

Machine Learning-Based Visual Health Monitoring System for Garment Factory Workers

by

Faiyaz Hossain Rusad

20101030

Ahanaf Ferdous

20201014

Rusab Mustafa Arif

20241012

Robiul Hasan Apu

20301323

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
Brac University
January 2025

© 2025. Brac University
All rights reserved.

Declaration

It is hereby declared that.

1. The thesis submitted is our original work while completing a degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material that has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Faiyaz Hossain Rusad

20101030

Ahanaf Ferdous

20201014

Rusab Mustafa Arif

20241012

Robiul Hasan Apu

20301323

Approval

The thesis titled “**Machine Learning-Based Visual Health Monitoring System for Garment Factory Workers**” submitted by:

1. Faiyaz Hossain Rusad (20101030)
2. Ahanaf Ferdous (20201014)
3. Rusab Mustafa Arif (20241012)
4. Robiul Hasan Apu (20301323)

of Fall 2024 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on January 27, 2025.

Examining Committee:

Supervisor:

Dr. Jannatun Noor Mukta

Assistant Professor
Department of Computer Science and Engineering
Director, BSDS
United International University

Program Coordinator:
(Member)

Dr. Golam Rabiul Alam

Professor
Department of Computer Science and Engineering
School of Data and Sciences
Brac University

Head of Department:

Dr. Sadia Hamid Kazi

Associate Professor
Department of Computer Science and Engineering
School of Data and Sciences
Brac University

Abstract

Garment factory workers are at high risk for health hazards due to their regular exposure to synthetic dyes and harmful chemicals, which can lead to various skin diseases. This study introduces a machine learning-based visual health monitoring system to address these challenges. The system employs advanced image processing techniques to analyze images of workers' skin, allowing for the detection and evaluation of the stage and severity of any condition. In addition to image analysis, a short survey will be conducted with garment workers to collect information about their symptoms, pain levels, and how severe they perceive their skin problems. The image analysis and the survey responses will then be combined into a robust machine learning framework. To improve the reliability of diagnosis, the research utilizes convolutional neural networks (CNNs) and Transformer models, which help address potential variations in image quality. For training the models, a diverse dataset of skin disease images has been gathered from sources such as HAM10000, Kaggle, and various other online platforms. These data sets are a solid foundation for creating reliable models for accurate detection. In addition, the research includes creating a mobile application designed to support real-time diagnostics and monitoring. This study aims to improve early skin disease detection and establish an effective health monitoring system for garment factory workers, ultimately contributing to their overall health and well-being.

Keywords: skin disease detection, machine learning, CNN, Transformer models, garment workers, health monitoring.

Contents

1	Introduction	1
1.1	History of Garment Workers' Diseases Caused by the Garment Factory . . .	1
1.2	Statistics on Losses Due to the Diseases of the Workers	2
1.3	Treatment and Prevention in Different Stages of Diseases	2
1.4	Problem Statement	3
1.5	Research Objective	3
1.6	Research Question	4
2	Literature Review	5
2.1	Related Work	5
3	Work Plan	11
3.1	Work Plan Diagram	11
3.2	Work Plan Description:	12
4	Survey and Field Work Analysis:	13
4.1	Survey Methodology	13
4.1.1	Demographic Overview:	13
4.1.2	Working Conditions:	14
4.1.3	Health Issues:	16
4.2	Workers' Understanding of Health & Safety	16
4.3	Consultation with the Factory-Appointed Doctor	18
5	Data	19
5.1	Data Splitting	19
5.1.1	Train-Test Split	19
5.2	Data Augmentation and Balancing	21
5.2.1	The below methods are used for Data Augmentation-	21
6	Model Implementations And Analysis:	26

6.1	Deep Learning Models:	26
6.2	Comparative Analysis:	31
6.3	Custom Model:	33
6.3.1	Introduction to TransNet-X	33
6.3.2	Why choosing MobileNetV3 & Swin Transformer-	34
6.3.3	Details of TansNet-X (Hybrid Model):	36
7	Results:	41
7.1	CNN Models Result:	41
7.1.1	DenseNet	41
7.1.2	InceptionV3	42
7.1.3	MobileNet	43
7.1.4	Xception	44
7.1.5	VGG16	45
7.1.6	ResNet50	46
7.2	Transformer Models Result:	47
7.2.1	Deit	47
7.2.2	Swin	48
7.2.3	ViT	49
7.3	Custom Model Results:	50
7.3.1	The Experimental report of TransNet-X is given below:	50
7.3.2	Advantages and Disadvantages of TransNet-X:	53
8	Approach	55
8.1	User Login	55
8.2	Skin Analysis (User Functionality)	57
8.3	Viewing Results	58
8.4	Admin Features	59
8.5	Managing Databases	60
8.6	Sitemap of the APP -	60
8.7	User Flow for Worker and Admin	61
9	Conclusion	62

Chapter 1

Introduction

Employee well-being has become a major concern for organizations in this fast-moving business world. As technology continues to advance, numerous innovative solutions have emerged to improve workplace safety. One notable innovation is the Visual Monitoring Health System (VMHS) [1], which is specifically made to address the health issues faced by garment manufacturing workers. With the state-of-the-art AI technology integrated into VMHS, employees' health is monitored in real-time with an early warning for health risks and better health management. The latest image recognition technology is used to analyze the collection of visual data at multiple industrial monitoring points [2]. Machine learning algorithms are deployed to detect early symptoms of fatigue, heat stress, and repetitive strain injuries, which are frequent problems for garment factory workers. VMHS takes it even further, providing real time surveillance, prescriptive recommendations, and insight into regulatory measures based on employee health profiles to enable proactive health programs [3]. With constant monitoring of worker health and attempts to maximize preventive care, VMHS can greatly limit the scope of workers coming down with illnesses that can be traced back to the workplace. Therefore, it influences a healthier work environment and enhances employee productivity and morale [4]. In an age of rapid technological change, VMHS is a great example of how AI can transform workplace safety and worker well-being in industries.

