Comprehensive Skin Lesion Analysis System: Integrating Machine Learning and Customized CNN Models for Dermatological Conditions

BY

Israt Jahan Nipa
ID: 191-19-2070

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Electronics and Telecommunication Engineering

Supervised By

Professor Dr. A. K. M. Fazlul Haque
Professor
Department of ICE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

SEPTEMBER 2023

APPROVAL

This Project titled **"Comprehensive Skin Lesion Analysis System: Integrating Machine Learning and Customized CNN Models for Dermatological Conditions"**, submitted by* **Israt Jahan Nipa, ID no:191-19-2070***to the Department of Information and Communication Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Electronics and Telecommunication Engineering and approved as to its style and contents. The presentation has been held on *24th September,2023 *.

BOARD OF EXAMINERS

Md. Taslim Arefin — **Chairman**
Associate Professor & Head
Department of ICE
Faculty of Engineering
Daffodil International University

Prof. Dr. A.K.M. Fazlul Haque — **InternalExaminer**
Professor
Department of ICE
Faculty of Engineering
Daffodil International University

Ms. Tasnuva Ali — **InternalExaminer**
Assistant Professor
Department of ICE
Faculty of Engineering
Daffodil International University

(Prof.Dr.M.Quamruzzaman) — **ExternalExaminer**
Professor & Head
Department of EEE
Faculty of Engineering
World University of Bangladesh

DECLARATION

I hereby declare that, this project has been done by me under the supervision of **Professor Dr. A. K. M. Fazlul Haque, Professor, Department of ICE** Daffodil International University. I also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

____

Professor Dr. A. K. M. Fazlul Haque
Professor
Department of ICE
Daffodil International University

Submitted by:

Israt Jahan Nipa

____

Israt Jahan Nipa
ID: 191-19-2070
Department of ETE
Daffodil International University

ACKNOWLEDGEMENT

First, we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project successfully.

We really grateful and wish our profound our indebtedness to **Professor Dr. A. K. M. Fazlul Haque**, **Professor** Department of ICE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "*Deep Learning and Image Processing*" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior draft and correcting them at all stage have made it possible to complete this project.

We would like to express our heartiest gratitude to Md. Taslim Arefin, Associate Professor, and Head**,** Department of ICE, for his kind help to finish our project and also to other faculty member and the staff of ICE department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

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

ABSTRACT

This work aims to explore the potential of automated skin lesion analysis in the context of detecting cancerous lesion and proposes a thorough investigation of a analysis system that uses two unique approaches. The goal is to evaluate the efficacy of using a analysis system to accurately identify and diagnose cancerous lesions and to investigate the challenges associated with this technology. Among the various types of cancer, skin cancer is considered to be highly perilous. Early detection and treatment of it can lead to a high cure rate, as most skin cancers can be eliminated before they metastasize. In particular, this paper will focus on the challenges of multiclass classification, feature extraction, and classifier design for skin lesion analysis. This study's primary goal was to create a fairly accurate automated system for diagnosing skin lesions to aid in early diagnosis utilizing both machine learning & deep learning-based techniques. The proposed work has two different approaches, in the deep learning approach a customized CNN model has been employed as the classification model and in the machine learning approach a CNN model has been used as a feature extractor, those features were fed into six different machine learning classifiers. Some statistical measures such as confusion matrix specificity, sensitivity, accuracy, precision, recall was also generated. The HAM10000 dataset which has seven different type of skin lesion, served as a platform for evaluating the proposed methodology. Different pre-processing steps has been employed to balance the dataset. Both suggested frameworks operated effectively and it has been concluded that implementation of the Deep learning approach outperforms other machine learning methods in terms of accuracy. The Deep learning approach has an 80% accuracy rate, and the highest accuracy among all the 6-machine learning classifier applied is 77 % obtained by the Xgboost classifier.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

1.1 Introduction

Human body's largest organ, the skin, acts as a barrier between the internal organs and the external environment. It contributes significantly towards sensory perception, vitamin D synthesis, and immune defense. Skin lesions are abnormalities that occur on the skin, such as growths, bumps, discolorations, or wounds and are distinct from the adjacent tissues. These can be brought on by a variety of conditions, including underlying ailments, trauma, allergies, and infections. Skin lesion can appear on many regions of the body and can vary in size, shape, or texture. There are primary and secondary skin lesions, respectively. Primary skin lesions are abnormal skin conditions that can develop over a person's lifetime or be present at birth. Primary skin lesions that have been navigated or inflamed will lead to secondary skin lesions. While the majority of skin lesions are benign and painless, some of them are malignant and have the potential risk of transforming into skin cancer. Unrepaired deoxyribonucleic acid (DNA) in skin cells results in genetic flaws or mutations on the skin, which ultimately causes skin cancer[1]. Skin cancer is by definition a malignant skin lesion. While each type of skin cancer has own characteristics, typical indications of skin cancer can include quickly expanding skin lesions, changes in the color or size of an already-existing lesion, a sore that won't heal or one that does but then comes back, the underside of a nail may have a brown or black stripe. A person of any ethnicity or gender can develop skin cancer. Though some groups benefit from it more than others. However, certain groups experience it far more frequently than others such as non-Hispanic white people. It affects women more frequently than men before the age of 50. The prevalence of it increases in men beyond the age of 50 also in the united states approximately 9,500 people are diagnosed with skin cancer everyday[2].In individuals with darker skin tones, skin cancer is frequently discovered in its later stages, when it is more challenging to treat. Skin cancer can be either melanoma type or non-melanoma type. On-melanoma includes basal cell carcinoma and squamous cell carcinoma which rarely spreads to another part of the body. A medical practitioner may do a number of tests to diagnose cancer, including: Visual examination, dermatoscopy (a procedure that involves studying the lesion using a dermatoscope,). Imaging tests such a CT scan, MRI, or PET scan, as well as biopsies (removing a small fragment of the lesion for inspection).skin cancer has the tendency to frequently spread to other organs of the body and this phase becomes more deadly and less curable than the first stages. Therefore, if it is discovered early it has a better chance of being completely cured. Due apparent resemblance between several lesions, efficient identification of it is often vigorous even for healthcare providers and it if often time consuming to diagnose the cancer. It takes about 2 to 3 weeks to get the results of your biopsy which is a long period. This study will act as an aid of the healthcare professionals to help them distinguish different lesion and categorize them as either harmful and non-harmful. this project is saves time in diagnosis also serves the most accurate results.

1.2 Motivation

skin cancer usually beings on the cells of the skin and it carries a greater risk of becoming dangerous and damaging in the later stages. Large number of individuals around the globe are impacted by this common health concern. It happens when skin cells grow abnormally and is typically brought on by direct exposure to ultraviolet (UV) light from the sun or tanning beds. Skin cancer can be life threatening as it can cause mortality or even disfigurement of the affected body parts if it is neglected over a long period of time. When the cancer is in its early stages it if easier to treat and it also don't have the potential to spread to other parts of the body. That's why early detection of skin cancer is crucial for effective treatment and improved patient outcomes. The frequency of its incidence is growing rapidly. In 2020, globally an estimated of 57,000 people passed away from the illness[3]. As per the most recent WHO data published in 2020, skin cancer deaths in Bangladesh accounted for 0.12% of all fatalities[4]. As skin cancer rates continue to rise, a precise, efficient, and impactful detection method is becoming more and more necessary to help the healthcare official in diagnosing the disease. The presented method could have a significant influence on skin cancer detection's precision and speed, which would ease the strain on healthcare systems and boost treatment outcomes. This study has the potential to steadily enhance the field of medical image analysis and cancer diagnosis while simultaneously helping to save lives by applying the technology to a real-life issue. This project offers doctors a reliable tool to identify skin cancer earlier to improve patient outcomes.

1.3 Problem Statement

Potentially life-threatening diseases such as cancer, require quick and accurate detection to guarantee the right kind of treatments. In the field of skin diseases, a precise and early diagnosis can have a big impact on the results. However, it is still very difficult to diagnose patients with this level of accuracy and accurately classifying skin lesions into multiple categories remains a challenging task due to the similarity in appearance between different types of lesions. Traditional diagnostic techniques frequently depend on human evaluation, which might be subject to errors. In order to mitigate this, the main goal of our study is to develop a reliable skin lesion analysis system by implementing the power of artificial intelligence, particularly deep learning and machine learning. By creating a skin lesion analysis system, there has been an shown an indication to provide clinicians with reliable tools for diagnosis. This study presents two distinct strategies: one involves designing a custom deep learning model, while the other employs machine learning techniques alongside a convolutional neural network (CNN) as a feature extractor. By utilizing these methods, we hope to increase the precision and effectiveness of skin lesion diagnosis, which will ultimately lead to better patient outcomes and more accurate medical decisions.

1.4 Research Questions

- What is the degree of reliability of deep learning models when it comes to detecting skin cancer?"
- What is the functioning process of a deep learning model for detecting skin cancer?
- What are the potential challenges associated with skin lesion analysis for cancer detection?
- What are the potential applications of skin lesion analysis using neural machine learning models?
- How can the accuracy of skin lesion analysis using pre-trained network models as a features extractor be improved?
- How do the machine learning algorithms perform as a classifier?

1.5 Expected Outcome

The performance of pre-trained convolutional neural network models in identifying cancer and two other kinds of skin lesions is thoroughly analyzed for readers in this study. The purpose of the report is to compare the performance of the suggested approach to that of already used methods in order to show how well it performs in accurately recognizing different types of skin lesions, including melanoma. The research also intends to provide light on the shortcomings and potential areas for enhancement of the suggested approach. This report's main objective is to aid in the creation of more precise and effective cancer detection methods, with the ultimate goal of enhancing skin cancer patients' prognoses. Readers are able to evaluate the research's contributions impartially.

1.6 Report Layout

Chapter 1:

In order to frame the study, the introductory chapter provides baseline information and research context. This usually includes a summary of the research questions and issue statement that the study is intended to solve. The chapter also describes the goals and significance of the work, as well as how it adds to our understanding of machine learning's application to the analysis of skin lesions. The reader can grasp the purpose of the study and the possible consequences of the research findings by presenting this information.

Chapter 2:

Readers are given a thorough overview of the project topic in this chapter. It highlights the types and prevalence of skin cancer as well as an overview of the disease. Additionally, it summarizes earlier studies on the use of machine learning methods to the investigation of skin lesions and describes how convolutional neural networks are used to classify images. An overview of the pre-trained CNN models that have been applied to skin lesion analysis is given, and the gaps in the literature are critically analyzed, emphasizing the demand for the suggested study.

Chapter 3:

The methodology chapter gives readers a visual representation of the methodology and thorough explanation of the steps taken to conduct the study. In addition, this chapter discusses how the dataset was acquired and how it was applied in this study, which included dataset's pre-processing, image scaling, normalization. The CNN model's customization for the specific task, and the assessment metrics used for evaluating the model's effectiveness were also covered.

Chapter 4:

The outcomes of this project are presented in full to readers in the results chapter. This includes a presentation of the experimental findings, which cover the model's accuracy, sensitivity, and specificity and other evaluation metrics are also discussed. The findings of the models and their consequences are also discussed in the results chapter. The study's conclusions and their importance can be understood by readers in a clear and objective manner.

Chapter 5:

The main findings from the research and their significance are outlined in the discussion and conclusion chapter. It contains a review of the shortcomings and how they affected the findings. It explains how the it contributes to current knowledge of skin lesion analysis. This chapter provides readers with a comprehensive grasp of the projects contributions to the field of skin lesion analysis by offering an overview of the research findings and their consequences

CHAPTER 2
BACKGROUND STUDY

2.1 Preliminaries

In recent time, artificial intelligence's (AI) adverse effects on medical field has been truly transformational especially in the identification of disease. traditional methods are more susceptible to human error which can cause a disruptive effect in the diagnosis. They may have to deal with thousands of picture scan and test results each day, making the procedure time-consuming. Medical professionals will soon be able to get around these restrictions thanks to developments in AI and ML, opening the door to automated, reliable, and quick analysis. Artificial intelligence is able to evaluate massive amounts of data with astounding speed and precision, which enables early disease detection in serious diseases when recognizing the disease early is essential for the patient's survival. it can also have more personalize approach for treatment by analyzing medical history and other relevant data. Researchers are advancing into a new era of effective and precise diagnosis by utilizing the power of artificial intelligence (AI) and machine learning (ML) approaches. This chapter explores the challenges, work in the related field and far-reaching scope inherent to the problem of skin lesion analysis.

2.1.1 Different Types Of Skin Lesion

there are numerous types of skin lesions are present in real life which has their own characteristics and implications. we have introduced some most common types of skin lesions here with their medical terminology

- **Melanocytic Nevi/Moles:** Commonly referred to as moles, melanocytic nevi are benign growths that emerge from pigment-producing cells. They are typically brown, tan, or black and can be of any shape and size.

- **Melanoma:** A malignant type of skin cancer that originates in melanocytes. It often presents as an irregular, dark lesion and has the potential to spread and metasis which is the cause of most deaths.
- **Benign Keratosis:** These benign skin growths often resemble warts or rough patches. They can vary in color and texture, often appearing scaly or rough on the surface.

