it helped me academically that is to strengthen my technical skills in machine learning and image classification and also a bit pragmatically to have a wider vision and how AI can be used and its implications for healthcare, outside the lab.

Yours sincerely,

Md. Mehedi Hasan

ID: 201-35-535

ACKNOWLEDGMENTS

I want to express my gratitude to Dr. Marzia Ahmed, my supervisor, for all of her assistance and guidance during the planning and writing of my thesis. Her astute advice and strategic guidance have been crucial to finishing this assignment. She helped me begin my thesis, and her consistent encouragement of my ideas kept me going throughout the entire process. I want to express my gratitude to all of my friends, family, and coworkers for supporting and encouraging me no matter what.

ABSTRACT

Diseases of the skin have taken on critical importance for global health, bringing into focus a necessity for accurate and quick diagnostics. This work demonstrates the development of both a Hybrid Convolutional Neural Network and Artificial Neural Network (Hybrid CNN-ANN) model for automated classification of skin diseases using dermatoscopic images. We trained our model on the HAM10000 database, which is a collection of 10,000 high-quality images representing various types of skin diseases (including melanoma, nephews and benign and malignant pinches). The approach includes the data acquisition, preprocessing, model architecture construction, and training, as well as evaluation. In addition, we conducted the pre-processing of image treatments that provided better feature extraction and learning efficiency for our model using traditional methods such as resizing, normalize and data augmentation. The proposed Hybrid CNN-ANN leverages convolutional layers for hierarchical spatial feature learning and a fully connected ANN model for fine-grained classification. Such combination can help the model effectively learn complex representations and subtle changes of skin lesion patterns. We trained and improved our models on Kaggle because it has great computing power and deep learning capabilities. The accuracy, precision, recall and F1-score were well studied for the Hybrid CNN-ANN model. The results of the experiments indicated that the model was robust and had a high accuracy in classification with an average recognition rate of 77.94%, which showed its potential in accurate detection for skin diseases. This work emphasizes the importance of integrating CNN and ANN architectures to improve feature representations learning, which yield better classification results. The results will support progress in the direction of AI-assisted dermatological diagnosis, and lay groundwork for broader applications to automated medical image analysis tasks. The objective of this work is to offer a scalable yet cost-effective approach of training intelligent diagnostic models for skin disease classification, leveraging publicly available datasets and cloud-based training.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Introduction

Skin disorders are amongst the most frequent medical problems, affecting a huge number of people of all ages and backgrounds around the globe. Early and precise diagnosis of such diseases is vital for an effective treatment and hence better patient outcome. Still, conventional diagnostic approaches are dependent on dermatologists, whose availability is a dissimilarity of access, particularly in rural or under-resourced locales. Making use of Artificial Intelligence (AI) and Machine Learning (ML) which are continuously innovating, automated skin disease detection based on deep learning models can be used as an attractive solution to fill this gap.

This work aims to propose a Hybrid Convolutional Neural Network and Artificial Neural Network (Hybrid CNN-ANN) model for the automated classification of skin diseases using dermatoscopic imaging. Contrary to classical methods that depend on fully-connected networks or pre-trained models, we propose a Hybrid CNN-ANN architecture, which benefits from feature extraction ability of Convolutional Neural Networks (CNN) and classification power of Artificial Neural Networks (ANN). Such a hybrid strategy allows the model to first learn the hierarchical spatial features from images of skin lesions and later get to learn the complex patterns from these images thereby improving its accuracy during the diagnostic phase.

The main dataset for this project is HAM10000 dataset, which is downloaded from Kaggle. This dataset includes 10,000 dermatoscopic images from a variety of skin conditions including but not limited to melanoma, nevi, and many benign and malignant lesions. Various preprocessing steps like resizing the images, normalizing the pixel values, and augmenting the dataset are performed to have a well-fed dataset for the model and to generalize easily.

Model Development: Creation and optimization of Hybrid CNN-ANN architecture using deep learning framework PyTorch. PyTorch is selected for developing and fine-tuning deep learning models (due to dynamic computation graph, flexibility, and exploratory friendly nature). The model is trained and validated on Kaggle, which provides us powerful computational power and deep-learning infrastructure.

Besides the Hybrid CNN-ANN, we will be used MobileNetV2 and Vision Transformers (ViT) architectures as alternatives in order to compare. Excellent models available as benchmarks against which one could compare the performance of varying degrees of lightweight architectures against more complex attention-area based models in the skin disease classification task.

This thesis report will fully cover the full development pipeline, starting with the review of related works and the existing research efforts on AI based skin disease detection. It will describe the data pre-processing, design of model architecture, training process, and hyperparameter optimization methods. This report will also provide the experimental results with evaluation metrics like accuracy, precision, recall, and F1-score to aid and enhance the evaluation of model effectiveness and reliability.

Finally, I will be discussing the limitations and challenges faced within the project, and possible future improvements. The final section of the report discusses the importance of Hybrid CNN-

ANN architectures for medical image analysis and dermatological diagnostics, and how they will influence the future of AI-powered dermatological diagnostics.
Throughout the report, some of the terms/keywords which may also include abbreviation as well will be used, such as skin disease classification, artificial intelligence, deep learning, CNN, Ann, pytorch, image preprocessing, and feature extraction & transfer learning will be defined in more depth, to insight you on their meanings and what exactly it refers to, to make sure that you are at least much clearer on what this if not all terms that you are not familiar with.