1.1 History of Garment Workers' Diseases Caused by the Garment Factory

Workers in the garment industry face various health problems due to the rigorous conditions they undergo in the factories. The workers in textile mills during the Industrial Revolution faced respiratory conditions caused by long exposure to cotton dust, and this is what the garment workers are experiencing today [4]. In addition, many garment workers suffer from skin-related problems due to constant exposure to chemicals in the fabric dyeing and finishing processes. Those engaged in repetitive tasks, like stitching, are also at risk for developing musculoskeletal injuries [5]. Despite the wealth of industrial laws and rules concerning workplace safety, several problems continue to exist due to ignorance. These include respiratory illnesses because of the lack of adherence to safety rules and the development of mental health problems brought about by the adverse conditions of work. The countries that host garment workers must keep their promises to develop what is known as a safe working place, enforcing, and ensuring compliance with safety protocols

that also involve structural control over the manufacturing of garments in restraint of possible injury or mishap [6].

1.2 Statistics on Losses Due to the Diseases of the Workers

The health issues of garment workers have consequences on individuals and are reflected in a wider economic impact. Although the overall financial loss resulting from worker illness is impossible to quantify from purely statistical data, several components form the sum of this burden. The first of the direct impacts is loss of productivity caused by employees being absent from work for infections and occupational diseases. As a result, whenever workers fall sick and fail to come to work, the result is disrupted production schedules, delayed orders, and increased overheads. A firm's reputation could also be damaged if such delays consistently cause inconvenience to customers. Even the workers who go to work despite sickness will work poorly hence further affecting overall productivity [5]. In addition to productivity losses, the financial burden of occupational diseases encompasses medical expenses for affected employees, including consultation fees, diagnostic tests, medications, and hospitalization costs. These expenses can add up quickly, particularly for chronic or severe conditions that require ongoing treatment. In some instances, employment contracts may obligate garment manufacturers to cover specific medical costs, which adds to the financial pressure on businesses. The economic impact also includes workforce turnover, as frequent illness-related absences can lead to higher costs for replacement and training. Increased turnover rates compel companies to invest in hiring and taking in new employees, further straining their resources [4]. However, financial losses could be countered through strategies for occupational health and safety measures. Healthcare strategies, in particular, have proved helpful to garment manufacturers and have decreased the rate of absenteeism while cost-effectively raising productivity, thereby guaranteeing a healthier and sustainable workforce.

1.3 Treatment and Prevention in Different Stages of Diseases

According to the type and stage of the disease, various strategies need to be emphasized on either prevention or management of diseases in garment factories. Early identification and timely interventions are critical in ensuring that proper treatment is administered to the concerned workers [6]. Diagnosis and preventive measures in workplace health policies should go together to take care of their working staff. A bit of prevention is worth more than a cure, which underlines the need for actions aimed at minimizing health risks before they become a serious condition [4]. Preventive measures include supplying protective gear like masks and gloves to limit workers' exposure to dust and harmful chemicals. Regular health screenings and ergonomic training can also help reduce the risk of musculoskeletal disorders (MSDs) and repetitive strain injuries (RSIs) by teaching workers proper posture and movement techniques during their tasks. Upholding high hygiene standards and ensuring adequate ventilation are essential in preventing respiratory illnesses linked to dust and chemical exposure. The awareness campaigns on workers' lifestyle health and early reporting, together will make a cultural shift toward employee well-being. Workplace

assessment regularly, and strict adherence to safety regulations, are very much necessary to save garment factories from the spread of occupational diseases. Safety measures enforced, early treatment, and healthier work conditions could help garment manufacturers reduce work-related illnesses by as much as in other sectors of industry, which would have the effect of increasing workplace safety [1].

1.4 Problem Statement

The handling, bleaching, and coloring of textile materials with the use of chemical dyes in garment factories often lead to skin diseases of the workers. Even with existing safety protocols, identifying skin diseases resulting from long-term chemical exposure continues to be a significant challenge [5]. Delayed diagnoses can lead to deteriorating health conditions, increased medical costs, and a decrease in the quality of life for workers. This calls for an immediate need for the development of a continuous monitoring and detection system that gives early signs of skin diseases arising among factory employees [1]. The proposed system will integrate various advanced imaging with machine learning algorithms that can continuously provide real-time data and alert the appearance of skin abnormalities, thus enabling their timely medical interventions and treatment according to [7].

The system will be trained on a comprehensive dataset that includes images of various skin diseases at different stages. When a new image is analyzed, it will be compared to the dataset to accurately identify the type and seriousness of the disease, suggest possible causes, and recommend suitable treatments [8]. Through preemptive work like this, health risks are lowered, disease progression is reduced, i.e., absenteeism is reduced, and workers' productivity increased [5]. Furthermore, the main objective of the system is to fill the gaps in health management that are caused by healthcare facilities' lack of equipment and poor workplace health management.

1.5 Research Objective

The main objective of this study is to create a system that minimizes health risks, especially skin diseases encountered by garment factory workers. The research will start with an examination of prevalent skin health issues among workers and the collection of data regarding the types and stages of diseases they face [1]. Skin images from workers will be captured and analyzed for this purpose using advanced techniques in image processing methods [8]. The images are further processed and classified to bring out an accurate detection and classification of the various skin conditions [7]. Workers will also be surveyed on their personal experiences of these health conditions and their feelings about them.