- **Dermatofibroma:** Benign skin nodules with a brownish hue are known as dermatofibromas. When pressed, they may feel hard or have dimples, and the specific reason why they occur is frequently unclear.
- **Vascular Lesions:** This category includes various vascular anomalies such as hemangiomas and angiokeratomas. They result from abnormal blood vessel development, leading to distinctive appearances like raised bumps or red marks.
- **Actinic Keratosis:** Also known as solar keratosis, these are precancerous growths triggered by sun exposure. They appear as rough, scaly patches and can range in color from pink to brown.
- **Basal Cell Carcinoma:** The most common type of skin cancer, basal cell carcinoma typically emerges in sun-exposed areas. It often presents as a pearly or waxy bump, sometimes with visible blood vessels.

2.1.2 Usual Detection Process

The detection of skin lesions involves a multi-step process that combines visual examination, dermatoscopy (dermoscopic imaging), and sometimes biopsy for accurate diagnosis. Dermatologists follow these steps:

- **Visual Inspection:** Dermatologists begin by visually examining the skin, looking for irregularities in color, shape, size, and texture.
- **Dermatoscopy:** Dermatoscopes, handheld devices with magnifying lenses and light sources, are used to closely examine lesions. This makes tiny details apparent that are invisible to the observer and helps with classification.
- **Pattern Analysis:** Dermatologists assess the patterns within a lesion, such as pigment network, globules, streaks, or vascular structures. These patterns help to identify the type of lesion.

- **Clinical Experience**: Dermatologists draw on their clinical experience to differentiate between benign and malignant lesions based on their appearance and patterns.
- **Biopsy (if necessary)**: For suspicious lesions, a biopsy may be performed. A small sample of the lesion is extracted for laboratory analysis, confirming the diagnosis.

2.2 Related Works

Melanoma , the most lethal type of skin cancer is detected using deep learning and fuzzy k-means clustering from dermoscopic images by (Nawaz et al., 2021)[5] . they tested their suggested methodology using the three often used datasets, ISBI-2016, ISIC-2017, and PH2.At first, preprocessing was done on the dataset images to improve the visual information by removing noise and illumination issues. Faster R-CNN which is an end-to-end single stage trained model, was employed in order to extract features of fixed length from the dataset images which were submitted for analysis. Subsequently, the above-mentioned extracted features were subjected to a complex and esoteric process of fuzzy k-means clustering, which later segmented the affected area of the skin. The proposed approach achieved accuracies of 95.40, 93.1, and 95.6% on the ISIC-2016, ISIC-2017, and PH2 datasets, respectively.

A dependable automated system for skin lesion analysis was pitched by (Li and Shen, 2018)[6] which was evaluated on the ISIC 2017 dataset .the images in the dataset are divided into three different class of skin lesion.A deep learning framework was proposed to address the three main problems of skin lesion image processing, namely lesion segmentation, dermoscopic feature extraction, and lesion classification. This framework consists of two fully convolutional residual networks (FCRN) Lesion segmentation and classification are both addressed simultaneously by the Lesion Indexing Network. It generates the results of segmentation and coarse classification. The task of dermoscopic feature extraction is handled by the Lesion Feature Network, the distance heat-map is created by a lesion index calculation unit (LICU), The AUC of the lesion indexing network exceeds the current deep learning systems for lesion segmentation and classification.

(Ningrum et al., 2021)[7] designed a low-resource artificial intelligence (AI) model for malignant melanoma identification utilizing dermoscopic pictures and patient metadata. The primary objective was to create a model and evaluate its effectiveness for binary categorization of malignant and nonmalignant melanomas. They have outlined an architecture that comprises two models: CNN and a CNN+ANN hybrid. Before feeding the data into the neural network, images were cropped as a part of preprocessing has been performed. MinMaxScaler was used for numerical variables (such as age) and one hot encoder for categorical factors (such as anatomy site, location, and gender).in the 2^{nd} stage different classification methods were used to make the correct decision. In this study the CNN model used the image data and the hybrid model used the patient's metadata along with dermoscopic images. The hybrid approach (CNN+ANN), can improve the classification's accuracy in detecting malignant melanoma.

In [8] authors proposed a gripping approach alongside the detection of melanoma .They used a CNN model for spotting the melanoma and at the same time the model performance was compared with a large group (58) of dermatologist.. They developed a 300-image test Which was from the Department of Dermatology at Heidelberg University in Germany's picture repository & it contains 20% melanoma images and 80% melanocytic nevi of various subtypes. Sensitivity, specificity, and area under the curve (AUC), (ROC) were the main criteria for evaluation for the diagnostic categorization. In the test set most dermatologists were surpassed by CNN. The deep learning algorithm attained a sensitivity of 95% and a specificity of 63.8%, compared to dermatologists' average sensitivity and specificity of 86.6% and 71.3%, respectively

(Pacheco & Krohling, 2020)[9] suggested employing deep learning models to identify skin cancer using patient-specific clinical data. With the help of the Federal University of Espirito Santo's Dermatological Assistant Program (PAD), they gathered a custom dataset which contains patient's clinical data along with the lesion images. Three different malignancies and illnesses were present among the eight different types of skin lesions in their sample. To merge the picture and clinical data, a number of models like Google Net, ResNet50/101, VGGNet13, and MobileNet were used. The findings showed that all models under investigation significantly improved their performance thanks to clinical aspects. Their framework contains a number of flaws, including its inability to handle missing data in the clinical characteristics and its confusion when attempting to differentiate between squamous cell carcinoma (SCC) and basal cell carcinoma (BCC) lesions due to their comparable clinical features.

(Shetty et al., 2022)[10] In their suggested research, the skin lesion photos were classified using CNN and machine learning approaches. The HAM1000 dataset served as the basis for the research which was carried out. Dataset images were preprocessed (scaled to 96 × 96) and augmented (Horizontal Flip augmentation) before feeding into the network. Several machine learning models such as Random Forest algorithm, logistic regression, support vector machine, k-nearest neighbor, naive Bayes, decision tree, Linear Discriminant analysis were used. According to the results, the customized CNN outperformed the suggested machine learning methods with an accuracy of 95.18 percent. The models were assessed according to their Accuracy, Precision, and Recall scores, as well as their F1-Score.

In (Jojoa Acosta et al., 2021)[11] the method outlined in this research begins with a stage where a bounding box is drawn around the skin lesion and the region of interest is automatically cropped using the Mask and Region-based Convolutional Neural Network methodology (Mask R_CNN). A ResNet152 structure is used in the subsequent stage to determine if lesions are "benign" or "malignant”. This study utilizes the ISIC 2017 database, which has over 2000 high resolution dermatoscopic images divided into three main categories: melanoma, nevus, and keratosis. The suggested model improves accuracy and balanced accuracy by 3.66% and 9.96% on the test data set, respectively.

(Cruz-Roa et al.,)[12] Employed a deep learning framework to differentiate between Basal cell Carcinoma(BCC) and normal tissue sequences on 1417 images extracted from 308 Region of Interest (ROI) images of skin histopathology. This detection method incorporates learning about picture representation, classifying images, and interpreting the results. They analyzed the use of feature descriptors, such as the bag of features, canonical wavelet transforms, and Haar-based wavelet transform, with the deep learning approach and machine learning techniques. With an F-Measure of 89.4% and a balanced accuracy of 91.4%, the deep learning framework outperformed standard techniques.

In (Singh et al., 2020)[13]the authors proposed a computer-aided design (CAD) system for early evaluation and diagnosis of benign or malignant skin lesions Using machine learning models. the ABCDE (“Asymmetrical, Border, Color, Diameter, Evolving) rule and the PH2 data set are used to diagnose melanoma, using SVM as a data mining and classification learning model. the segmentations include the use of adaptive thresholding. The ABCD rule requires certain characteristics to detect melanoma. Extraction of the skin lesions asymmetry, border, color, and diameter constitutes the main processing. The suggested SVM classifier then progressively analyzes the images to determine whether the lesion is a benign nevus or a malignant melanoma. The model's performance was demonstrated by obtaining sensitivity of 100%, specificity of 87.5%, and accuracy of 92.2%.

(Brinker et al., 2019)[14] tested the performance of the CNN model with a group of 145 dermatologists in the task of classifying clinical melanoma images. a convolutional neural network method (ResNet50) was trained with 12,378 open-source dermoscopic images. The test case of one hundred clinical skin lesion photos (MClass-ND) includes eighty cases of nevi and twenty cases of melanoma that have been confirmed by a biopsy. While the dermatologists were able to reach a total sensitivity of 89.4%, a specificity of 64.4%, and an AUROC of 0.769, their proposed approach was able to achieve the same sensitivity while also achieving a superior specificity score of 69.2%.

The primary objective of all the related studies has been discussed up to this point has been to identify and categorize skin-related diseases, distinguishing between benign and malignant skin lesions, thus making an effort to decrease the number of fatalities from skin cancer. They have explained how deep neural networks can help medical personnel and healthcare professionals distinguish between cancerous and non-cancerous lesions with accuracy.

2.3 Comparative Analysis & Summary

In this project, we aimed to demonstrate how machine learning and deep learning approaches compared in terms of multiclass disease identification. the dataset used contains a diverse collection of dermoscopic images of skin lesions which was thoroughly assembled and curated by a team of dermatologists, scientists, and machine learning specialists. different approaches were deployed for both ML & DL. according to the desired architecture data preprocessing has been done. a customizes CNN model was introduced as the DL approach. for the ML approach, a pre trained CNN model was used as a feature extractor a 4 different type of ML algorithm was deployed as a classifier. Several trial with different hyperperameter were done in order to get the highest accuracy. as the memory constrains was a big issue while deployed the model, several methods to ease the constrains were also used

Table 2.3: Comparative analysis of previous work

Reference	overview of the technique & classifier used	Dataset Used	Scope for the future development
using deep learning and fuzzy k-means clustering for skin cancer detection[5]	melanoma detection from dermoscopic images. Faster R_CNN used to extract features, fuzzy means clustering use to segment affected area	ISBI2016, ISIC-2017, and PH2	despite achieving high accuracy, it is not used in real time dues to time complexity which could be improved
Utilizes patients clinical information for computer aided diagnosis [9]	justifying skin cancer using patient clinical data. several pretrained model such as Google Net, ResNet50/101, VGGNet, and MobileNet were used	Custom dataset (images + Clinical Data)	development could involve accurately distinguishing between SCC & BCC.
A Convolutional neural network outperforming 145 dermatologist in a melanoma classification task[14]	they tested the performance of the CNN model with a group of 145 dermatologists in the task of classifying clinical melanoma images.	The HAM1000 dataset	Interpretability improving. Focusing on classifying other types of skin cancers and non-cancerous skin conditions

2.4 Scope of the Problem

The issue at stake is the precise categorization of various skin lesion types, a job that has been left to the knowledge of dermatologists. The downsides of this manual method include its time commitment, vulnerability to human mistake, and potential for inconsistent diagnoses. The goal of this study was to create an automated analysis system that can achieve diagnostic accuracy that is on par with or better than human analysts by focusing on a wide variety of seven different types of skin lesions. We wanted to improve the effectiveness and accuracy of skin lesion classification by utilizing both machine learning and deep learning methodologies, thereby increasing patient care through early and precise diagnosis.

2.5 Challenges

Despite being the highest potential of artificial intelligence there are always some challenges that’s need to be addressed. in this study we illustrated the supplication of both machine learning and deep learning approach. The main challenge of this study was to find the suitable balance of the number of images in different classes of the dataset. the dataset was highly imbalanced which is the cause of the bad performance of the model. We came across the interpretability problem of the CNN models which are often known as "black boxes," problem. we had to preprocess and balance the data to get a fair accuracy. we also had to try different combinations of hyperperameter to find the best results. We had to be experimental with the different layers of the CNN model also.

CHAPTER 3

Research Methodology

3.1 Introduction

The organized technique used to address the research queries and goals is described in the methodology chapter. Here is a detailed explanation of the methods used in this study. Readers will be able to visualize the suggested methodology by viewing a full illustration of our methodology of choice. This image highlights the evolution and activity flow while illustrating the crucial steps of the research process.

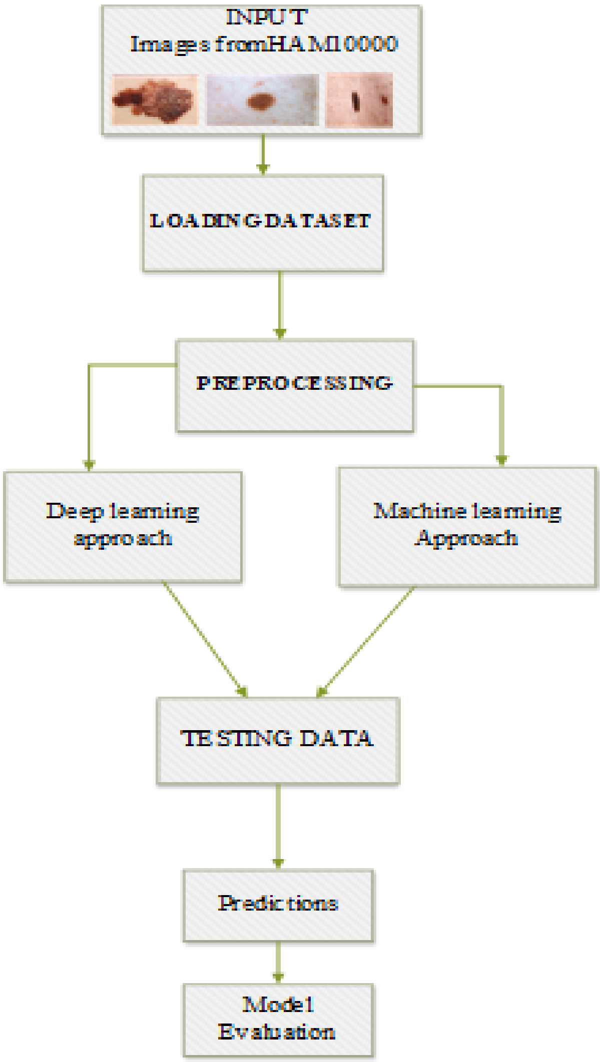

Figure3.1: Proposed Methodology

3.2 Subject of Research and Equipment

The subject of this study is the creation and application of a cutting-edge method for analyzing skin lesion analysis system. The project intends to develop a reliable and precise classification system for differentiating between seven different types of skin lesions by utilizing machine learning and deep learning approaches. The study makes use of the vast collection of dermatoscopic pictures known as the HAM10000 dataset, which is frequently used in research. The ultimate goal is to improve medical diagnosis and patient care by automating the classification of skin lesions, hence advancing dermatological analysis methods.