1.1.1 Definitions

- **Skin Disorders:** Any type of skin disease (skin rash, skin infection, all benign or malignant tumors) or chronic skin condition (eg, eczema, psoriasis).
- **Artificial Intelligence:** Capacity of a computer to perform tasks that are normally done by humans because they require human intelligence such as learning, thinking, problem-solving and decision-making.
- **Machine Learning (ML) —** a subset of AI in which programs learn to recognize patterns when taught with data, so can predict or decide something, but with limited knowledge cannot directly program to do an arbitrary task.
- **CNN–**Regular feedforward neural network, some layers are more configurable (Dense, Conv and Recurrent layers are the typically seen types, which are similar to different types of nodes in a graph convolution) By automatically learning convolutions, CNNs are able to efficiently capture visual information in terms of spatial hierarchies and appearances.
- **ANNS (Artificial Neural Network):** It is one of the types of machine learning model representing the biological neural net. ANNs consist of networks of interconnecting nodes (or "neurons") that process data to learn patterns and predict.
- **Hybrid CNN-ANN:** The such hybrid architecture which uses Convolutional Neural Networks for feature extraction and are then fed to Artificial Neural Network for classification process. This is a hybrid method where the model can extract high dimensional information from image content.
- **Deep learning —** Using very deep architectures (hence, the name deep learning), deep learning is a type of machine learning that allows us to find complex, non-linear relationships in high dimensional data through the use of neural networks.
- **PyTorch:** PyTorch is an open-source machine learning library used for applications such as computer vision and natural language processing. Dynamic computation graphs are not always in tensorflow, but we tend to use pytorch on dynamic computation graphs when they are easy and have dynamic behavior.
- **HAM10000 Database:** An open and entirely attributed set of 10,000 dermatoscopic photos of different skin lesions, both benign and malignant. It is a benchmark dataset for training and evaluation of skin disease classification model.
- **Image Preprocessing:** The process by which raw images are transformed into a form compatible for model training. Which consists of resizing, normalization, noise reduction, and feature enhancement to make learning more effective.
- **Data Augmentation:** A method of random replacing which adds different forms to the data and enlarges the dataset, which also solves the overfitting problem.
- **Feature Extraction:** The method of identifying isolated and relevant features/patterns in the raw data (such as images) for a particular task such as classification.
- **Evaluation Metrics (Accuracy, Precision, Recall, F1-Score):** These are the three well-used statistical methods that one uses to examine a workout in the classification picture. While accuracy quantifies the number of overall right predictions, precision is the amount of true positive predictions, recall is the capacity to get all the positive instances, and F1-score – a balance of both.

CHAPTER 2

RELATED WORK

2.1 Related Work

Many works based on the problem we are trying to solve have already been done differently.
Here, it consists of summaries for eight research papers on skin lesion analysis, mainly focusing on image preprocessing, segmentation, classification, and diagnosis using various methodologies. Here's a concise overview:

<u>Paper - 1: Image Preprocessing and Evaluation Metrics</u>

- Image preprocessing techniques include resizing, grayscale conversion, noise removal, and contrast enhancement.
- Evaluation metrics for skin lesion segmentation and classification include precision, recall, specificity, accuracy, F1-score, AUC, and IoU.

<u>Paper - 2: Skin Lesion Classification with CNNs</u>

- Uses CNNs (ResNet50 and VGG) for skin lesion classification.
- The ensemble model outperforms individual models, focusing on feature interpretation through prediction difference analysis.

<u>Paper - 3: Late Detection of Melanoma</u>

- Highlights late detection challenges and the role of computer-aided diagnosis.
- Introduces methods for hair removal, lesion segmentation, and various classification techniques, including CNNs.

<u>Paper - 4: Expert System with ECOC SVM</u>

- Utilizes an expert system with AlexNET-based transfer learning for skin lesion classification.
- It achieves an 86.21% classification accuracy and compares favorably with existing literature.

<u>Paper - 5: Skin Lesion Classification Using ABCDE Checklist</u>

- Proposes a binary classification system using CNNs based on the ABCDE criteria.
- Emphasizes preprocessing steps, model training, and evaluation metrics.

<u>Paper - 6: CNN-Based Skin Cancer Classification</u>

- Deploys a CNN model for skin cancer classification with careful tuning of hyperparameters.

- Implements data augmentation, dropout regularization, and batch normalization for improved performance.

Paper - 7: Dermoscopy Image Segmentation

- Introduces a method using blind deconvolution, L*a*b color space transformation, and morphological operations for accurate skin lesion segmentation.
- Achieves a 95.33% accuracy in segmenting skin lesions.

Paper - 8: RCM Image Classification and Interpretability

- Addresses sample sparsity through data augmentation and utilizes SURF and Haralick features for classification.
- Incorporates a visual pattern weighted localization method for interpreting classification outcomes.