After collecting the data, it will be analyzed using machine learning frameworks to better understand the prevalence and severity of these health concerns [9]. The research will also explore the causes of skin diseases and suggest several prevention strategies that might bring down these complications in the future. Ultimately, the research aims at promoting better health and safety at the workplace in garment factories by offering comprehensive information about the health problems that workers experience and developing remedies that promote better health outcomes for enhanced productivity, and improving working conditions [3].

1.6 Research Question

The objective of this study is to gain a total understanding of health issues, specifically skin diseases, that garment factory workers endure. To achieve this, the study will address the following key research questions:

1. **What are the most common skin diseases among garment factory workers and how do they evolve?**
2. **How do these skin conditions affect workers health, well-being and productivity?**
3. **What are the workplace factors (e.g. chemical exposure, working conditions, job roles) that contribute to skin diseases?**
4. **What is the level of efficiency concerning the already-developed safety measures and prevention arising from these identified health hazards?**
5. **What can we do to prevent skin diseases and improve their workers' health?**

So that these questions can be explored, the study will employ a mixed methods approach, like a combination of qualitative and quantitative research methods. We will also listen to first-hand insights from workers, through interviews, to understand how they view skin diseases and the struggles they encounter. Averaging the prevalence and severity of these conditions across different levels of jobs and factory environments will be one of our quantitative data analyses. To examine the relationship between workplace factors and skin disease incidence and to evaluate the use of existing preventive measures, statistical techniques such as regression analysis will be used.

By integrating both qualitative and quantitative methods, this research will provide a 'comprehensive view' of the health challenges that garment factory workers encounter. The results will assist in formulating targeted interventions and policy recommendations aimed at improving workplace health and safety, ultimately boosting workers' well-being and productivity.

Chapter 2

Literature Review

2.1 Related Work

This study focuses on creating a thorough health monitoring system specifically for garment factory workers, taking into account their distinct occupational challenges. These workers encounter significant health risks, especially in developing countries such as Bangladesh. The research combines insights about health monitoring, diagnostic innovations, and occupational health frameworks.

2.1.1 Occupational Health Challenges in Garment Factories:

Garment factory workers often face various occupational health hazards due to harmful chemical exposure, poor working conditions, and the absence of safety measures. The incidence of ailments such as respiratory diseases, musculoskeletal disorders, skin diseases, and other occupational diseases has been documented in several studies.

For instance, a cross-sectional study in Ethiopia [1] revealed that 66% of garment workers reported being diagnosed mainly for respiratory diseases, musculoskeletal disorders, or injuries. Major respiratory problems stemmed from exposure to endotoxins and inhalable cotton dust during textile processing. Similarly, repetitive manual labor and heavy machinery caused traumatic injuries and MSDs. Research in Kasur, Pakistan [3], noted that long-term exposure to cotton dust increased workers' risk of respiratory and pulmonary diseases by 5-6% annually. Skin conditions such as contact dermatitis, allergic reactions, and infections were linked to poor hygiene and a lack of protection gear.

In Bangladesh, many reports indicated high rates of musculoskeletal pain, headaches, and body pain [10], driven by long working hours and a lack of ergonomic support. Surveys conducted with workers in combination with clinical assessments indicated that more than 74% of workers suffer from injuries, such as cuts and pricks from unsafe working environments [11]. A total of 162 female garment workers participated in a cross-sectional survey that was done in Tumkur, India, about their health and QOL using WHO-QOL and GHQ-28 questionnaires. High prevalence of musculoskeletal disorders, 23.4% back pain, 8.6% joint pain, and stress-related tension headaches were seen at a rate of 41.9% [12].

The methods of data collection differed from one study to another, including interviews of workers, clinical diagnosis, and factory-level surveys. Some analysis techniques like SPSS and descriptive statistics were used in analyzing the factors that influence health out-

comes. The research highlighted the critical need for implementing Occupational Safety and Health Management Systems (OSHMS) [13] and enhancing safety measures to reduce these risks. The garment factories are in dire need of effective improvement of workplace conditions, good ventilation, use of protective gear, and access to health facilities as possible mitigation measures for these occupational hazards [6].

2.1.2 Advancements in Skin Disease Diagnostics:

Newly diagnosed skin diseases increasingly rely on advanced image processing and machine learning methods that allow more accurate identification. Many have researched the effect of these promising approaches for skin disease diagnosis, employing a variety of methods for feature extraction, classification, and segmentation.

Deep learning models, particularly Convolutional Neural Networks (CNNs), have become popular for diagnosing skin diseases automatically. Research has shown that CNNs can classify skin conditions with impressive accuracy [9]. For example, the MobileNet model was utilized for diagnosing skin diseases on mobile devices, featuring lightweight architectures that allow for quicker processing. Hybrid models that merge CNNs with advanced methods, such as Error Correcting Output Codes (ECOC) and Social Spider Optimization (SSO), have achieved remarkable accuracy rates of over 97% [14]. ResNet-50 and AlexNet were able to achieve better diagnostic performance using transfer learning on datasets like ISIC 2018 [15]. Classification frameworks have also incorporated machine learning algorithms, such as Random Forests, Support Vector Machines (SVM), and K-Nearest Neighbors (KNN) [7], achieving accuracies up to 97% on dermoscopic image datasets like HAM10000 [16]. These approaches emphasize early detection and precise classification of conditions like melanoma, contact dermatitis, and other skin disorders. An automated image-based method for the diagnosis of skin diseases using CNNs for feature extraction and a softmax classifier for classification was proposed in a research paper [17]. The Xiangya-Derm-based dataset yielded an 87% accuracy for the model, manifestly exhibiting a faster and more reliable alternative to traditional dermatological diagnostic methodologies.