Here is a list of the instruments needed for this model.

- Windows 10
- Python
- Google drive
- Google Colab
- NumPy
- Matplotlib
- Sklearn library
- Tensor flow Keras
- Label Encoder
- Splitting function
- Classification Report
- Confusion matrix

3.3 Dataset Utilized

The dataset used in this study is HAM1000("Human Against Machine with 10000 training images")[15] which includes a varied collection of dermatoscopic images of skin lesions. This section gives a thorough summary of the HAM10000 dataset, covering its history, constitute up, annotation, and importance to the scientific community.

1. **Dataset Composition:** The data set consists of 10,015 dermatoscopic images, each of which has been thoroughly classified into one of seven different lesion classes:

Table 3.3: detailed no. of images in Dataset

Lesion Type	Lesion type	Number of Images
Melanocytic nevi (nv)	Benign	6705
Melanoma (mel)	3malignant	1113
Benign keratosis-like lesions(bkl)	Benign	1099
Basal cell carcinoma (bcc)	Malignant	514
Actinic keratoses (akiec)	Benign	327
Vascular lesions(vasc)	Benign	142
Dermatofibroma(df)	Benign	115
Total Images		10015

2. **Metadata and Annotation:** Each image has metadata stored alongside it, which provides details like the location on the body, sex, age, and more. In order to categorize and examine the skin lesions, machine learning models can be trained with this metadata, which adds extra context for the lesions. Histopathology (histo) is the primary method used to confirm lesions in more than 50% of the cases. The remaining cases rely on follow-up exams, expert consensus, or in-vivo confocal microscopy (follow-up, consensus, or confocal) to provide the final word. The lesion_id-column found in the HAM10000_metadata file allows users to monitor lesions that have many pictures.

3.4 Statistical Analysis

the dataset has a metadata file which stores several information like age, gender, cell type, and where it appeared on the body alongside the lesion_id and image_id. Exploratory data analysis was done here to visualize the dataset before performing any sorts of preprocessing and model training. the following statistics are based on the cell type, age, gender and the lesion localization on the body

1. Figure 3.4.1 visualizes the total cell count according to the cell/lesion types. Melanocytic nevi have the highest number of instances among the 10,015 images. Melanoma has the second highest instances.

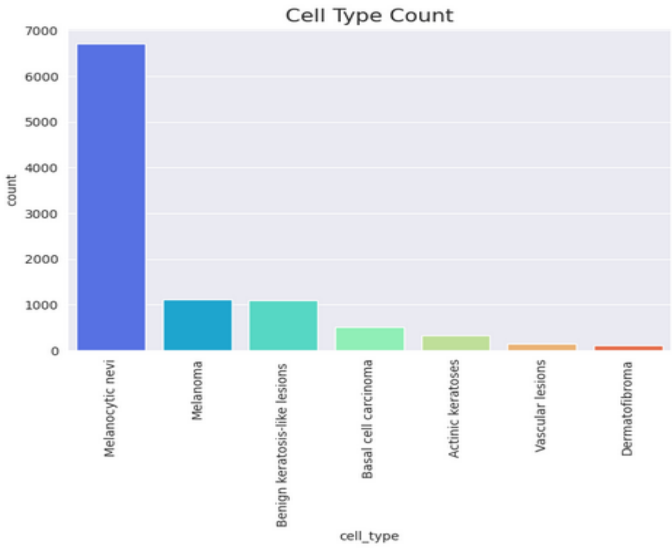

Figure 3.4.1 cell type count

2. Figure3.4.2 shows where the lesions are located on the patient's body.it lest the readers visualize the general site where the lesions mostly occur. The back and the lower extremity has the two most frequently occurred places of lesions.

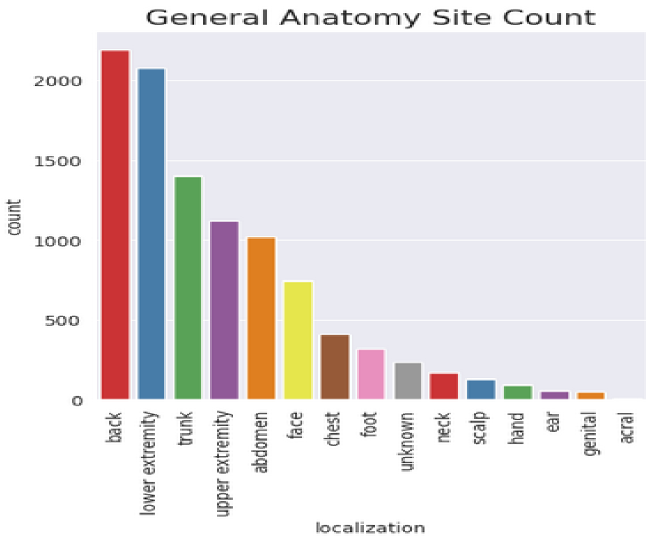

Figure 3.4.2 localization on the body

3. Figure 3.4.3 represents the age distribution according to the gender, which gives .an idea that women are more effect than man regardless of age by skin lesion. Patient from zero to ninety-five years. The highest number of cases has been recorded in the 40-55 age .

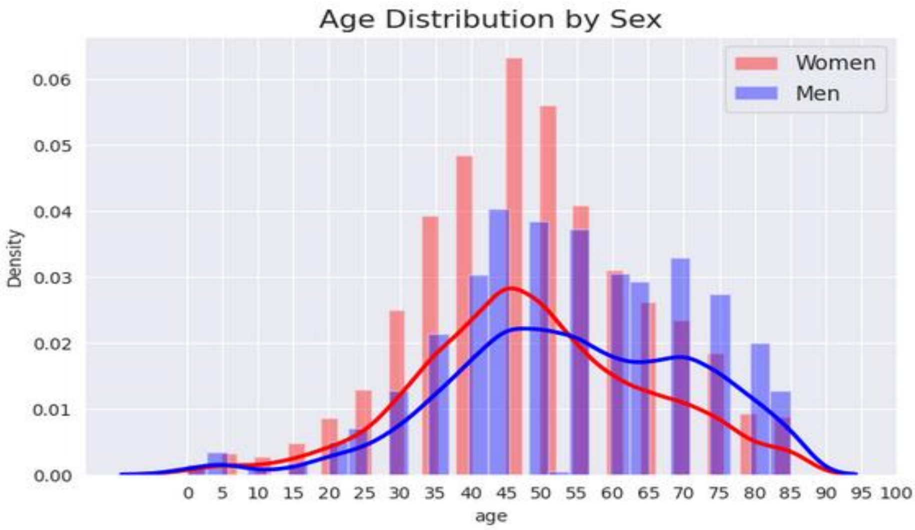

Figure 3.4.3 age distribution alongside gender

4. Figure 3.4.5 lets us visualize the how females tend to develop lesions at an earlier age than males among all the different types of lesions. In the seven discrete categories patients who are female, suffers more than the male patients

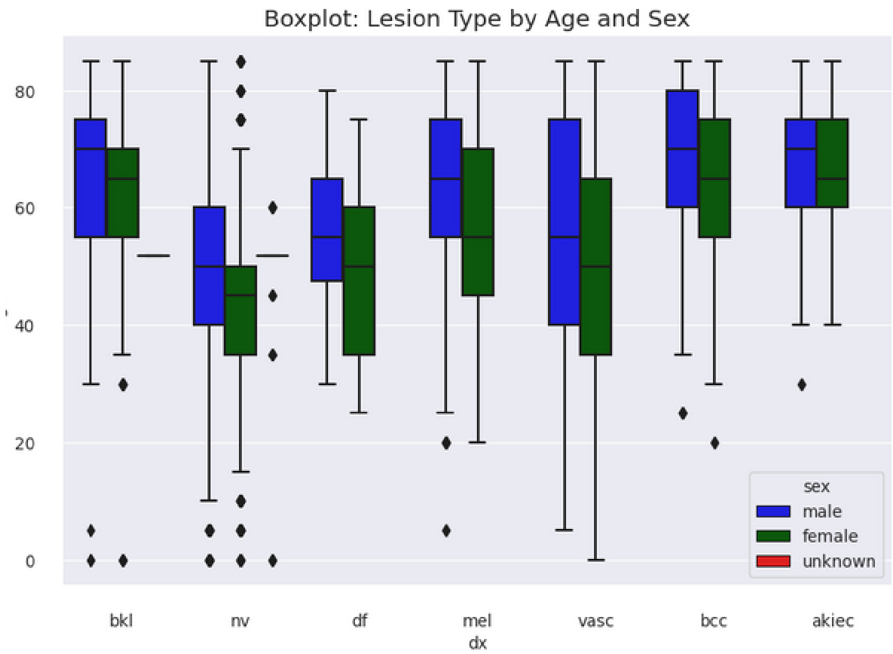

Figure 3.4.5 lesion type analysis by age and gender

3.5 Data Preprocessing

In order to ensure that raw data is in an appropriate format and quality for efficient machine learning or statistical processing, data preprocessing is a critical step in preparing the data for analysis and modeling. In order to make the data more useful for later activities, this frequently entails converting and cleaning it. The dataset has undergone several forms of preprocessing in the study. several steps such as label encoding, image resizing, balancing the dataset etc. was applied to turn the data into an manageable, and aligned format so that it could be used efficiently in training the model

3.5.1 Label Encoding

Label encoding is a technique where categorical labels or classes are converted into numeric format. Unique integers are assigned to each distinct categorical label as machine learning models and deep learning models can only understand and operate on numeric data it is crucial to convert any categorical labels to numeric format and to assign each unique categorical label with a unique integer. The lesion types /diagnosis classes (dx) were converted into numeric labels using the scikit-learn Label Encoder for the sole purpose of making classification model training easier. An identification number is given to each distinct diagnosis class.

table3.5.1 identification number / numerical value for each class

Diagnosis class	Identification number
Akiec	0
bcc	1
Bkl	2
Df	3
Mel	4
Nv	5
vasc	6

3.5.2 Data Balancing

Data balancing, also known as class balancing, is the process of distributing several classes within a dataset in an equitable manner. Balancing helps avoid biases during training in situations when some classes have noticeably more samples than others. Resampling, SMOTE, k-fold cross validation, and bagging classifier are just a few of the methods used to achieve a balanced dataset. The initial dataset is highly imbalanced where's among 1015 images 6705 belongs to one class named melanocytic nevus.to balance the data random oversampling technique was used to the minority class. For each diagnosis class, balanced datasets are created by randomly resampling instances from each class. By doing this, the model is less likely to be trained with a bias toward the dominant class.

Table 3.5.2: before and after resampling

Class identification number	no of images before resampling
5	6705
4	1113
2	1099
1	514
0	327
6	142
3	115

Class identification number	No of images after resampling
0	500
1	500
2	500
3	500
4	500
5	500
6	500

3.5.3 Image Reading and Resizing

Image resizing involves making adjustment to an image's width and height while keeping its original composition. This is often necessary when images have different sizes and need to be standardized for processing in machine learning models. Same dimension for all the images is crucial for feeding them into a training model .at the same time models run faster on smaller size of images. the original image has the size of 450 × 600 pixel. the images were read using the file paths provided in the CSV file. a dictionary of image paths with their respective image IDs were constructed. These paths are then used to read the images, and each image is resized to a common size of 32x32 pixels using the "Image. Open" function. Resizing ensures that all images have the same dimensions for compatibility with neural network architectures.

3.5.4 Normalization

Normalization refers to the method of scaling data to more normal or regular .it is often used in order to increase convergence during training and avoid features from dictating the learning process due to differing scales. pixel values are frequently normalized to the range [0, 1] by multiplying each value by the highest number. In the study the images were first loaded and resized and then it was normalized to range by dividing the pixel values by 255. It is a usual practice in deep learning to aid convergence during training.

3.5.5 One-Hot Encoding

one-hot encoding in machine learning involves transforming structured information into a format that can be used by machine learning algorithms to boost the accuracy of predictions. To train a multi-class classification model, where each label is represented by a binary vector, this is technique is often required. Each categorical label in the dataset is converted into a binary vector with a single class represented by each member. All other items were set to 0, except for the element designating the class, which is set to 1. One-hot encoding was used to encode the labels for each image. The “to_categorical” function converts the numeric labels acquired from the label encoding step into category arrays.