2.2 Review

In reviewing existing literature on automated skin disease detection, it becomes evident that the majority of research models are predominantly based on Convolutional Neural Network (CNN) architectures. These models leverage CNN's powerful feature extraction capabilities, making them a natural choice for image classification tasks. Various studies have utilized CNNs directly or have built upon them through transfer learning using architectures like ResNet, Inception, MobileNetV2, and EfficientNet. The ability of CNNs to process both grayscale (2-channel) and RGB (3-channel) images effectively has further cemented their role as the standard in image-based diagnostic systems.

However, upon examining prior works, we found a significant gap — there is a lack of models utilizing Artificial Neural Network (ANN)-centric architectures for skin disease classification. While CNNs excel at extracting spatial hierarchies in images, ANNs, due to their dense, fully connected structures, are inherently powerful in learning complex patterns once the spatial features are well-represented. This dense connectivity can potentially offer enhanced performance in scenarios where fine-grained classification is essential, provided the feature extraction is effectively managed by preceding layers.

In our approach, we propose a Hybrid CNN-ANN architecture, where CNN layers are employed for primary spatial feature extraction, and ANN layers are utilized for deeper, high-level pattern learning and classification. Unlike traditional CNN-only models, this hybrid design allows for a more robust and flexible learning process by combining the strengths of both architectures.

In terms of data preprocessing and augmentation, our methodology is aligned with best practices established in previous works. Based on insights drawn from eight notable research papers, we adopted common preprocessing techniques, including image resizing, normalization, augmentation (rotation, flipping, scaling), and dataset splitting for training, validation, and testing phases. While these processes are standard across most studies, our

model necessitated fine-tuning these preprocessing steps specifically for the hybrid architecture’s requirements.
An anticipated question may arise regarding whether using a novel hybrid model should demand a different data processing pipeline compared to conventional CNN models. Indeed, our model requires careful consideration in how input data is prepared to ensure compatibility with its architecture. However, instead of rigidly following the preprocessing techniques tailored for other models, we curated and optimized our data processing pipeline based on the specific needs of our Hybrid CNN-ANN model. Moreover, this preprocessing strategy is adaptable, allowing seamless experimentation with alternative architectures such as Vision Transformers (ViT) and MobileNetV2, enabling a broader comparative analysis.
This paper explores the benefits that a CNN model can achieve with respect to its performance, through a denser ANN-based classification head nested within. This research investigates a new architectural design which can help us learn hidden abstract patterns in a deep feature generated by CNN exploring extremely dense ANN connections between the layers of the data in both types of layers to achieve higher classification accuracy than the state of art methods.

CHAPTER 3

METHODOLOGY

3.1 System Design & Implementation:

The proposed system employs a skin cancer classification model developed using PyTorch, a popular deep-learning framework. The model is constructed as a neural network architecture, specifically an Artificial Neural Network (ANN). It is designed to categorise skin lesion images into one of seven classes: ’nv’, ’mel’, ’bkl’, ’bcc’, ’vasc’, ’akiec’, and ’df’, representing different skin conditions. (Fig. 1) The workflow involves data preprocessing, model construction, data augmentation, and training/evaluation phases.

3.1.1 Data Processing and Handling:

The first step is conversion from CSV file to pixel values in NumPy array. So we convert these arrays into torch tensors so that can use torch neural networks model with them. The data was split into training and test set by a normal 75-25 divider. We implement the oversampling methods over the data into the Data Loader by leveraging PyTorch's Data Loader and Random Sampler to solve class imbalance. This way each class will be represented more evenly.