Major image processing techniques were used, including contrast enhancement, noise reduction, Otsu thresholding, and DeepLabv3+, enhancing the quality of the input images, and hence improving model performance. Hair removal, Gaussian filtering, and feature extraction techniques involving LBP and Gray Level Co-occurrence Matrix have been pivotal. One of the studies found the technique of both intensity-based registration methods, with mutual information as the similarity metric, to overlay thermal infrared and visual images for improved diagnostics clinically [18]. Segmentation methods, including the Grabcut algorithm and K-means clustering, have been used to accurately define lesion boundaries [19]. These strategies enhance the accuracy of diagnostic tools and minimize variability caused by image artifacts. Another study focused on extracting the mean and standard deviation of color channels, creating histograms, and applying a custom Sobel operator for edge detection to analyze and segment skin lesion images [20]. We also discovered a study that examines advanced non-machine-learning image processing methods for detecting skin lesions [21]. This research focuses on automated skin lesion detection, using techniques like CLAHE for enhancing contrast, Otsu's thresholding, and active contour segmentation to accurately identify lesions. It applied the ABCD rule for feature extraction and utilized a rule-based classification system for detecting melanoma skin cancer.

Various researchers have targeted analyzing the health and issues of garment factory workers via surveys and questionnaires. For example, a geographically extensive survey was carried out of over 1100 employees in Bangladesh to determine their health and satisfaction

with their work [4]. This study came out with important findings such as the existence of musculoskeletal pain, headaches, and the lack of appropriate health insurance. Another study [6] focused on seeking healthcare behavior and knowledge of health insurance of 486 employees and found low awareness of healthcare rights and access to health services. Following the same trend, a questionnaire designed to evaluate the low back pain reported by garment workers showed that over 90% of respondents complained of pain at some stage in their work lives [22]. Such a survey provides a good starting point to target root causes of health issues prevalent among workers which is a good base for designing appropriate interventions and policy recommendations. A survey on DL and AI in COVID-19 detection [23] noted that they can be used as an adjunct to traditional diagnostic procedures like RT-PCR tests and achieve similar accuracy (82% to 86%) and better sensitivity (92% to 95%). In addition to existing surveys, diagnostic methods are employed in this study to fill the gap in monitoring the health of garment factory workers.

2.1.3 Machine Learning Techniques in Health Monitoring:

Machine learning has largely brought much-needed improvement in health monitoring systems, giving a new perspective toward a novel early diagnosis and intervention solution. This section discusses the ML methods that transcend deep learning models, insights from existing health-monitoring systems and their operational deficiencies, and proposals for further improvements to the existing systems.

Traditional machine learning algorithms, including Support Vector Machines, Random Forests, and Decision Trees, have been widely applied in health monitoring. For example, the SVMs showed very high accuracy, up to 95% in skin disease classification on datasets like ISIC and HAM10000, which underlines their robustness and simplicity [7]. Random Forest classifiers were used for the classification of health conditions like respiratory and musculoskeletal disorders, which showed reliable results in occupational health studies [24].

In some cases, K-Nearest Neighbors (KNN) and Naive Bayes classifiers have helped to spot specific health risks for workers such as injuries from poor posture. These methods work well with smaller sets of data [25]. Methods that use [26] multiple weak classifiers together, like Gradient Boosting and AdaBoost, look promising. They can boost the prediction power of health monitoring systems [27]. Techniques to pull out key features and reduce data size such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA), have also been used. These methods clean up and improve the data that goes into ML models, leading to better sorting results [28]. Numerous studies even mention the importance of AI in enhancing the reliability of diagnoses and improving patient outcomes in skin cancer detection. One of them confirms that artificial intelligence (AI) prediction models can effectively diagnose skin cancer through dermoscopy images [5]. The presented models automate the diagnosis process, enhancing early detection rates to improve patient care and point out the great benefit to society from AI applications in dermatology. AI techniques, like transfer-learning, are used for the detection of retinal disorders in the study [29], yielding expert-level accuracy tested on a dataset of more than 100,000 retinal images.

Most of the health monitoring systems currently available are not accessible, scalable, and integrated with multimodal data. There is a shortage of an effective [30] health monitoring system in the RMG sector of Bangladesh due to problems such as insufficient lighting and ventilation, dust exposure, accidents, and safety measures, affecting the health and overall satisfaction of the workers [11]. Many experts have pointed out problems with tracking health in real-time [31] and the limited use of AI-based answers in places with few resources. The lack of full health records and reliance on collecting data by hand make

these systems less effective. For instance, a study on [32] garment workers in Bangladesh showed how hard it is to keep an eye on muscle and bone problems and other job-related health risks. This is because of poor ergonomic measures and no digital health records [33]. Also, current systems can't predict problems well, which limits taking action to lower health risks. Workers have little access to good healthcare, with most turning to local drug sellers and not knowing about health insurance [6]. There can also be a system focusing on diagnostic techniques, that shows how thermal and visual imaging combinations can deliver more precise medical imaging and diagnostics [34].

It has been suggested by researchers that integrating machine learning frameworks with wearable sensors and IoT devices can allow real-time health tracking and predictive analytics. Authors in [35] proposed IoT-based systems integrated with ML algorithms to enhance remote health monitoring. Integrating traditional ML models with AI-driven analytics may overcome scalability issues and improve the diagnostic accuracy of existing systems. Other recommendations include the use of mobile applications to capture self-reported health data, as in [4], which highlighted the use of questionnaires in capturing the health condition of workers. The implementation of automated diagnostic tools using ML may reduce reliance on manual monitoring, thus making health systems more efficient and accessible. Addressing these limitations and incorporating machine learning techniques in health monitoring will contribute to increasingly effective, scalable, and user-centric solutions for workplace health and well-being improvement.