3.5.6. Dataset Splitting

Dataset splitting involves dividing a dataset into two distinct subsets: training, testing. The testing set evaluates how effectively the model generalizes to new data, whereas the training set is used to train the model. Carefully splitting the dataset ensures that our model is robust and performs well in real-world scenarios. The scikit-learn “train_test_split” function was used to divide the preprocessed dataset into training and testing sets. We divided the dataset into 75% of training data and 25% of testing data. this is crucial to assess the model's performance on hypothetical data. The training set consists of a portion of the preprocessed images and their corresponding categorical labels, while the testing set contains the remaining images and labels

3.5.7 Data Separation and Combination

The process of combining or concatenating various datasets into a single, cohesive dataset is referred to as data combination.as the dataset was highly imbalance, it was resampled using random oversampling, after balancing balanced subsets of different classes are combined to create a single balanced dataset that can be used for training a model. At the Initial phase, distinct data frames (df_0, df_1, etc.) were created for the data of each class. After balancing the data all the balanced data frames were concatenated together to create the one particular data frame which contains the balanced dataset

3.6 Implementation Requirements

a customized convolutional neural network model with several different layer was used a deep learning approach. For the machine learning approach images features were first extracted using a pre-trained neural network model and then several machine learning algorithms were used as a classifier

3.6.1 Deep Learning Approach

A customized CNN model is proposed for this study purpose of classifying each lesion correctly. the Keras Sequential API was used in this. Each layer in this sequential model, which consists of several interconnected ones, contributes to the procedure. Convolutional Layer 1, which uses 32 filters with a 2x2 kernel size and ReLU activation to capture low-level patterns, is the first layer in the architecture. Activations are then stabilized by a batch normalization layer, and overfitting is reduced by a dropout layer with a 20% dropout rate. Following Convolutional Layer 1, ReLU activation is combined with 64 filters and a 2x2 kernel to continue feature extraction. Again, a 50% Dropout Layer improves regularization while Batch Normalization improves training. For the completely connected layers, the Flatten Layer resizes extracted features into a 1D vector. With 64 units and ReLU activation, the first Dense Layer further abstracts learnt features. Generalization is aided by a Dropout Layer with a 50% dropout rate. Finally, the second Dense Layer, serving as the output layer, comprises 7 units for the multiclass classification task. A probability score is generated for each class by the Softmax Activation, allowing for precise predictions.

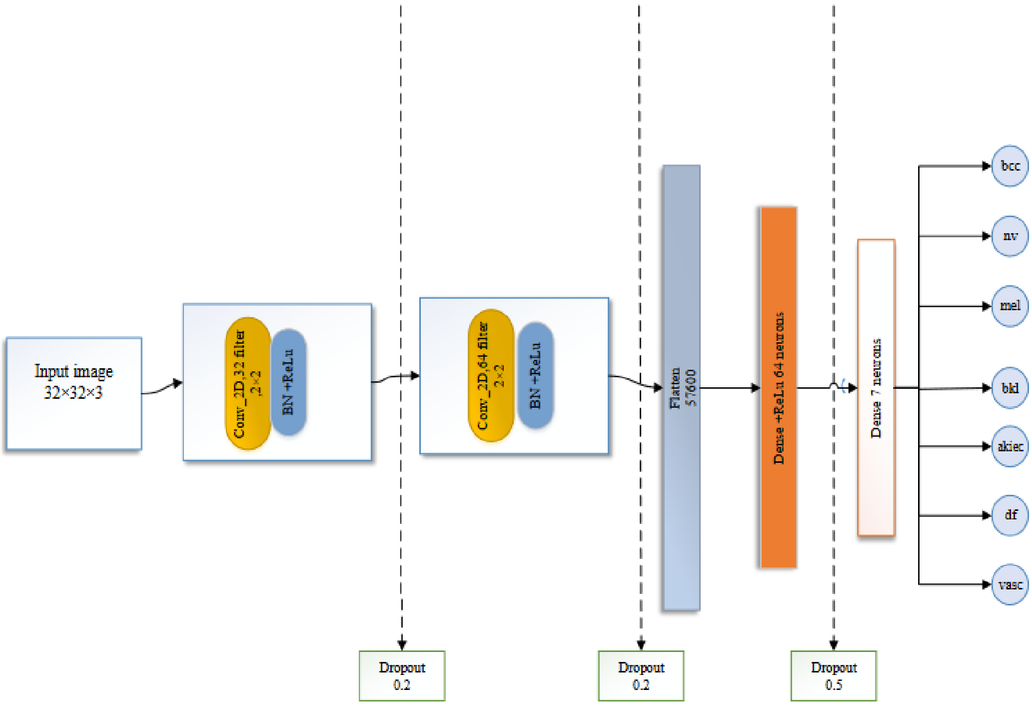

Figure3.6.1: proposed deep learning approach

3.6.1(a) Different layers

Layers which are used in the proposed CNN model is described below:

1. **The input layer**: input layer specifies how the neural network's input data will be shaped. the Conv2D layer that is inserted as the initial layer to the model implicitly defines the input layer. The input shape is set to (32, 32, 3), where 32 stands for the input photos' file sizes and 3 denotes the RGB color space's three-color channels.
2. **Convolutional Layers (Conv2D)**: Two Conv2D layers are used. The first layer applies 32 filters with a kernel size of 2x2, using the ReLU activation function. This layer processes the input images, extracting local features through

   convolutional operations. The second Conv2D layer follows a similar pattern with 64 filters, enhancing feature abstraction.

3. **Batch Normalization Layers**: Two Batch Normalization layers are inserted after each Conv2D layer. Activations from the previous layer are normalized

via batch normalization, which reduces internal covariate shifts and stabilizes training while enhancing convergence. As a result, training becomes more efficient and stable.

4. **Dropout Layers**: Dropout layers are employed twice in the model. A dropout rate of 0.2 is applied after the first Conv2D layer, and a higher dropout rate of 0.5 follows the second Conv2D layer and the subsequent Dense layer. Dropout randomly deactivates a fraction of neurons during training, which mitigates overfitting by enhancing model's generalization.
5. **Flatten Layer:** The Flatten layer is used to reshape the output from the convolutional layers into a 1D vector. This prepares the data for the fully connected (Dense) layers that follow, effectively transitioning from convolutional feature maps to a format suitable for classification.
6. **Dense Layers:** Two Dense layers are employed. The first Dense layer consists of 64 units with the ReLU activation function. The convolutional layers' learned characteristics are further integrated in this layer. The second Dense layer has 7 units (equal to the number of classes) with a softmax activation function, enabling the network to produce probability scores for each class.

3.6.2 Machine learning approach

as the machine learning approach, a comprehensive workflow that combines a convolutional neural network as a feature extractor with various classifiers to achieve accurate predictions has been designed .at first the data pre-processing is done to ensure the correct format of the data. The pre-processed data is then sent into the CNN models architecture, a potent feature extractor known for its effectiveness in extracting complex patterns from images. The extracted features are then directed towards multiple classifier models, namely Support Vector Machines (SVM), k-Nearest Neighbors (KNN), Random Forest (RF), and XGBoost. Each classifier harnesses the rich feature representations from the model to make informed predictions. These predictions are then aggregated, and evaluation metrics such as accuracy, F1-score, and others are computed to gauge the performance of each classifier. The holistic diagram of this workflow elegantly encapsulates the journey from data pre-processing to classification and assessment, emphasizing the symbiotic integration of convolutional neural networks feature extraction prowess with the diverse strengths of the chosen classifiers.

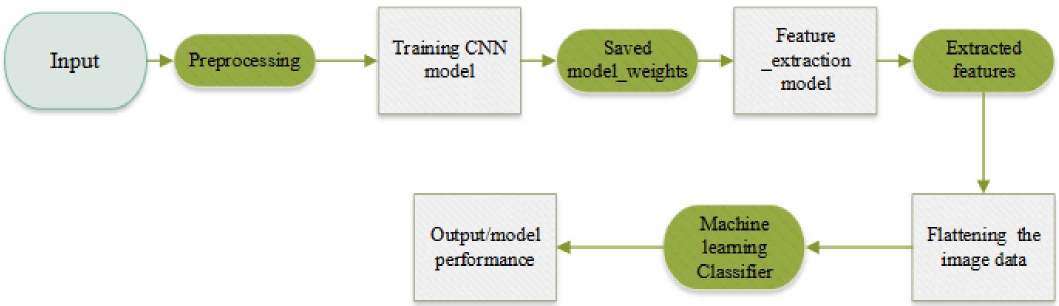

Figure 3.6.2 workflow of the ml approach

3.6.3 Feature Extraction

In the machine learning approach, a particular method was used to explicitly extract the features from the images. They aid in extracting pertinent data from images and producing meaningful structures that machine learning algorithms can utilize to classify images. There are several ways of extracting features while using traditional machine learning algorithm such as Histogram of Oriented Gradients (HOG), color histograms, local binary patterns, Gabor filters, color moments, edge detection, texture features, principal component analysis, SIFT (Scale-Invariant Feature Transform). They are often known as traditional feature extraction method. Convolutional Neural Networks (CNNs) are an additional effective feature extraction technique for classifying images. they automatically extract from raw pixel data the hierarchical features. since CNNs are made up of many layers. Through a succession of convolutions and non-linear activation functions, they automatically learn features from the original image pixels. CNNs are excellent at identifying small-scale patterns, edges, and textures in images, which makes them a excellent option for image classification tasks. As the dataset we are using is a moderately large dataset, with complex features in the raw images present, as customized CNN model has been used as a feature extractor to leverage the power of learning features automatically of the convolutional neural networks. they can capture complex features from the raw data which reduces the need of manual feature engineering. a customized CNN model consisting of different layers named two convolutional layer first one with 32 filters and second one with 64 filters, batch normalization layer, dropout layer, dense layer, flatten layer and SoftMax activation function has been used as the core model. After compiling the model with categorical_crossentropy' loss and 'Adam' optimizer, the weights of the model have been saved. now the initialization of the main model starts, the weights of the previous trained model has been loaded and a new model has been created which takes the same input as the original model and outputs the activations of the layer just before the final Flatten layer. The new model extracts the feature representation of the images up to the flatten layer, after the feature extraction the extracted data for both training and testing has been reshaped to a flattened format which prepares the data This prepares the data as a input into a machine learning classifier. In this manner, CNN model's learnt features have been used to enhance the performance of a following machine learning classifier.

3.6.4 Support Vector machine

A Support Vector Machine is a powerful supervised learning algorithm used for classification and regression tasks. SVM operates by identifying a hyperplane that optimally separates different classes of data points in feature space while maximizing the margin between classes. This hyperplane is determined by selecting support vectors, which are data points lying closest to the decision boundary. The objective of SVM is to find the hyperplane that maximizes the margin between support vectors while minimizing classification error. Mathematically, for a linearly separable case, the decision boundary can be represented as:

$$wTx+b=0wTx+b=0 \quad [16]$$

where ww is the weight vector and bb is the bias term.

3.6.5 k –nearest neighbor

k-Nearest Neighbors is a simple and intuitive classification algorithm. Given a new data point, KNN classifies it by considering the classes of its kk nearest neighbors in the feature space. The class with the highest frequency among these kk neighbors becomes the predicted class for the new data point. KNN is based on the assumption that data points of the same class are often close to each other. The algorithm operates effectively when the appropriate value of kk is chosen, balancing between overfitting (small kk) and underfitting (large kk).

Mathematically, the predicted class CC for a new point can be defined as:

$$C = argmaxc \sum_{i=1}^{k} [\mathrm{y}i = \mathrm{c}] \qquad [17]$$

where yi is the class label of the ith neighbor.

3.6.6 Random Forest

Random Forest is an ensemble learning technique that constructs multiple decision trees during training and combines their outputs for improved classification performance. Each decision tree in the forest is built on a random subset of the training data and features. During classification, the class predicted by each tree is counted, and the class with the majority vote across all trees is chosen as the final prediction.

This ensemble approach helps to mitigate overfitting and enhance the model's generalization. Mathematically, the prediction of the Random Forest can be represented as c

$$C = \mathrm{argmax}_{\mathrm{c}} \sum_{i=1}^{N} [Ti\,(x) = c \qquad [18]$$

where Ti(x) is the class predicted by the ith decision tree.

3.6.7 XGBoost

XGBoost (Extreme Gradient Boosting) is a boosting algorithm that combines the predictions of multiple weak learners (usually decision trees) to create a strong learner. It works by iteratively adding new trees that correct the errors made by the previous trees. XGBoost optimizes a loss function using gradient descent to minimize the errors in predictions. The algorithm's strength lies in its regularization techniques, handling missing values, and scalability. Mathematically, the prediction of XGBoost can be defined as:

$$F\,(x) = \sum_{m=1}^{M} \gamma_{\mathrm{m}}\, \mathrm{h}_{\mathrm{m}}\,(\mathrm{x}) \qquad [19]$$

where *F(x)* the final prediction, γm is the weight assigned to the *mth* tree, and *hm(x)* is the prediction of the *m*th tree.

3.6.8 Decision Tree

Decision Tree is a supervised machine learning algorithm used for both classification and regression tasks. In order to optimize the homogeneity of the target variable within each segment, it divides the feature space into segments by recursively splitting the data based on the most informative characteristics. The model structure resembles a tree, with leaf nodes representing projected outcomes and core nodes representing decisions based on attributes. At each node, features and thresholds are chosen using parameters like mean squared error, entropy, and Gini impurity. The goal of the splitting procedure is to increase information gain or decrease impurity[20].