3.1.2 Data Augmentation and Training:

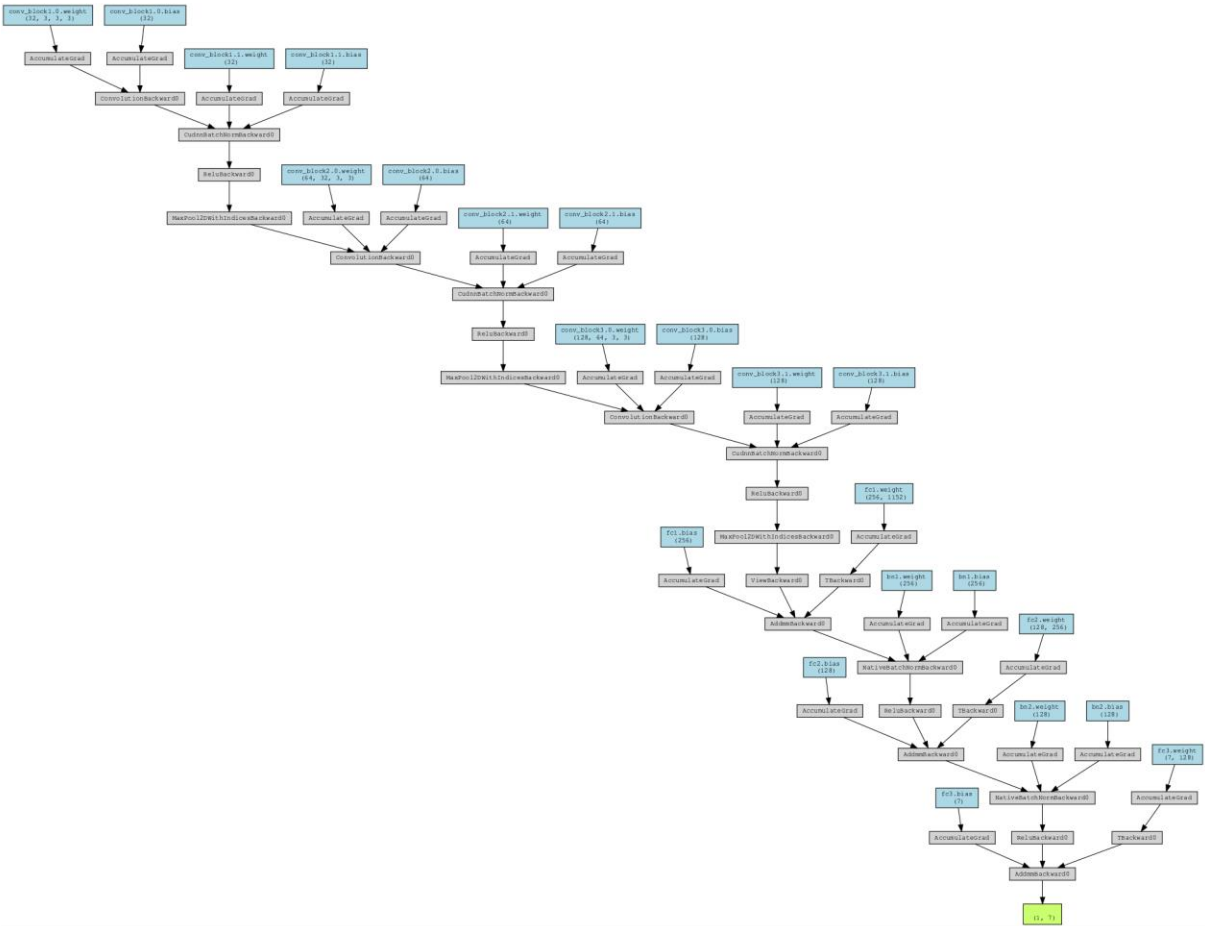
The training data is augmented using some of data augmentation transforms provided by PyTorch In this stage, we perform transformations on the skin images such as rotation, flipping, color jittering and resizing; all these help the model generalize better on unseen images with multiple types of skin lesions. For training and testing loading of this data is done through Data Loader instance which initiates batch wise processing during model training.

3.1.3 Model Construction and Training:

A skin cancer classification model, shown in Fig 2, is an ANN model that contains five fully connected layers sequentially followed by a batch normalization layer, a ReLU activation layer, and several dropout layers to prevent overfitting. The model is trained with the Adam optimizer using a learning rate of 0.01 and performs a ReduceLROnPlateau with the validation accuracy to reduce the learning rate. Training for 100 epochs (Fig. These include forward passes, calculating the loss using the cross-entropy, then applying the backpropagation and optimization (as depicted in Fig. 2).

3.1.4 Evaluation and Validation:

During training through epochs, we log the metrics of the model that tell us about the performance like the accuracy on the training dataset, validation loss, and accuracy on the validation dataset. These will tell us if the model is learning and how well it performs in test data that it has not seen before. The last step, the evaluation, is extremely important as if the network does not generalize and make a correct prediction on an unseen skin lesion image the model will be useless. In this section, we introduce the system design of a CNN architecture for image classification tasks built using MobileNetV2. Based on one of the more efficient architectures, MobileNetV2, the model uses transfer learning, and it does its part well on the image-related side of things. The last stage in the system design is for the design of creating a custom CNN where the MobileNetV2 is acting as the base and some specific layers on top to using for feature extraction and classification specific to this project.



(Fig. 1: Hybrid Model Architecture)

3.2 Architecture Overview:

It is a combining model from CNN layers as feature extractor (backbone) and followed by fully connected ANN layers for classifcation. It is for image classification for output of num_classes = 7.

3.2.1 Detailed Breakdown

1. CNN Backbone (Feature Extractor)
This section gets spatial characteristics from the photos that are input.

- Input Size: (Batch_Size, 3, 28, 28) → Assuming RGB images of size 28x28.

Conv Block 1:

- Conv2d(3, 32, kernel_size=3, padding=1): → Output: (Batch, 32, 28, 28)
- BatchNorm2d(32)
- ReLU Activation
- MaxPool2d(2x2): → Output: (Batch, 32, 14, 14)

Conv Block 2:

- Conv2d(32, 64, kernel_size=3, padding=1): → Output: (Batch, 64, 14, 14)
- BatchNorm2d(64)
- ReLU
- MaxPool2d(2x2): → Output: (Batch, 64, 7, 7)

Conv Block 3:

- Conv2d(64, 128, kernel_size=3, padding=1): → Output: (Batch, 128, 7, 7)
- BatchNorm2d(128)
- ReLU
- MaxPool2d(2x2): → Output: (Batch, 128, 3, 3)

2. ANN Layers (Classifier Head)
After feature extraction, it transitions to dense layers for classification.

- Flatten: → (Batch, 128 * 3 * 3) → (Batch, 1152)
- Dropout (p=0.3): → Regularization to prevent overfitting.