2.1.4 Applications of Deep Learning in Disease Detection:

Deep learning has emerged as one of the methodology paradigms that is changing the face of disease detection, especially in dermatology, where most practices are performed through CNN, and Transformer models are commonly employed for image classification and diagnosis.

Most of the CNN models, including ResNet50, VGG16, InceptionV3, EfficientNet, and DenseNet, have shown outstanding performance in diagnosing skin diseases. For example, works using ResNet and InceptionV3 have reported high accuracy in detecting melanoma, based on deep feature extraction and transfer learning from their respective pre-trained models [14]. VGG16 and DenseNet were utilized for their ability to extract hierarchical features, which allowed for accurate classification of dermoscopic images [35]. In another study, some of the above-mentioned models and VGG19 are also used to classify skin cancer images, with InceptionV3 obtaining the best diagnostic performances. In a few cases, these beat dermatological specialists in accuracy and hence, image augmentation could improve generalization and early melanoma detection [36].

The use of CNN-based methods has been demonstrated to be useful in both disease diagnosis and classification across various fields, according to several studies. A research study emphasized the effectiveness of CNNs, including an InceptionV3-based classifier, in dermoscopy image classification to identify skin diseases with an 83% accuracy and aid tumor diagnosis [37]. An article cited in [22] also revealed that methods utilizing CNNs with high accuracy, up to 99.04% of detection rates, were found to be effective in detecting soybean leaf diseases, with more robust techniques like data augmentation, transfer learning, and hybrid models that utilized both CNN technology and shallow classifiers, particularly for large or very small datasets. A separate study emphasized the importance of segmentation in enhancing CNN-based disease detection, with one experiment using S-CNN to achieve higher accuracy (98.6%) and confidence levels than F-PNN models [38].

Recent advancements in transformer-based models have greatly enhanced the classification and segmentation of skin diseases. Vision Transformers (ViTs) are among the most

powerful and majestic architecture that have been used to classify complex skin lesions. A new lightweight Vision Transformer, known as SkinDistilViT, was developed for classifying skin lesions and achieved impressive accuracy (82.34%) on the ISIC 2019 dataset, all while maintaining computational efficiency through knowledge distillation [32]. The unique feature of being capable of capturing global dependencies within the images guarantees them to have a strong place among other CNNs. Hybrid models are proposed as combining CNN and Transformer architectures, where CNNs will offer strength in feature extraction and transformers are their high-performance modeling in long-range dependency [31]. Such a realization will boost diagnostic accuracy and robustness, especially over a larger dataset. Different architectures of transformers were compared by another study [30], where the DinoV2 achieved state-of-the-art accuracy of 96.48% on the 31-class skin diseases dataset, further showcasing the probable possibility of diagnostic enhancements with transformers. HMT-Net, a hybrid model that utilized Transformers and MLP-based modules, provided an edge in segmentation tasks on datasets like ISIC2018, particularly when dealing with segments of irregular shape and occlusions [26]. These studies indicate that transformer-based approaches in dermatology are both versatile and effective.

The datasets utilized in various studies include HAM10000, ISIC 2018, and even agricultural datasets from Kaggle. The HAM10000 dataset, featuring over 10,000 labeled dermoscopic images, has been fundamental for training deep learning models. Furthermore, other works have made use of augmentation techniques and pre-trained models for better performance in disease classification [39]. ISIC datasets have been vital in benchmarking, providing heterogeneous and high-quality images [24]. Some have incorporated datasets created by combined results from online repositories and clinical environments. Similar to our study, a study [40] proposed a LeNet-based technique for 10000 image datasets for plant disease detection, getting 95% accuracy, higher compared to traditional CNN methods.

Deep learning models have shown impressive accuracy in detecting diseases. ResNet50 and EfficientNet surpassed 95% accuracy on the ISIC and HAM10000 datasets, highlighting their effectiveness [15]. The CNN-based DeepLabv3+ model reached a 90% detection rate for pathogenic agents and malignancies in cases of atopic dermatitis, successfully segmenting skin images into distinct regions for precise classification. The performance of the model was hence comparable to manually cropped images by dermatologists and therefore proved to be reliable in skin disease diagnosis [2]. While equivalent performances were sometimes performed, transformer-based models such as ViTs and CNN-Transformer hybrids outperformed traditional CNNs in handling such imbalanced datasets. For example, a hybrid CNN-Transformer model utilizing Focal Loss attained an accuracy of 89.48% on the ISIC 2018 dataset, underscoring its potential for clinical use [31].

The introduction of deep learning models into mobile and smartphone applications has facilitated easier diagnosis of skin diseases. Systems like [41] have developed Android applications using CNN models that were fed large datasets like HAM10000 that achieved almost 99% accuracy in the 7 classes of skin diseases. Lightweight models such as MobileNet were used to facilitate real-time diagnosis on resource-constrained devices, which in turn enhances usage when addressing the needs of far-flung areas or those with fewer services [9]. On the other hand, there was an ensemble machine classification approach for targeted IoT-based systems in detecting skin diseases remotely [35]. This evidence established that deep learning model deployment is very feasible on mobile platforms; hence, there has been a narrowing of the gap from development to clinical applications. This therefore implies the need for coming up with a complete and scalable skin lesion detection system, incorporating CNNs, transformers, and mobile applications, hence agreeing with developments in the subject area.