3.6.9 Logistic Regression

Logistic Regression is a widely used binary classification algorithm that models the probability of an instance belonging to a particular class. Utilizing the logistic function, it translates the linear combination of input features to provide anticipated estimations. To forecast binary classes, these probabilities are then thresholded**.**

$$\boldsymbol{Logit(p) = log\left(1 - \frac{p}{pp}\right) = \beta 0 + \beta 1 + \beta 2 + \cdots \ldots + \beta n X n}$$

Where *p* is the probability of the positive class, x1,x2,…,xnx1,x2,…,xn are input features, and *β0,β1,…,βnβ0,β1,…,βn* are coefficients to be learned [21].

3.7 Model parameters & hyper parameters

The internal variables that the model learns during training in order to make predictions are referred to as model parameters. These variables reflect the connections and trends found in the training data. On the other hand, hyperparameters are external setups and settings that are made before the training phase starts. Despite not being acquired through the training data itself, they have an impact on how the model adapts and generalizes. They are picked by the data scientist or machine learning engineer. model parameters are the internal values that the model learns from the data, while hyperparameters are external settings that govern how the learning process occurs

Table 3.7.1 displays the Parameters of the proposed DL approach:

Table 3.7.1: proposed DL approach parameters

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 31 ,31 ,32)	416
batch_normalization	(None, 31, 31, 32)	128
dropout (Dropout)	(None ,31 ,31 ,32)	0
conv2d_1 (Conv2D)	(None ,30 ,30 ,64)	8256
batch_normalization_1	(None ,30, 30, 64)	256
dropout_1 (Dropout)	(None ,30 ,30 ,64)	0
flatten (Flatten)	(None ,57600)	0
dense (Dense)	(None ,64)	3686464
dropout_2 (Dropout)	(None ,64)	0
dense_1 (Dense)	(None ,7)	455
Total params: 3,695,975		
Trainable params: 3,695,783		
Non-trainable params: 192		

For The ML approach the model which was used as a feature extractor has some different parameters too which are displayed below:

table :3.7.2: ML approach parameters

Layer (type)	Output Shape	Param #
conv2d_input (Input Layer)	(None, 32, 32 ,3)]	0
conv2d (Conv2D)	(None, 31, 31 ,32)	416
batch normalization (Batch normalization)	(None ,31 ,31, 32)	128
dropout (Dropout)	(None, 31, 31 ,32)	0
conv2d_1 (Conv2D)	(None ,30 ,30 ,64)	8256
batch_normalization_1 (Batch normalization)	(None ,30 ,30 ,64)	256
Total params: 9,056		
Trainable params: 8,864		
Non-trainable params: 192		

A few standard hyperparameter settings are employed to provide a better model evaluation. Table highlights the values of the hyperparameters used in the CNN model. Through experimentation, it was discovered that 50 epochs produced a model with little loss and no overfitting to the training set. The loss value was set to "categorical cross-entropy" because the study was focused on a multiclass classification problem.
Table .3.7.3. visualizes the hyperparameters of the DL approach

Table 3.7.3: hyper parameters of Deep learning approach

HYPPERPATAMETERS	**VALUES**
Learning Rate**:**	0.0005
Batch Size	16
Optimizer**:**	Adam
loss	Categorical_crossentropy
Epoch	50

Table 3.7.4 lists the hyperparameters that were employed by the machine learning algorithms.

Table 3.7.4 hyperparameter of machine learning classifier

HYPERPARAMETERS	**VALUES**
SVM	Kernel = ’linear’, c = 1,random_state = 0
KNN	n_neighbors = 5
Xgboost	lr(learning rate)=0.1, n_estimators=100, max_depth=5
RF	n_estimators = 200, random_state = 42
LR	random_state = 0
DT	random_state = 0

3.8 Evaluation & Predictions

model evaluation and predictions are primary conception of both machine learning and deep learning. They involve analyzing a model's capacity for generalization and applying the model to provide precise estimations based on fresh data. Model evaluation is the procedure of determining how well a deep learning or machine learning model performs on raw data.it uses different metrics to helps us analyze the performance of the model and generalize it for pattern for more precise predictions. The purpose of model evaluation is to identify the model's advantages, disadvantages, and future growth areas. Prediction essentially includes programming computer algorithms to gain knowledge from the past and then apply that understanding to make predictions or anticipate future events regarding the results of fresh, unforeseen data points. The model employs the patterns it learnt during training to produce forecasts for unknown input data.

- When evaluating a machine learning model, several performance metrics including accuracy, precision, recall, F1-score, and ROC curves are calculated as well as strategies like dividing the dataset into training and testing sets. Predictions often include employing a trained model to forecast a target variable based on input information (e.g., regression or classification).

- Due to the complexity of deep neural networks, model evaluation in deep learning frequently uses methods similar to those used in machine learning, but with extra considerations. The division of data into training, validation, and testing sets, the observation of the training and validation loss curves, the use of measures like categorical cross-entropy and accuracy, and the use of early stopping procedures to avoid overfitting are common practices. Forecasts entail making predictions on fresh, unforeseen data using a trained neural network.

3.9 Evaluation metrics

1) **Classification Report:** A classification report is a thorough assessment of how well a model performed in a classification task. It offers significant measures for each class, including F1-score, recall (sensitivity), accuracy, and precision. The categorization report aids in your understanding of the advantages and disadvantages of the model's predictions by offering information on how well it performs across various classes.

- The percentage of overall accurate predictions is known as accuracy.
- Precision measures the model's capacity to reliably distinguish between positive cases among those it identifies as such.
- On the other hand, recall quantifies the percentage of real positive cases that the model accurately predicted.
- The harmonic mean of precision and recall, the F1-score, strikes a balance between erroneous positives and false negatives. Since it provides information on how well the model performs for each class, the classification report is very helpful when dealing with multi-class classification issues.

2) **Sensitivity:** also known as recall or the true positive rate, quantifies the model's ability to correctly identify positive instances out of all actual positive instances in the dataset. It is particularly important in circumstances when there is a high financial cost to missing positive cases, like in the diagnosis of illnesses or the identification of fraud. A high sensitivity implies a lower rate of false negatives, ensuring that positive cases are not overlooked.

3) **Specificity**: also referred to as the true negative rate, measures the model's capability to correctly identify negative instances out of all actual negative instances. When precisely identifying negatives is important, like in the case of disease screening tests, specificity is advantageous. A high specificity indicates a lower rate of false positives, ensuring that negative cases are correctly identified

4) **Confusion Matrix**: In a classification problem, a confusion matrix is a table that lists a model's predictions and actual results for each class. False positives, true positives, true negatives, and false negatives are all included. This matrix offers thorough insights into the model's performance, emphasizing areas where it excels and those where it needs work.

CHAPTER 4

EXPERIMENT RESULTS AND DISCUSSION

4.1 Experimental Setup

Images processing is an emerging field which is creating a huge significance in the medical sphere. Ai-driven image analysis is helping the healthcare workers to diagnose disease more swiftly and accurately. as a result, early detection in life threatening cases are also becoming low each day. Machine learning models make predictions based on the probabilities and patterns they've discovered during training. While high probabilities can indicate significant confidence, due to the inherent inconsistency and complexity of real-world data, models rarely guarantee 100% assurance. The dataset in question was highly imbalanced as majority of the images belonged to one particular the class. Before being fed to the model, the dataset has been pre-processed and balanced. The desired predictions were obtained using both machine learning and deep learning models. Several tools, including Tensorflow, Sklearn, Numpy, Pandas, Matplotlib, and Seaborn, were used to carry out the experiment on Google Colab. several machine learning algorithms such as SVM, Random Forest, XGBoost has been used as a classifier after explicitly extracting the features using a pre trained model. at the same time a customized CNN model was used as a deep learning approach.

4.2Experimental Results & Analysis:

Even though there are several types of skin lesion, because of their distinctive characteristics Doctors, particularly dermatologists, may find it difficult to correctly diagnose skin lesions through physical examination alone. Recent advances in technology like machine learning and artificial intelligence have the potential to help dermatologists with diagnosis. Large datasets of skin lesion photos can be used to train computer vision models, which can then be used to spot patterns and traits that may be challenging to spot by visual inspection alone.to help aid the healthcare professionals we deployed a machine learning and deep learning model to help categorize different type of lesion precisely. Additionally, Tansorflow and Matplotlib were utilized to graphically illustrate different data accuracy. The model is sufficient to deliver respectable accuracy.

Name	Abbreviation
Logistic Regression	LR
Decision Tree	DT
Random Forest	RF
K-Nearest Neighbor	KNN
Support Vector Machine	SVM
Machine learning	ML
Deep learning	DL
Convolutional Neural network	CNN

4.3 Training Outcome

1. Deep learning model:

this customized model was run till 50 epochs. Model Check pointing is used as Callbacks to save the model's weights during training at certain intervals which is useful to capture the best model achieved so far. Batch size 16 and learning rate 0.00005 was used. The accuracy and loss curve of training and validation of proposed Deep learning approach is given below

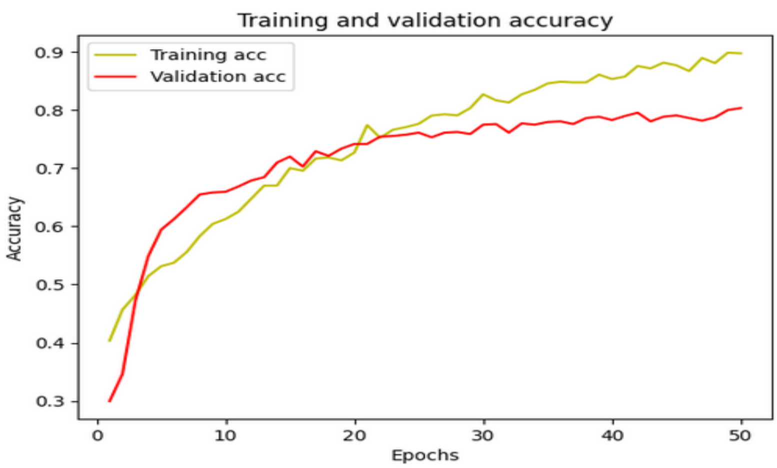

Figure.4.3.1: Training & Validation Accuracy

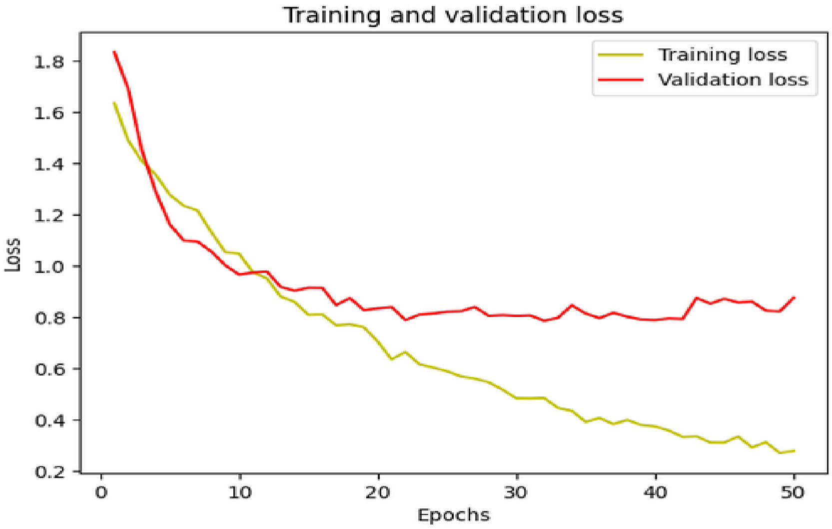

Figure.4.3.2: Training & Validation Loss

Fig 4.3.3 displays the accuracy and loss of the DL approach:

```
28/28 [==============================] - 1s 33ms/step - loss: 0.8754 - acc: 0.8034
Test loss: 0.8754022717475891
Test accuracy: 0.803428590297699
```

Fig 4.3.3: accuracy and loss of DL approach

Test accuracy of DL approach: 80%

Training accuracy for the **ML classifiers** are shown in the below table:

Table :4.3.1 training accuracy for ML classifier

Classifier Name	Accuracy	Accuracy (%)
Support vector machine	0.744	74%
K-nearest neighbor	0.464	46%
Decision Tree	0.73142855	73%
Random Forest	0.632	63%
Logistic Regression	0.531428	53%
XGBoost	0.7794285	77%

4.4 Confusion matrix

The confusion matrix is a representation of the accuracy and outcomes metrics of the algorithm or classifier in the deep learning classification process.