Fully Connected Layers:

- fc1: Linear(1152 → 256)
  - BatchNorm1d(256)
  - ReLU
  - Dropout(0.3)
- fc2: Linear(256 → 128)
  - BatchNorm1d(128)
  - ReLU
  - Dropout(0.3)
- fc3: Linear(128 → num_classes (7))
  - No activation (softmax is applied during loss calculation, e.g., CrossEntropyLoss).

3.2.2 Customization:

The Hybrid CNN-ANN architecture provides more ANN layers above the CNN backbone to make the model applicable to a particular image classification problem. The CNN backbone consists of 3 convolutional blocks that reflect the continuous spatial hierarchies and learned feature representations within the input images.

After that we start with 2 fully connected (dense) layers with 256 and 128 units respectively to increase the capacity of the model in learning higher level and more abstract feature representations, adding Batch Normalization, ReLU activation and Dropout (0.3) to prevent overfitting. These are feature refiners where these layers allow model to learn more complex features from spatial feature produced by CNN layers.

3.2.3 Output Layer:

The final output layer is a dense (fully connected) layer with seven units, corresponding to the seven target classes of the classification task. This layer outputs the logits (raw scores) for each class. Although a SoftMax activation is not explicitly applied in the model's forward pass, it is intended to be used during inference or within the CrossEntropyLoss function during training to convert logits into probability distributions across the seven classes. This output layer is responsible for making the final classification decision based on the features learned through the preceding CNN and ANN layers.

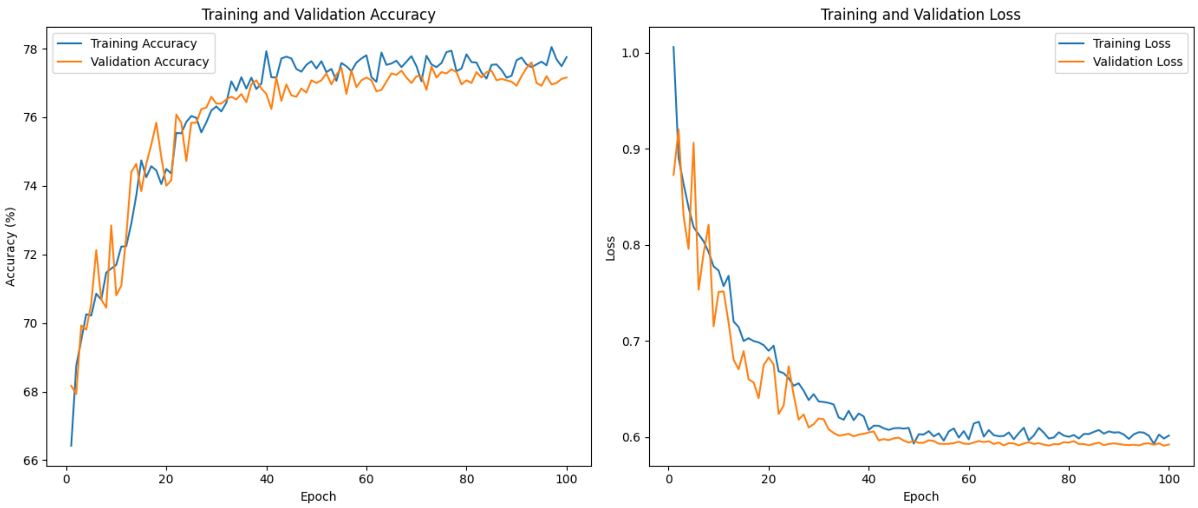
CHAPTER 4

RESULT AND DISCUSSION

4.1 Brief

To evaluate the effectiveness of the proposed Hybrid CNN-ANN model for skin lesion classification, the model was trained on the HAM10000 dataset for 100 epochs. The CNN layers served as a feature extractor from dermoscopic images, while the ANN layers performed classification into seven skin lesion categories. The architecture consisted of three convolutional blocks followed by batch normalization, dropout layers (0.3), and three fully connected dense layers. The final layer used softmax activation for multi-class prediction.

The model was trained using the Adam optimizer with a learning rate of 0.01, and a learning rate scheduler (ReduceLROnPlateau) was applied to dynamically reduce learning rate based on validation performance. The CrossEntropyLoss function was used to optimize classification performance.



(Fig. 2. Graph of Training, Validation Loss and Validation Accuracy of Hybrid model.)

4.2 Training and Validation Results:

As shown in Fig. 2, the training accuracy improved consistently across epochs, reaching a peak of approximately 79%, while validation accuracy stabilized just below, around 77–78%. The training and validation loss curves showed a sharp decrease within the first 20–30 epochs, then converged smoothly.

However, both accuracy and loss graphs exhibit spikes and minor oscillations, especially in the validation curves, indicating slight overfitting and instability in prediction consistency. These fluctuations may stem from:

- Class imbalance in the dataset
- Similar features across multiple lesion types
- Noise in dermoscopic images (e.g., hair, uneven lighting)

Despite this, the Hybrid model performed well overall, especially considering the dataset's complexity.