The literature reviewed emphasizes the urgent need for better occupational health measures and the transformative capabilities of machine learning and deep learning in health monitoring and disease diagnostics. By bringing together insights from various studies, this review establishes a foundation for developing advanced, technology-driven solutions to tackle health issues faced by garment factory workers and others.

Work Plan

3.1 Work Plan Diagram

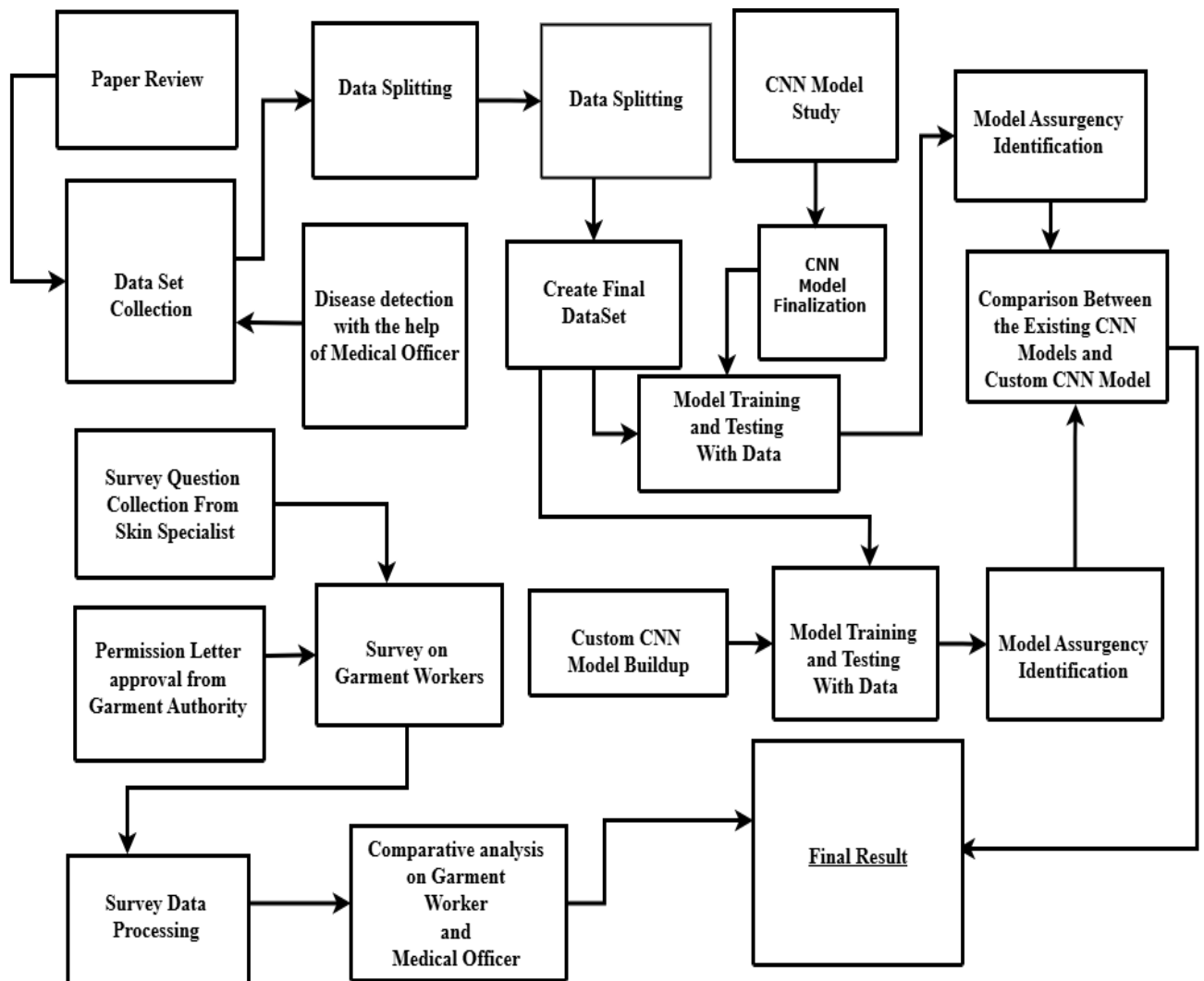


Figure 3.1: Modeling enhancement and validation procedures

3.2 Work Plan Description:

This is the work plan of the proposed research as captured in Figure 3.1. The thesis project follows a sequential workflow in developing an AI-based system for detecting skin diseases among garment workers. First of all, there is a clear outline for carrying out a literature review to study the research publications on the subject of skin disease detection using the deep learning method. After this, data collection requires the collection of different types of images of skin diseases. The images will be preprocessed and then will undergo data splitting into training, validation, testing sets, and data augmentation for enabling generation in models. Since the data needs to be as accurate as possible, getting a dermatologist's expertise in the identification of the diseases will further strengthen the reliability of the dataset.

The next part presents the development of the models: studying and testing the CNN models for their performance in classifying images and then finally selecting the models. After that comes Transformer-based models, which are optimized in terms of their performance for detecting features in skin disease images. Then both models have been decided upon, and training and testing against the prepared dataset needs to be done. Furthermore, the research is directed to the combination of both approaches through the development of an integrated hybrid model using CNN and Transformer architecture. The performances of individual models using CNN and Transformer are assessed separately, and then the accuracy testing of the hybrid model is done to validate its efficacy. On the other hand, the comparative analysis in terms of accuracy and efficiency of the already existing CNN model, Transformer model, and hybrid approach ensures the selection of an appropriate model suitable for real-world applications.