1. The CNN model correctly predicted 114 instances of akiec correctly while it misclassified only a few numbers of images as other classes. the highest performance was shown in the df class where 127 instances were classified correctly and only 1 was misclassify as bkl. And the lowest performance was seen in the mel class, where it predicted a lot of other instances as mel class instance. The confusion matrix for the deep learning approach is shown below figure

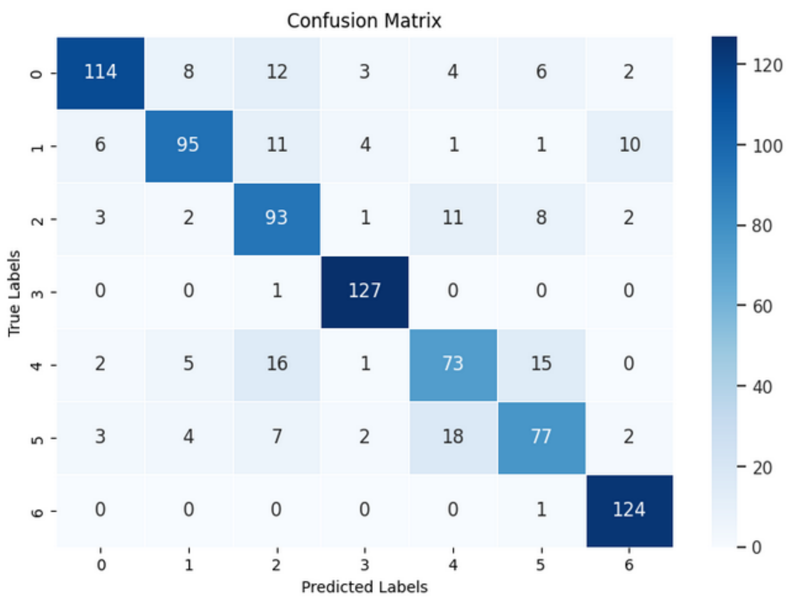

Figure.4.4.1: Confusion matrix of Dl approach

The confusion matrix for ML classifiers is shown below.

K-Nearest Neighbors demonstrated strong performance in other categories while accurately recognizing akiec and df occurrences. It correctly identified 103 cases as akiec, misclassified 10 as bcc, 12 as bkl, and other instances into different classes. 40 cases were correctly predicted as bkl, 40 incorrectly classed as akiec, 4 incorrectly labeled as bcc, and others were incorrectly placed into various classifications. Correctly predicted 102 occurrences as df, incorrectly classified 10 instances as akiec, 6 instances as bcc, and other instances into other groups. Correctly forecasted 99 cases as vasc, incorrectly classified 15 as akiec, 2 as bcc, and others into different classes. The model occasionally misclassified bkl and mel and showed some difficulty distinguishing between the two. Despite these difficulties, KNN offers a trustworthy method for classifying skin lesions, particularly when proximity-based approaches are needed. Figure 4.4.2 gives us the visual representation of the k-nearest neighbor classifier

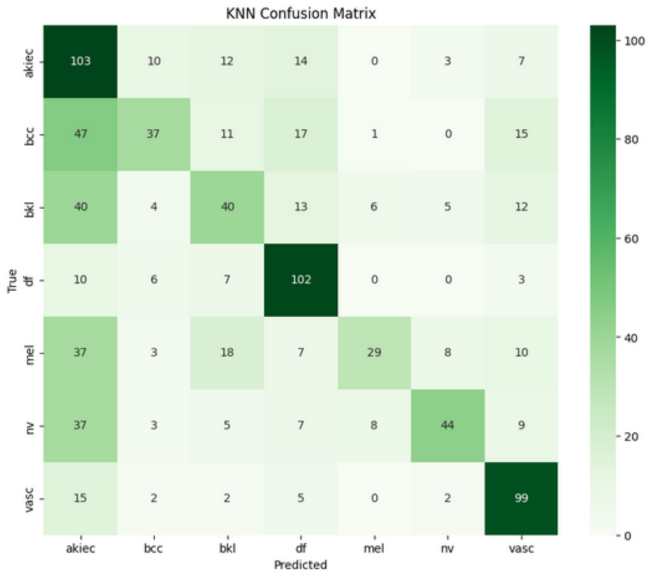

Figure4.4.2: confusion matrix of KNN classifier

The **SVM classifier** maintained competitive performance in other categories while doing exceptionally well at correctly detecting akiec and df instances. It accurately predicted 111 cases as akiec, incorrectly classified 11 as bcc, 7 as bkl, and other instances into different groups. Misclassified 2 instances as bkl, 2 instances as mel, and other instances into different classes while correctly predicting 123 instances as df. 9 occurrences was misclassified as akiec, 5 instances as bcc, and other instances into different classifications while correctly predicting 66 instances as nv. Vasc was a problem for the model; it would occasionally label it as bcc or mel. SVM, however, provides strong classification abilities, notably in important classes like akiec and df.

Figure4.4.3 visualizes the confusion matrix for support vector machine classifier:

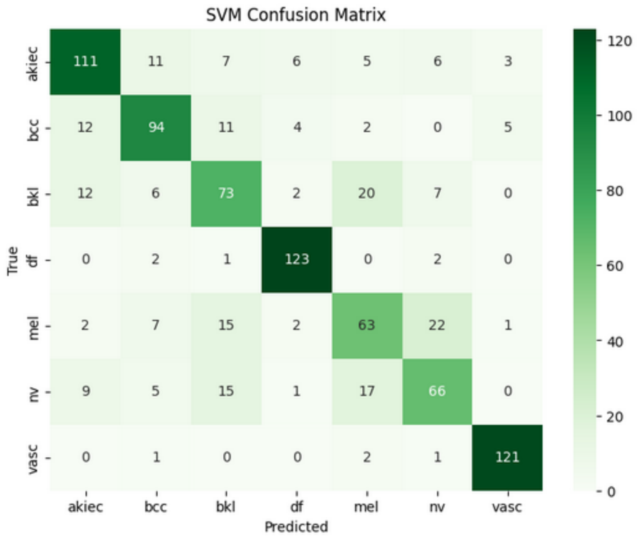

Figure 4.4.3 confusion matrix of support vector machine

Logistic Regression performed well in distinguishing df and vasc cases and showed moderate success with akiec and bcc. It misclassified 15 occurrences as akiec, 9 instances as bcc, and others into different groups while correctly predicting 57 instances as bkl.82 cases were correctly predicted as df, whereas the remaining examples—6 as akiec, 22 as bcc, and others—were incorrectly classified.101 cases were correctly predicted as vasc, whereas 15 were misclassified as bcc, 2 as mel, and 4 were placed in different classes. it faced challenges with differentiating between bkl and mel, leading to misclassifications in these categories. Overall, Logistic Regression provides a straightforward approach to skin lesion classification with promising results in specific classes. Figure 4.4.4 visualizes the confusion matrix for logistic regression classifier

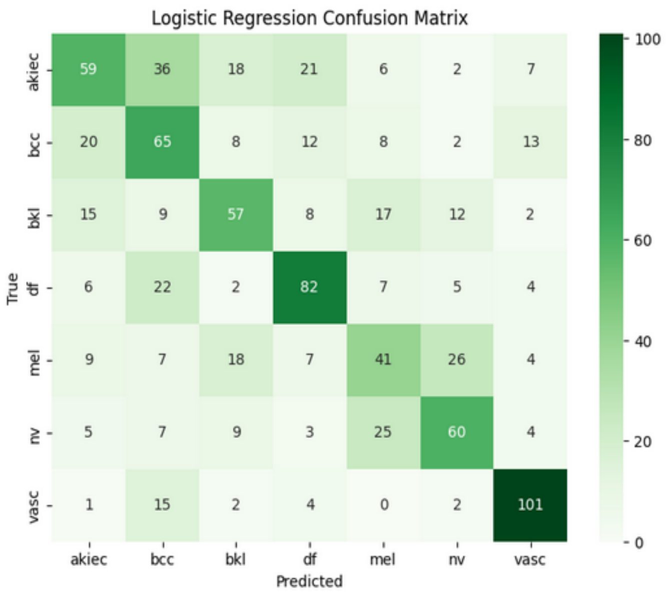

Fig 4.4.4: confusion matrix of logistic regression

A balanced performance across several classes was displayed by the Decision Tree classifier. It correctly identified 110 occurrences as akiec, misclassified 13 as bcc, 11 as bkl, and other instances into different classes. Bcc was correctly predicted in 77 cases, while akiec and bkl were incorrectly assigned to 8 and 17, respectively. 126 cases were correctly forecasted as df, incorrectly classified 2 instances as bkl, and classified others into different classifications. It misclassified 8 occurrences as akiec, 4 instances as bcc, and others into other groups while correctly predicting 71 instances as mel.2 instance was misclassified as bcc and 1 as mel, but correctly identified 121 occurrences as vasc. It correctly categorized akiec, bcc, and df instances but had some issues with mel and vasc. Misclassifications occurred as a result of the model's difficulty distinguishing between these two classesFigure 4.4.5 exhibits the confusion matrix for Decision Tree classifier:

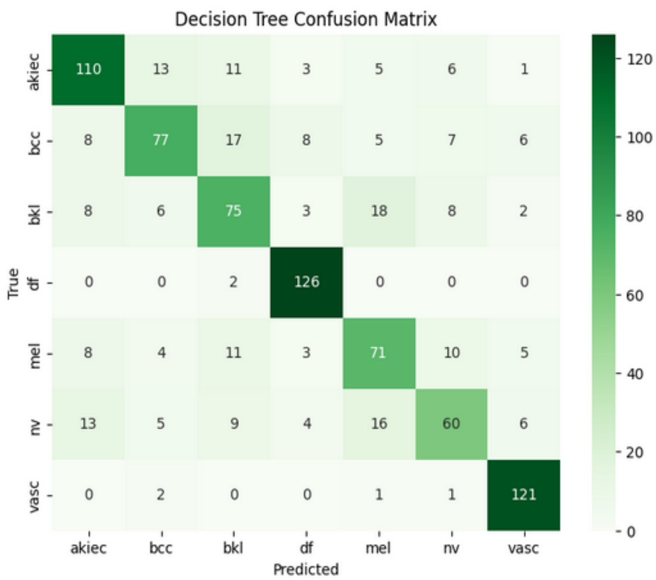

Fig4.4.5: confusion matrix of Decision Tree

The XGBoost classifier demonstrated efficient efficiency in the categorization of skin lesions. It correctly recognized akiec and df situations and performed admirably compared to other classes. It Predicted 106 cases as akiec correctly, but incorrectly classified 14 as bcc, 18 as bkl, and other instances into other classes. Where 87 instances of bkl were correctly predicted, whereas 2 and 4 examples were incorrectly categorized as akiec and bcc, respectively. At the same time, it Predicted 72 instances of mel correctly, misclassified 7 as akiec, 5 as bkl, and other instances into different classes. With only scattered misclassifications in these categories, the model had little trouble discriminating between bkl and mel. Overall, XGBoost provides an effective ensemble-based strategy for classifying skin lesions, especially in important classes like akiec and df. Figure 4.4.6 displays the confusion matrix for XGboost classifier:

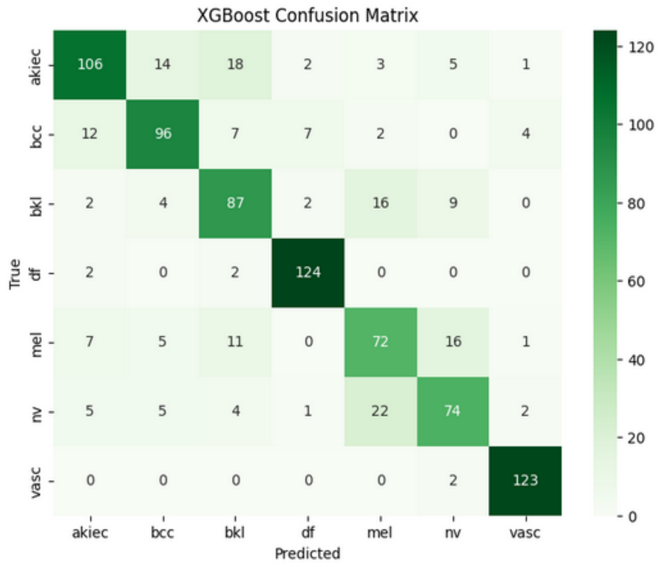

Fig 4.4.6: confusion matrix of XGBoost classifier

The Random Forest classifier performed satisfactorily. It correctly detected cases of akiec, bcc, and df while making few categorization errors. akiec has been predicted correctly 143 times. There are 0 instances of akiec that were misclassified. For bcc, 66 instances have been predicted correctly (True Positives). There are 60 instances of bcc that were misclassified as akiec, 2 as bkl, and 4 as mel. And as For df, 124 instances have been predicted correctly (True Positives). There are 66 mel instances that were incorrectly labeled as akiec and 5 as nv. For nv, 51 cases were correctly predicted. In vasc, 118 cases were successfully predicted. There are 7 vasc instances that were misclassified as akiec. For bkl, 54 cases were properly predicted. it struggled to distinguish between bcc and bkl as well as between bkl and mel. Overall, the Random Forest model performed admirably, especially in important areas like akiec and df.
Figure 4.4.7 exhibits the confusion matrix for Random Forest classifier

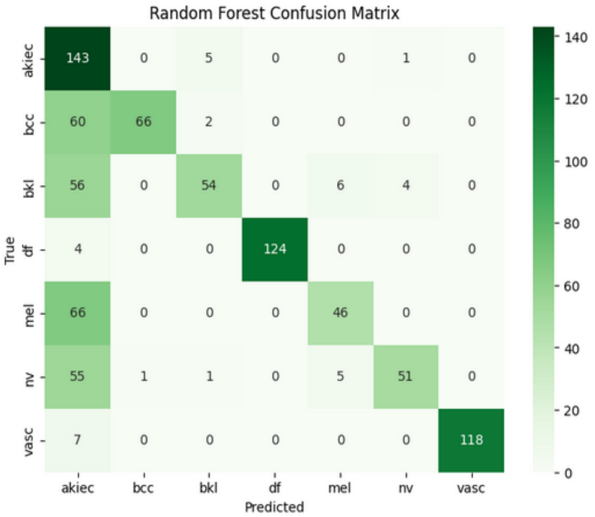

Figure 4.4.7: confusion matrix of Random Forest

4.5 Classification Report:

Classification report for the **Deep learning approach** is shown in the table 4.5.1 with an overall accuracy of 80%. For the majority of classes, it displays a balance between precision and recall, underscoring its aptitude for correctly identifying and categorizing various skin disorders. In particular, it does exceptionally well in "df" and "vasc" situations, earning high precision and recall scores that lead to remarkable F1-scores. The classifier retains a strong overall performance despite some classes, like "mel" and "nv," having slightly lower F1-scores than others.