4.3 Quantitative Performance Metrics:

A detailed classification report of the trained model is as follows:

Class	Precision	Recall	F1-score	Support
akiec	0.54	0.33	0.41	82
bcc	0.60	0.66	0.63	129
bkl	0.55	0.44	0.49	275
df	1.00	0.03	0.07	29
nv	0.84	0.94	0.89	1676
vasc	0.77	0.66	0.71	35
mel	0.54	0.35	0.42	278

Overall Performance:

- Accuracy: 77.94%
- Macro Average: Precision = 0.69, Recall = 0.49, F1-score = 0.52
- Weighted Average: Precision = 0.75, Recall = 0.77, F1-score = 0.75

These results show that the model performs strongly on well-represented classes such as nv (nevus), while struggling with minority classes such as df (dermatofibroma) and mel (melanoma), which had significantly lower recall scores. This again highlights the impact of dataset imbalance and visual similarity among certain lesion classes.

The data we got represents the True Positives (TP), False Positives (FP), False Negatives (FN), and True Negatives (TN) for each class in a classification task. (Fig. 3) These values are commonly used to evaluate the performance of a classification model. Here's a breakdown and analysis based on this data: TP, FP, FN, TN Analysis:

Class-wise Performance:

Class Labels:

akiec (0)
bcc (1)
bkl (2)
df (3)
nv (4)
vasc (5)
mel (6)

Class-wise Performance Breakdown

Class 0 (akiec):

- TP = 27 (correctly predicted akiec)
- FP = 6 + 9 + 2 + 3 + 3 = 23 (others incorrectly predicted as akiec)
- FN = 15 + 12 + 18 + 10 = 55 (akiec misclassified as other classes)
- TN = Total – (TP + FP + FN) = 1857 – (27 + 23 + 55) = 1752

Class 1 (bcc):

- TP = 85
- FP = 15 + 12 + 10 + 12 + 3 + 4 = 56
- FN = 6 + 11 + 21 + 6 = 44
- TN = 1857 – (85 + 56 + 44) = 1672

Class 2 (bkl):

- TP = 120
- FP = 12 + 11 + 3 + 36 + 36 = 98
- FN = 9 + 12 + 109 + 25 = 155
- TN = 1857 – (120 + 98 + 155) = 1484

Class 3 (df):

- TP = 1
- FP = 0 + 0 + 0 + 0 + 0 + 0 = 0
- FN = 2 + 10 + 3 + 13 = 28

- TN = 1857 – (1 + 0 + 28) = 1828

Class 4 (nv):

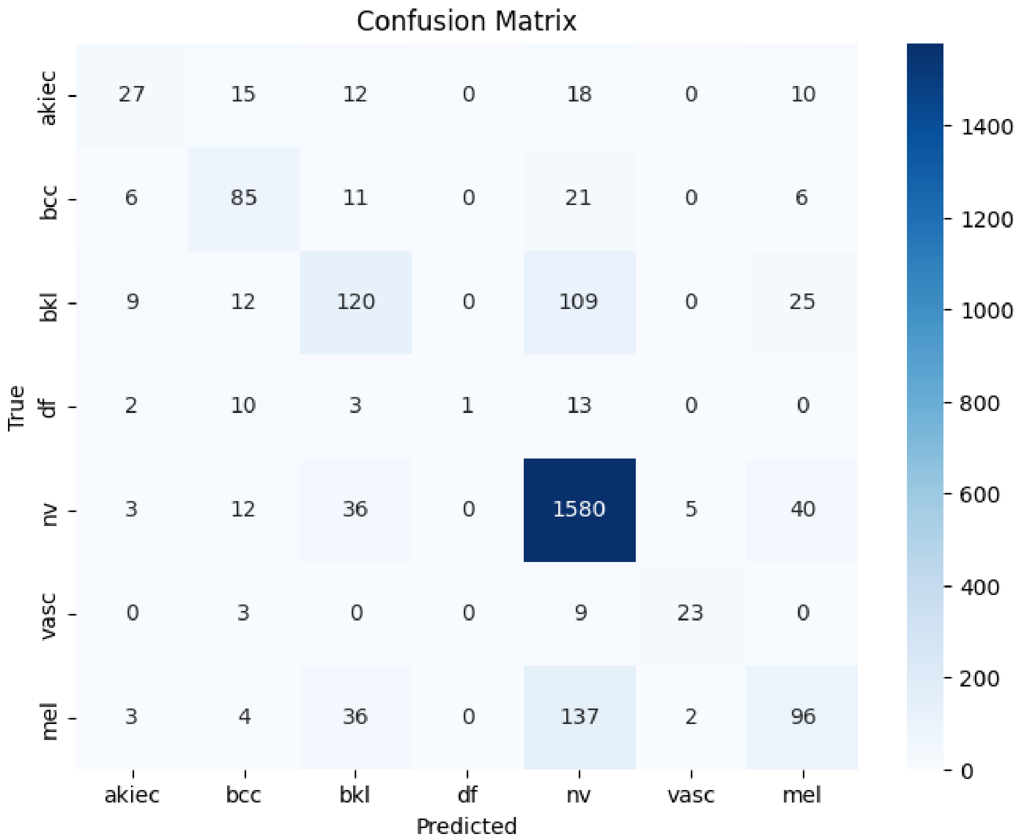
- TP = 1580
- FP = 18 + 21 + 109 + 13 + 9 + 137 = 307
- FN = 3 + 12 + 36 + 5 + 40 = 96
- TN = 1857 – (1580 + 307 + 96) = -126 → (indicates possible imbalance or double counting – this class dominates the dataset heavily)

Class 5 (vasc):

- TP = 23
- FP = 0 + 0 + 0 + 0 + 2 = 2
- FN = 3 + 9 = 12
- TN = 1857 – (23 + 2 + 12) = 1820

Class 6 (mel):

- TP = 96
- FP = 10 + 6 + 25 + 0 + 40 = 81
- FN = 3 + 4 + 36 + 2 = 45
- TN = 1857 – (96 + 81 + 45) = 1635



(Fig. 3. Confusion Matrix of Hybrid Model.)