A survey is designed for this study, under the dermatologist's supervision, for validation of the system in realistic settings; thus, gathering original responses from the garment workers becomes feasible. Approval is sought beforehand from authorities dealing with garments, for both ethics and organizational policy purposes. Workers' responses provide information about workers' exposure to chemicals, work conditions, and experiences of their health effects relating to skin-related issues. In addition, a comparison of experiences in the workplace will be made against clinical observations by medical officers. Survey data will be processed and analyzed to unveil trends and risk factors leading to occupational skin diseases.

With the final model and inclusion of survey information, the application development is proceeded with. Application Interface Development - Creating a friendly application interface by implementing the deducted logic from the hybrid model so that garment workers can identify skin health by uploading images related to disease detection along with expert suggestions for its prevention or treatment.

Final report writing and reporting aims at the acquisition of the overall performance of the proposed system. By merging AI-driven diagnostics with real-world data from garment workers, this study seeks to improve occupational health and provide a practical approach to reducing skin disease risks in the garment industry.

Chapter 4

Survey and Field Work Analysis:

4.1 Survey Methodology

In our study, on the development of a healthcare care system for garment factory employees, we used a questionnaire survey and analyzed the field work being exposed to serious health risks, especially skin diseases. In particular, we interviewed the workers concerning their work situations, getting near dangerous chemicals, and the health effects the working population faces. Further, the interviewing of the doctor appointed at the factory was done to support a medical standpoint for health problems generally being confronted by these workers. We surveyed the garment factory, interviewing about 70 workers from different departments, including cutting, sewing, dyeing/printing, and finishing. The questionnaire covered demographic information, working conditions, health issues, and possible improvements. We scored and analyzed the answers, identifying key trends and challenges. At the end of the survey, we asked the garment factory workers if we could take pictures of their skin disease. Some agreed we could, some did not.

4.1.1 Demographic Overview:

- Age Distribution: Most respondents (45.5%) were in the 26-35 age range, with 36.4% being under 18. The presence of underage workers raises potential child labor issues.
- Gender Distribution: The workforce was mainly male (77.3%), with females making up 22.7%, which points to a gender disparity.
- Work Experience: A notable number of workers (31.8%) had 4-6 years of experience, while 21.2% had 6-10 years.
- Education Level: The majority of workers (53%) had completed higher secondary education, whereas 13.6% had no formal education, indicating limited opportunities for advancement.

4.1.2 Working Conditions:

- Job Roles- The job of dyeing /printing work was mostly performed (45.5%) and other included finishing (21.2%)
- Working Hours- The average working hours for 45.5% were 8-10 hours. 40.9% worked less than 6 hours, reflecting a difference in their working schedules.
- The Environment of Work Satisfaction- Quite a large number of respondents (69.7%) rated their work as satisfactory with ratings of 4 or 5, while a small percentage rated it as unsatisfactory (4.5%).
- Ventilation & Air Quality- Opinions ranged from 31.8% giving it as neutral to 27.3% considering it as good, with 10.6% rating it poor.
- Exposure to Chemicals and Dust- Reports on exposure include 40.9% occasional exposure and 16.7% continuous exposure leading to health problems.



Figure 4.1: Garments Factory



Figure 4.2: Garments Factory

4.1.3 Health Issues:

- Frequency of Health Problems: 36.4% of the participants reported health problems often, and 30.3% always faced them.
- Skin Problems: The prevalence of skin problems was always 65.2% and frequently 24.2%.
- Types of Skin Issues: Workers mentioned rashes, itching, blisters, dry skin, infections most likely because of chemical exposure.

4.2 Workers' Understanding of Health & Safety

Some of the common complaints from a large number of workers were unbearable heat, very poor ventilation systems, working with chemicals, and the need for protective gear against such elements. A significant section of workers also describes their distress due to working conditions that entail long hours that are not supportive of safety. Some access to medical consultations was there, but many would not even go near a doctor or hospital, time or money being at the forefront of their minds. Their main suggestions included better ventilation, increased protective gear, shorter working hours, cleanliness, and regular check-ups for health. Their main suggestions for improvement included better ventilation, upgraded protective gear, shorter working hours, improved cleanliness, and regular health check-ups. Although 55.4% of the workers felt neutral about the idea of a structured health monitoring system, 23.1% strongly agreed, and 21.5% agreed, indicating a favorable attitude towards the concept.



Figure 4.3: Garments Factory Worker's Skin Diseases

4.3 Consultation with the Factory-Appointed Doctor

The factory's designated doctor shared important insights into the health issues faced by workers, especially concerning skin conditions like scabies, eczema, psoriasis, and taeniasis, a fungal infection. The doctor noted that these ailments are common due to extended exposure to dust, chemicals, and poor hygiene practices. A common problem observed is that workers in most cases visit the hospital too late, probably after trying medications from local chemists, so that they wait for the ailments to disappear from their systems. This innate tendency results in aggravation and makes treatment hard to handle. Another major shortcoming was observed in the improper awareness among laborers about communicable skin-related diseases, the result of which was the widespread occurrence among others. Other conditions also included sexually transmitted diseases amongst workers, which again were not related to factory conditions but a result of the lack of education and awareness regarding sexual hygiene and health. The doctor claims that timely medical aid, routine check-ups, and increased awareness are necessary to improve the overall health conditions of workers.

In the analysis of our survey and fieldwork, serious health hazards came to light, especially skin diseases due to exposure to chemicals and poor hygiene in garment factory workers. The environments in factories, long working hours, and lack of protective means are what push these workers to the front of health perils. To safeguard their health, it is essential to implement a thorough health monitoring system, enhance workplace safety protocols, and educate workers about these risks.

Accordingly, our results strengthen the notion of a more active approach toward embedding AI-based health monitoring gadgets in the factory and create an area for proposing system development.