Table 4.5.1classification report of the DL approach.

Lesion type	Precision	recall	f1-score	support
akiec	0.89	0.77	0.82	149
bcc	0.83	0.74	0.79	128
bkl	0.66	0.78	0.72	120
df	0.92	0.99	0.95	128
mel	0.68	0.65	0.67	112
nv	0.71	0.68	0.70	113
vasc	0.89	0.99	0.94	125
macro avg	0.80	0.80	0.80	875
weighted avg	0.81	0.80	0.80	875
Accuracy			0.80	875

- We have used 6 different ML classifiers as the **ML approach**. The Classification report of the ML classifiers are exhibited below. Table 4.5.2 shows the classification report of **K-nearest neighbors' classifier**. The accuracy was 52%. Its low precision, recall, and F1-scores for the majority of classes, however, showed that it had trouble correctly classifying skin conditions. 'akiec,' 'bcc,' 'bkl,''mel,' and 'nv' categories fared badly, but 'df' and 'vasc' examples showed some success.

Table 4.5.2 classification report of K-nearest neighbors' classifier

Lesion type	precision	recall	f1-score	support
akiec	0.36	0.69	0.47	149
bcc	0.57	0.29	0.38	128
bkl	0.42	0.33	0.37	120
df	0.62	0.80	0.70	128
mel	0.66	0.26	0.37	112
nv	0.71	0.39	0.50	113
vasc	0.64	0.79	0.71	125
macro avg	0.57	0.51	0.50	875
weighted avg	0.56	0.52	0.50	875
Accuracy			0.52	875

Table 4.5.3 shows the classification report of **Support vector machine classifier**:
The SVM classifier demonstrated equivalent balanced precision and recall scores across classes and obtained an accuracy of 74%. It performed well for the "df" and "vasc" examples, outperforming the Decision Tree classifier in terms of accuracy. It could have done a better job of recognizing 'akiec' and'mel' situations, though.

Table 4.5.3classification report of Support vector machine classifier

Lesion type	precision	recall	f1-score	support
akiec	0.76	0.74	0.75	149
bcc	0.75	0.73	0.74	128
bkl	0.60	0.61	0.60	120
df	0.89	0.96	0.92	128
mel	0.58	0.56	0.57	112
nv	0.63	0.58	0.61	113
vasc	0.93	0.97	0.95	125
macro avg	0.73	0.74	0.74	875
weighted avg	0.74	0.74	0.74	875
accuracy			0.74	875

Table 4.5.4 shows the classification report of **Logistic Regression classifier**: A 53% accuracy rate was attained by the logistic regression classifier, which also showed balanced precision and recall scores. It did ok, but there was space for improvement, as seen by its moderate F1-scores. In comparison to other classes, it produced better results for the 'df,' 'nv,' and 'vasc' situations.

Table 4.5.4 classification report of Logistic Regression classifier

Lesion type	precision	recall	f1-score	support
akiec	0.51	0.40	0.45	149
bcc	0.40	0.51	0.45	128
bkl	0.50	0.47	0.49	120
df	0.60	0.64	0.62	128
mel	0.39	0.37	0.38	112
nv	0.55	0.53	0.54	113
vasc	0.75	0.81	0.78	125
macro avg	0.53	0.53	0.53	875
weighted avg	0.53	0.53	0.53	875
accuracy			0.53	875

Table 4.5.5 shows the classification report of **Decision Tree classifier**. 73% accuracy was demonstrated with the Decision Tree classifier. Across classes, it generally showed balanced precision and recall scores, with better performance on "df" and "vasc" examples. There is still space for improvement, especially for the classifications of "akiec" and "bkl," even though it produced overall F1-scores that were better than those of the Random Forest classifier.

Table 4.5.5 classification report of Decision Tree classifier

Lesion type	precision	recall	f1-score	support
akiec	0.75	0.74	0.74	149
bcc	0.72	0.60	0.66	128
bkl	0.60	0.62	0.61	120
df	0.86	0.98	0.92	128
mel	0.61	0.63	0.62	112
nv	0.65	0.53	0.59	113
vasc	0.86	0.97	0.91	125
macro avg	0.72	0.73	0.72	875
weighted avg	0.73	0.73	0.73	875
accuracy			0.73	875

Table 4.5.6 shows the classification report of **XGBoost classifier**. With an accuracy of 78%, the XGBoost classifier outperformed other models. It achieved strong F1-scores and showed balanced precision and recall scores across classes, excelling especially in the 'df,' 'akiec,' 'bcc,' and 'vasc' classifications. This classifier performed admirably overall.

Table 4.5.6 classification report of **XGBoost** classifier

Lesion type	precision	recall	f1-score	support
akiec	0.79	0.71	0.75	149
bcc	0.77	0.75	0.76	128
bkl	0.67	0.72	0.70	120
df	0.91	0.97	0.94	128
mel	0.63	0.64	0.63	112
nv	0.70	0.65	0.68	113
vasc	0.94	0.98	0.96	125
macro avg	0.77	0.78	0.77	875
weighted avg	0.78	0.78	0.78	875
accuracy			0.78	875

Table 4.5.7 shows the classification report of **Random Forest classifier**. With considerable differences in precision and recall across various skin disease classes, the Random Forest classifier attained an accuracy of 69%. It did well at classifying situations as "df" and "vasc," but had trouble doing so for "akiec" and "mel." With space for improvement in recall and precision for some classes, it had a moderate F1-score overall.

Table 4.5.7 classification report of Random Forest classifier

Lesion Type	precision	recall	f1-score	support
akiec	0.37	0.96	0.53	149
bcc	0.99	0.52	0.68	128
bkl	0.87	0.45	0.59	120
df	1.00	0.97	0.98	128
mel	0.81	0.41	0.54	112
nv	0.91	0.45	0.60	113
vasc	1.00	0.94	0.97	125
macro avg	0.85	0.67	0.70	875
weighted avg	0.84	0.69	0.70	875
accuracy			0.69	875

Fig 4.5.8 displays the visualization which summarizes the important metrics (accuracy, precision, recall, and F1-score) from the six classification reports of the 6 ML classifier used. Performance of six different machine learning classifiers was assessed in the analytic system, including Logistic Regression, Support Vector Machine (SVM), k-Nearest Neighbors (KNN), XGBoost, Decision Tree, and Random Forest. The accuracy with which each classifier was able to identify skin lesions in photos from seven different classifications was thoroughly evaluated. The precision metric, varied from 36% for KNN to 79% for XGBoost. SVM had the highest recall (74%), but KNN showed the lowest recall (69%). The precision and recall balancing F1-scores ranged from 37% for KNN to 98% for Random Forest. Additionally, the accuracy rates for all classifiers ranged from 52% for KNN to 78% for XGBoost. These in-depth analyses offer insightful information about the classifiers' propensity to accurately categorize skin lesions and

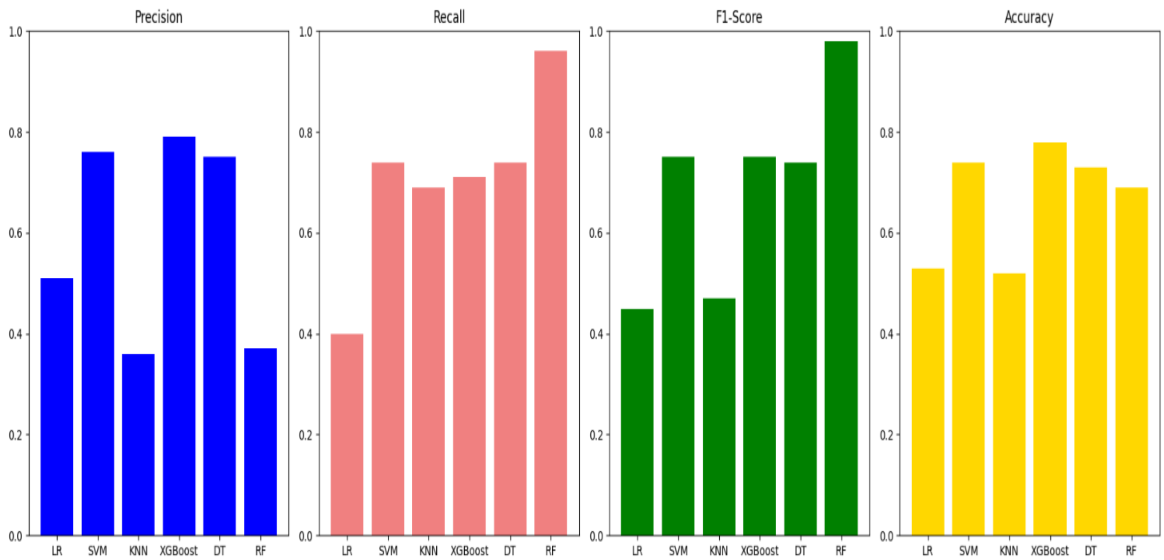

Figure 4.5.8: summary of the classification report of all ML classifier

4.6 Sensitivity & specificity

Sensitivity describes how effectively a model can identify positive cases. While specificity shows how accurately the model can classify negative cases.

The deep learning approach performs differently in each class, showing a range of performance. It achieves excellent specificity in a number of classes, including 'df' (0.98), 'akiec' (0.98), and 'vasc' (0.97), demonstrating its capacity to accurately identify situations in which these characteristics are absent. While there is space for growth in some classes, like "mel" (0.65) and "nv," others, like "df" (0.99) and "vasc," display excellent performance in reliably detecting affirmative cases. With a specificity of 0.93 and sensitivity of 0.78, the model does a respectable job at recognizing 'bkl' cases. The report's findings generally imply that the model's performance varies depending on the type of skin diseases, and that additional fine-tuning may be required to obtain a more appropriate trade-off between specificity and sensitivity for all classes. The score for the **DL approach** is shown below:

table 4.6.1: sensitivity and specificity score of DL approach

Class: akiec	
Specificity:	0.98
Sensitivity (Recall):	0.77
Class: bcc	
Specificity:	0.97
Sensitivity (Recall):	0.74
Class: bkl	
Specificity:	0.93
Sensitivity (Recall):	0.78
Class: df	
Specificity:	0.98
Sensitivity (Recall):	0.99
Class: mel	
Specificity:	0.95
Sensitivity (Recall):	0.65
Class: nv	
Specificity:	0.95
Sensitivity (Recall):	0.68
Class: vasc	
Specificity:	0.97
Sensitivity (Recall):	0.99

“Sensitivity and specificity score for all the **ML classifier** below:

Specifically, for "akiec," "df," and "vasc," Random Forest tends to produce superior sensitivity when compared to Logistic Regression. For other classes, including "akiec" and "mel," it also has poorer specificity, which suggests a larger incidence of false positives. On the other hand, Logistic Regression maintains moderate sensitivity across most classes and typically higher specificity, making it a reasonable option for reducing false positives. When sensitivity is a top priority, Random Forest is a good choice, while Logistic Regression is preferable when sensitivity and specificity must be balanced. For the intended trade-off to be realized, additional model fine-tuning may be required. Logistic Regression offers a balanced trade-off between sensitivity and specificity.

Table 4.6.2 sensitivity, specificity for (a) DT &(b) RF

Class: akiec	
Sensitivity:	0.9597
Specificity:	0.9945
Class: bcc	
Sensitivity:	0.5156
Specificity:	0.9451
Class: bkl	
Sensitivity:	0.4500
Specificity:	0.9060
Class: df	
Sensitivity:	0.9688
Specificity:	0.9692
Class: mel	
Sensitivity:	0.4107
Specificity:	0.9083
Class: nv	
Sensitivity:	0.4513
Specificity:	0.9462
Class: vasc	
Sensitivity:	0.9440
Specificity:	0.9693

(a)DT

Class: akiec	
Sensitivity:	0.9597
Specificity:	0.6584
Class: bcc	
Sensitivity:	0.5156
Specificity:	0.9987
Class: bkl	
Sensitivity:	0.4500
Specificity:	0.9894
Class: df	
Sensitivity:	0.9688
Specificity:	1.0000
Class: mel	
Sensitivity:	0.4107
Specificity:	0.9856
Class: nv	
Sensitivity:	0.4513
Specificity:	0.9934
Class: vasc	
Sensitivity:	0.9440
Specificity:	1.0000

(b) RF

For the majority of cases, logistic regression is well-balanced because well balanced and continually maintains high sensitivity and specificity. This balance makes it a dependable choice for most cases, effectively minimizing both false positives and false negatives SVM performs exceptionally well in 'akiec' with great specificity, but slightly lower specificity in other classes. Therefore, SVM becomes an attractive option when the utmost priority is to minimize false positives, making it an excellent choice for scenarios where the cost of misdiagnosis is high.