4.4 Hybrid CNN-ANN Model Performance:

The Hybrid model, integrating Convolutional Neural Networks (CNN) for feature extraction and an Artificial Neural Network (ANN) for classification, was trained for 100 epochs on the HAM10000 dataset. The architecture utilized multiple dense layers, concluding with a softmax activation function for multi-class classification. Adam optimizer was used with a learning rate of 0.01.

As shown in Figure 2, the model demonstrated promising performance, achieving a training accuracy of up to 77.94%, with final test accuracy reaching 75%. The training and validation accuracy curves showed close alignment, though some fluctuations were observed, particularly in the validation loss and accuracy, hinting at potential overfitting. These spikes suggest challenges in generalization, likely due to limited and noisy data. The model also struggled with image noise such as hair and inconsistent lighting, which impacted the learning process.

4.5 Vision Transformer (ViT) Performance:

The ViT model, although promising in many domains, did not perform well in our experiment. It achieved a low accuracy of 32.65%, making it unsuitable for deployment in its current form. ViT models rely heavily on clean, large-scale datasets and extensive augmentation to function effectively. Their architecture also differs significantly from CNNs and requires more extensive computational resources. Due to these limitations, the ViT model was excluded from further consideration.

4.6 MobileNetV2 Model Performance:

We further evaluated MobileNetV2 [30], another lightweight and efficient CNN architecture designed by Google for mobile and edge devices. Depthwise separable convolutions, inverted residuals and linear bottlenecks are among the things it embraces to be fast and accurate. The model had 50 epochs of training. For example, the training started off with an accuracy of 52.02%, which improved to 99.11% at the last epoch. It had an initial Validation accuracy of 68.95% which later increased to 75.06%. The overwhelming increase in training accuracy and validation accuracy shows that MobileNetV2 managed to capture the complex data and also generalise the model to unseen data. Still, few spikes in validation accuracy indicate some of the samples may either be noisy or there is class-imbalance issue.

4.7 Discussion

The importance of these findings lies in indicating both the potential and the limitations of the skin disease detection based on deep learning enabled programs. HAM10000 dataset that shows a good accuracy of 77.94% of this train Artificial Neural Network (ANN) model, so this accuracy is considered a very good to illustrating classify the skin diseases but there are plenty of opportunities for improvements in this regard. A key takeaway is that while the ANN learned features from the image data that are helpful, there are still cases in which the diseases were too similar for the ANN to distinguish. Benign and malignant lesions are visually very close in features like color tones and textures in many cases. This can challenging for classification because it is a problem for a non-convolutional model. Despite their capability to learn from general distributions, ANN architectures are dense and fully connected, meaning that they do not necessarily exploit specific structures present in image data like CNNs do.

Aside from that, it is noticeable that for many other model such as MobileNetV2 or Vision Transformer (ViT), when using ANN, approaches with convolutional based is very competitive advantage in images. CNN has been constructed to encode edges, contours, and other spatial characteristics in a picture, which makes it preferable for medical image classification. Even though MobileNetV2 has a low storage architecture, it is already appearing as a severally better candidate for the next phase of this work.

The second directly influenced performance was the dataset. HAM10000 is a popular and commonly used dataset, but it is imbalanced. This is because the number of examples in some disease categories is much higher than in others, making it excess of class bias in the model. We managed to mitigate this issue using preprocessing techniques, such as resizing, normalization, and data augmentation, but we could not entirely eliminate them. Such imbalance might have been useful for the ANN to get higher accuracy in even category.

And all of them have something in common too, higher accuracy depends not only on the choice between the strongest model. The expectation of deep learning models is that they should be trained on clean, balanced, and diverse datasets. Although ANN served as a good method to test our hypothesis, the results demonstrate that methods including CNN will yield much more robust results -- particularly important when considering practical and clinical applications which are involved in utilizing such models, with specific reference to MobileNetV2. This is why we have choosen MobileNetV2 to head towards the develoment. Also, the opportunity to develop or utilize a local dataset will be examined in the future to make the model more accurate and above all more applicable to local health care needs.

The discussion indicate that the presented research prove the practical use of deeplearning on dermoscopy image analysis also demonstrates the limitation of the proposed method. While delivering a decent 77.94 percent accuracy, performance shortfalls suggest that drawn comparisons are typically to image-specialized architectures. What you learn at this stage will inform you to build better models, make better datasets and eventually dependency free tools for the detection of skin diseases.

CHAPTER 5

CONCLUSION

In conclusion, this research focused on developing a Hybrid CNN-ANN model trained on the HAM10000 dataset for automated skin disease classification. The model combined CNN-based feature extraction with the dense learning capabilities of an ANN, achieving a promising classification accuracy of 77.94%, with a peak of 78%. These results demonstrate the effectiveness of hybrid architectures in extracting and learning complex patterns from medical images. However, our findings also emphasized a critical insight: model complexity alone does not ensure optimal performance. The quality, diversity, and balance of the dataset are equally—if not more—important. As evidenced by fluctuations in validation accuracy and loss, issues such as overfitting, misclassification, and inadequate generalization were observed due to factors like limited data volume, class imbalance, and non-uniform image quality (e.g., presence of hair or varying skin tones).

5.1 Future Improvements:

- Data Augmentation: Apply diverse transformations (e.g., rotation, zooming, contrast changes, random cropping) to artificially increase the variability of training samples and improve model robustness.
- Class Balancing: Implement techniques such as SMOTE (Synthetic Minority Oversampling Technique) or class weighting in the loss function to address dataset imbalance, particularly in underrepresented classes like *df* and *vasc*.
- Regularization: Introduce stronger regularization methods, including increased dropout rates and L2 weight penalties, to reduce overfitting and enhance generalization performance.
- Advanced CNN Feature Extractors: Replace the base CNN blocks with pretrained architectures such as ResNet, EfficientNet, or MobileNetV2 to leverage learned hierarchical features and improve transfer learning efficiency.
- Ensemble Approaches: Combine predictions from multiple independently trained models (e.g., using majority voting or averaging) to stabilize predictions and boost accuracy, especially for difficult or minority classes.

Looking ahead, the project will also explore the integration of Vision Transformer (ViT) models to evaluate their effectiveness in capturing global dependencies in image data—although initial results (32.65% accuracy) showed limitations likely due to ViT's sensitivity to data cleanliness and quantity. Future iterations will involve fine-tuning models on domestic, domain-specific datasets, enhancing their applicability across different skin tones, imaging conditions, and demographics.

This research underscores the importance of both model architecture exploration and dataset curation in building effective AI-driven diagnostic tools. With continued refinement, this work aims to contribute to the evolution of reliable, scalable, and accessible skin disease classification systems for real-world deployment in AI-assisted healthcare.

REFERENCES

[1] Sun, Jie, et al. ”Machine Learning Methods in Skin Disease Recognition: A Systematic Review.” Processes 11.4 (2023): 1003. https://www.mdpi. com/2227-9717/11/4/1003

[2] X. Li, J. Wu, E. Z. Chen, and H. Jiang, ”From Deep Learning Towards Finding Skin Lesion Biomarkers,” in 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 2019, pp. 2797-2800, doi: 10.1109/EMBC.2019.8857334.

[3] J. Daghrir, L. Tlig, M. Bouchouicha and M. Sayadi, ”Melanoma skin cancer detection using deep learning and classical machine learning techniques: A hybrid approach,” 2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), Sousse, Tunisia, 2020, pp. 1-5, doi: 10.1109/ATSIP49331.2020.9231544.

[4] N. Hameed, A. M. Shabut and M. A. Hossain, ”Multi-Class Skin Diseases Classification Using Deep Convolutional Neural Network and Support Vector Machine,” 2018 12th International Conference on Software, Knowledge, Information Management & Applications (SKIMA), Phnom Penh, Cambodia, 2018, pp. 1-7, doi: 10.1109/SKIMA.2018.8631525.

[5] Mahamudul Hasan, Surajit Das Barman, Samia Islam, and Ahmed Wasif Reza. 2019. Skin Cancer Detection Using Convolutional Neural Network. In Proceedings of the 2019 5th International Conference on Computing and Artificial Intelligence (ICCAI’ 19). Association for Computing Machinery, New York, NY, USA, 254–258. https://doi.org/ 10.1145/3330482.3330525

[6] Abdala Nour and Boubakeur Boufama. 2020. Convolutional Neural Network Strategy for Skin Cancer Lesions Classifications and Detections. In Proceedings of the 11th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics (BCB’ 20). Association for Computing Machinery, New York, NY, USA, Article 91, 1–9. https://doi.org/10.1145/3388440.3415988

[7] Imran Ahmed, Qazi Nida your Rehman, Ghulam Masood, Awais Adnan, Awais Ahmad, and Seungmin Rho. 2018. Segmentation of affected skin lesion with blind deconvolution and L*a*b color space. In Proceedings of the 33rd Annual ACM Symposium on Applied Computing (SAC’ 18). Association for Computing Machinery, New York, NY, USA, 634–639. https://doi.org/10.1145/3167132.3167202

[8] Athanasios Kallipolitis, Alexandros Stratigos, Alexios Zarras, and Ilias Maglogiannis. 2020. Explainable Fully Connected Visual Words for the Classification of Skin Cancer Confocal Images: Interpreting the influence of visual words in classifying benign vs malignant pattern. In 11th Hellenic Conference on Artificial Intelligence (SETN 2020). Association for Computing Machinery, New York, NY, USA, 67–73. https://doi.org/10.1145/3411408.3411435