Table 4.6.2 sensitivity, specificity for (a) LR &(b) SVM

Class: akiec	
Sensitivity:	0.9597
Specificity:	1.0386
Class: bcc	
Sensitivity:	0.5156
Specificity:	0.8728
Class: bkl	
Sensitivity:	0.4500
Specificity:	0.9205
Class: df	
Sensitivity:	0.9688
Specificity:	0.9826
Class: mel	
Sensitivity:	0.4107
Specificity:	0.9240
Class: nv	
Sensitivity:	0.4513
Specificity:	0.9239
Class: vasc	
Sensitivity:	0.9440
Specificity:	0.9773

(a) LR

Class: akiec	
Sensitivity:	0.9597
Specificity:	0.9959
Class: bcc	
Sensitivity:	0.5156
Specificity:	0.9197
Class: bkl	
Sensitivity:	0.4500
Specificity:	0.9099
Class: df	
Sensitivity:	0.9688
Specificity:	0.9813
Class: mel	
Sensitivity:	0.4107
Specificity:	0.9174
Class: nv	
Sensitivity:	0.4513
Specificity:	0.9304
Class: vasc	
Sensitivity:	0.9440
Specificity:	0.9840

(b) SVM

KNN excels in 'bcc' but displaying varied specificity, which might compromise sensitivity in other classes while still delivering excellent sensitivity. KNN can be a useful tool for optimizing for particular scenarios, with an emphasis on avoiding false negatives; nevertheless, the trade-off is found in other classes, where it may compromise sensitivity. Across several skin condition classes, the XGBoost classification model performs inconsistently. It is noteworthy that it demonstrates great sensitivity in classes like "akiec" and "df," demonstrating its competence in accurately detecting affirmative examples for these conditions. In most classes, it also keeps a decent level of specificity, proving its accuracy in identifying real negatives. Although it shows slightly lower values for classes like "bkl" and "mel," there is still opportunity for improvement in specificity for these classes. To obtain a more equitable trade-off between sensitivity and specificity across all skin conditions, it may be helpful to fine-tune the model for these particular classes.

Table 4.6.3 sensitivity, specificity for (a) KNN &(b) XGBoost

Class: akiec	
Sensitivity:	0.9597
Specificity:	0.7989
Class: bcc	
Sensitivity:	0.5156
Specificity:	1.0013
Class: bkl	
Sensitivity:	0.4500
Specificity:	0.9457
Class: df	
Sensitivity:	0.9688
Specificity:	0.9451
Class: mel	
Sensitivity:	0.4107
Specificity:	1.0026
Class: nv	
Sensitivity:	0.4513
Specificity:	0.9856
Class: vasc	
Sensitivity:	0.9440
Specificity:	0.9507

(a)KNN

Class: akiec	
Sensitivity:	0.9597
Specificity:	1.0124
Class: bcc	
Sensitivity:	0.5156
Specificity:	0.9224
Class: bkl	
Sensitivity:	0.4500
Specificity:	0.9007
Class: df	
Sensitivity:	0.9688
Specificity:	0.9839
Class: mel	
Sensitivity:	0.4107
Specificity:	0.9096
Class: nv	
Sensitivity:	0.4513
Specificity:	0.9278
Class: vasc	
Sensitivity:	0.9440
Specificity:	0.9827

(b) XGBoost

Each model has distinctive qualities in the thorough examination. The high sensitivity and specificity of logistic regression continuously maintains a balanced approach. SVM performs well in the 'akiec' class with excellent specificity. KNN gives flexibility with high sensitivity but variable specificity, excelling particularly in "bcc." SVM achieves a decent balance between sensitivity and specificity, whereas Random Forest promotes sensitivity. Performance from XGBoost is inconsistent and might use some fine-tuning. The selection of the classifier should be in accordance with the requirements of the work at hand. Reliable all-rounders' Logistic Regression and SVM should be used, while KNN should be used for tuning, Random Forest and XGBoost should be used for sensitivity, and SVM should be used for balancing.

CHAPTER 5
Conclusion and Future Research

5.1 Conclusion

In the outlined work, skin lesion images were assessed using machine learning and CNN algorithms, and then accurately classified into the appropriate class. The HAM10000 dataset was used for the study. To acquire an impacting result from the dataset, which was severely imbalanced and had a majority of instances belonging to one class, multiple preprocessing steps were employed. label encoding, image resizing, dataset splitting, dataset balancing was some of the pre-processing steps used before the training/testing phase. Two different approach was introduced to get the best result. As the Deep Learning approach, a customized CNN model was deployed and as the ML approach the features of the images were extracted first then several machine learning algorithms were used as a classifier to get the desired result. The findings indicate that the modified CNN outperformed the suggested machine learning algorithms, achieving an accuracy of 80% and the highest accuracy from all the seven types of ML classifier is 74 %. This indicates that the proposed CNN performs better in terms of classification for the HAM10000 data set.

5.2 Future Research

In crucial situations, the accuracy of the detecting the disease is essential. Delay in diagnosis or a false positive result might have serious consequences, including death. Early diagnosis is crucial for determining treatment outcomes and patient survival rates in conditions like cancer. The well-being of patients may suffer as a result of misdiagnoses leading to ineffective therapies or pointless operations. We are looking for various ways to improve the accuracy and improve our model in our future research. We are aiming to explore other traditional features extraction methods as well as the ensemble learning methodology for the classifier which will build a stronger, more precise predictive model. We are also concentrating on the future potential of fine-tune pre-trained CNN models, which are thought to be more effective models. This model may be applied to other datasets and can be a high-performance identification or medical diagnosis system.

Reference:

[1] M. Dildar *et al.*, “Skin cancer detection: A review using deep learning techniques,” *Int. J. Environ. Res. Public Health*, vol. 18, no. 10, 2021, doi: 10.3390/ijerph18105479.

[2] R. L. Siegel, K. D. Miller, H. E. Fuchs, and A. Jemal, “Cancer statistics, 2022,” *CA. Cancer J. Clin.*, vol. 72, no. 1, pp. 7–33, Jan. 2022, doi: 10.3322/caac.21708.

[3] “Skin cancer – IARC.” https://www.iarc.who.int/cancer-type/skin-cancer/ (accessed Jun. 18, 2023).

[4] “Skin Cancers in Bangladesh.” https://www.worldlifeexpectancy.com/bangladesh-skin-cancers (accessed Jun. 18, 2023).

[5] M. Nawaz *et al.*, “Skin cancer detection from dermoscopic images using deep learning and fuzzy k-means clustering,” *Microsc. Res. Tech.*, vol. 85, no. 1, pp. 339–351, Jan. 2022, doi: 10.1002/jemt.23908.

[6] Y. Li and L. Shen, “Skin lesion analysis towards melanoma detection using deep learning network,” *Sensors (Switzerland)*, vol. 18, no. 2, Feb. 2018, doi: 10.3390/s18020556.

[7] D. N. A. Ningrum *et al.*, “Deep Learning Classifier with Patient’s Metadata of Dermoscopic Images in Malignant Melanoma Detection,” *J. Multidiscip. Healthc.*, vol. Volume 14, pp. 877–885, Apr. 2021, doi: 10.2147/JMDH.S306284.

[8] H. A. Haenssle *et al.*, “Man against Machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists,” *Ann. Oncol.*, vol. 29, no. 8, pp. 1836–1842, Aug. 2018, doi: 10.1093/annonc/mdy166.

[9] A. G. C. Pacheco and R. A. Krohling, “The impact of patient clinical information on automated skin cancer detection,” *Comput. Biol. Med.*, vol. 116, p. 103545, Jan. 2020, doi: 10.1016/j.compbiomed.2019.103545.

[10] B. Shetty, R. Fernandes, A. P. Rodrigues, R. Chengoden, S. Bhattacharya, and K. Lakshmanna, “Skin lesion classification of dermoscopic images using machine learning and convolutional neural network,” *Sci. Rep.*, vol. 12, no. 1, Dec. 2022, doi: 10.1038/s41598-022-22644-9.

[11] M. F. Jojoa Acosta, L. Y. Caballero Tovar, M. B. Garcia-Zapirain, and W. S. Percybrooks, “Melanoma diagnosis using deep learning techniques on dermatoscopic images,” *BMC Med. Imaging*, vol. 21, no. 1, p. 6, Dec. 2021, doi: 10.1186/s12880-020-00534-8.

[12] A. A. Cruz-Roa, J. E. Arevalo Ovalle, A. Madabhushi, and F. A. González Osorio, “A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2013, vol. 8150 LNCS, no. PART 2, pp. 403–410. doi: 10.1007/978-3-642-40763-5_50.

[13] L. Singh, R. R. Janghel, and S. P. Sahu, “Designing a Retrieval-Based Diagnostic Aid using Effective Features to Classify Skin Lesion in Dermoscopic Images,” *Procedia Comput. Sci.*, vol. 167, no. 2019, pp. 2172–2180, 2020, doi: 10.1016/j.procs.2020.03.267.

[14] T. J. Brinker *et al.*, “A convolutional neural network trained with dermoscopic images performed

on par with 145 dermatologists in a clinical melanoma image classification task," *Eur. J. Cancer*, vol. 111, pp. 148–154, Apr. 2019, doi: 10.1016/j.ejca.2019.02.005.

[15] P. Tschandl, C. Rosendahl, and H. Kittler, "Data descriptor: The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Sci. Data*, vol. 5, Aug. 2018, doi: 10.1038/sdata.2018.161.

[16] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, 2000. doi: 10.1017/cbo9780511801389.

[17] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. New York, NY: Springer New York, 2009. doi: 10.1007/978-0-387-84858-7.

[18] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.

[19] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Aug. 2016, vol. 13-17-August-2016, pp. 785–794. doi: 10.1145/2939672.2939785.

[20] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*. CRC Press, 2017. doi: 10.1201/9781315139470.

[21] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression: Third Edition*. wiley, 2013. doi: 10.1002/9781118548387.

Turnitin Originality Report

Processed on: 10-Sep-2023 11:12 +06
ID: 2161779783
Word Count: 15472
Submitted: 1

191-19-2070 By Israt Jahan Nipa

Similarity Index	Similarity by Source	
18%	Internet Sources:	15%
	Publications:	11%
	Student Papers:	11%

1% match (Internet from 21-Nov-2022)
http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/20.500.11948/1086/053-19-315_P03936.pdf?isAllowed=y&sequence=1

1% match (Internet from 22-Jul-2023)
http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/10314/22981.pdf?isAllowed=y&sequence=1

1% match (Internet from 23-Sep-2022)
http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/20.500.11948/1098/071-19-661_P03923.pdf?isAllowed=y&sequence=1

1% match (Internet from 29-Jun-2023)
http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/9790/22506.pdf?isAllowed=y&sequence=1

1% match (student papers from 09-May-2019)
Submitted to Daffodil International University on 2019-05-09

1% match (Internet from 13-Dec-2022)
https://huggingface.co/spaces/HugsVision/Skin-Cancer/commit/9573c25e489378d99f88df55e847dbd774aa292d.diff

< 1% match (Internet from 18-Jan-2023)
http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/20.500.11948/1084/p03943.pdf?isAllowed=y&sequence=1

< 1% match (Internet from 05-Jan-2022)
http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/5901/171-15-9084%20%2814-27%25%29.pdf?isAllowed=y&sequence=1

< 1% match (Internet from 05-Jan-2022)
http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/5831/171-15-1422%2017%25.pdf?isAllowed=y&sequence=1

< 1% match (Internet from 21-Jul-2023)
http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/10074/22725.pdf?isAllowed=y&sequence=2

< 1% match (Internet from 29-Jun-2023)
http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/9817/22528.pdf?isAllowed=y&sequence=1

< 1% match (Internet from 06-Feb-2022)
http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/6725/172-15-9708%20%2813%25%29%20clearance.pdf?isAllowed=y&sequence=1

< 1% match (Internet from 06-May-2023)
http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/9011/21441.pdf?isAllowed=y&sequence=1

< 1% match (Internet from 29-Jun-2023)
http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/8867/181-15-10699.pdf?isAllowed=y&sequence=1

< 1% match (Internet from 22-Jul-2023)
http://dspace.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/10260/22954.pdf?isAllowed=y&sequence=1

< 1% match (student papers from 24-Jun-2019)
Submitted to Daffodil International University on 2019-06-24

< 1% match (student papers from 30-Jun-2020)
Submitted to Daffodil International University on 2020-06-30

< 1% match (Internet from 15-Feb-2023)
https://www.researchgate.net/profile/Manu-Goyal-3/publication/346894346_Artificial_intelligence-based_image_classification_methods_for_diagnosis_of_skin_cancer_Challenges_and_opportunities/links/5fd1ce64299bf188d406f643/Artificial-intelligence-based-image-classification-methods-for-diagnosis-of-skin-cancer-Challenges-and-opportunities.pdf

< 1% match (Internet from 04-Feb-2023)
https://www.researchgate.net/publication/368159157_LesionAid_Vision_Transformers-based_Skin_Lesion_Generation_and_Classification

< 1% match (Internet from 06-Nov-2022)
https://www.researchgate.net/publication/364557258_Malignant_melanoma_diagnosis_applying_a_machine_learning_method_based_on_the_c

< 1% match (Internet from 12-Nov-2022)