Chapter 5

Data

After consulting with two dermatologists and the doctor at the garments factory, we narrowed down 6 skin diseases, mostly infecting garment workers. Based on that, we collected a dataset of those 6 diseases from Kaggle and other sources, since we were not able to find the data on all of these diseases in one place we had to merge 3 datasets [42] [43] [44] to make the dataset we would be using to conduct our experiment, and then we performed data augmentation on the train and validation data to increase the quantity and balance the dataset equally. Lastly, to run the CNN models efficiently, we split the data into Train = 70%, Test = 10%, and Validation = 20% ratios. Finally, we got our expected dataset. More details are explained below with figures.

5.1 Data Splitting

The dataset was divided into three subsets: **Training**, **Validation**, and **Testing**. The splitting process followed these steps:

5.1.1 Train-Test Split

The first action divided the dataset into a Training set along with Validation + Test set portions. The Training set received 70% of the available data for training purposes yet validation and testing both took 30%. The split process for the dataset required the `train_test_split()` function from `sklearn.model_selection`. To achieve consistent splitting the `random_state` parameter was established with a constant value (`random_state=42`). The random state process maintains a consistent data split pattern every time the procedure runs.

Validation and Test Split

After performing the first division 30% of data underwent another split to produce Validation and Test sets from it. The second distribution of data assigned 50% to each subset from remaining data which yielded 15% of data for Validation and 15% for testing purposes. The `train_test_split()` function performed this second split of data while `random_state` determined the reproducible results. The specified proportions were preserved during the random distribution of the validation and test sets.

The final distribution of the dataset was as follows:

- **Training Set:** 70% of the total dataset
- **Validation Set:** 15% of the total dataset
- **Test Set:** 15% of the total dataset

The use of both `random_state` parameters removes ambiguity from data division while ensuring consistency across evaluations and training processes.

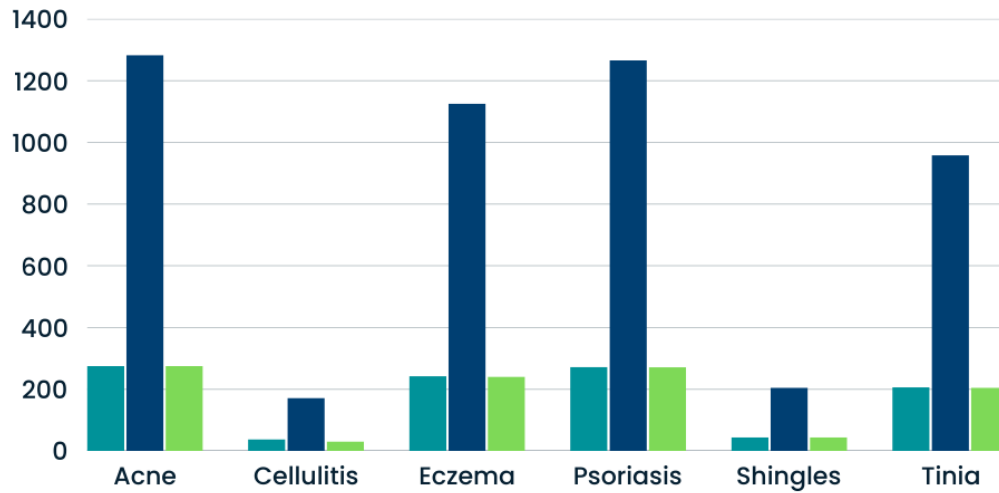


Figure 5.1: Data Splitting Original

5.2 Data Augmentation and Balancing

After splitting we augmented our data to balance them out. Our initial dataset consisted of 5686 images divided into 6 classes. Each class contained a different number of images. First, to increase the data, we augmented only the train and validation classes separately. We augmented the data so that most of the classes in train and validation had the same amount of data, and we left the test data untouched. After augmentation, we obtained a total of 8819 images. Essentially, we increased the dataset by 55.1% compared to the original dataset. The figure below provides a clearer understanding of the data augmentation process.

5.2.1 The below methods are used for Data Augmentation-

Data augmentation functions by modifying existing samples from a dataset to grow its overall size through new versions of data. The process requires researchers to randomly apply flipping, rotating, cropping and brightness adjustment to original images. By applying these alterations to the data the model becomes more powerful at recognizing various patterns across multiple conditions. Using data augmentation methods allows the production of various image versions which helps prevent model overfitting and enhances its ability to determine new untested patterns particularly valuable with small datasets.

1. Horizontal Flip (`HorizontalFlip`):

The purpose of this transformation is to horizontally reverse images which imitates orientation modifications. This transformation contains a chance of 50% that the image will receive a horizontal flip ($p=0.5$). The transformation enables object recognition under any orientation condition.

2. Vertical Flip (`VerticalFlip`):

This transformation creates a vertical image flip comparable to a top-to-bottom inversion. According to $p = 0.3$, the probability of image flipping is 30%. The effect helps process images whose top-bottom orientation changes between pictures (such as aerial or underwater captures).

3. Rotation (`Rotate`):

The function rotates images at random degrees inside defined parameter boundaries. The angle of rotation for the image does not exceed ± 30 degrees. The image rotation process has a probability of 0.7, which translates to a 70% likelihood of image rotation. Rotating images according to this operation allows the model to become insensitive to rotational changes that frequently appear across real-world data sets with different object orientations.

4. Random Brightness and Contrast Adjustments (RandomBrightness-Contrast):

Adjusts the brightness and contrast of the image randomly. The brightness control extends from 0.2 to 0.4 wherein it allows a maximum 20% brightness change. The settings allow users to change contrast values between +20% to -20% through the contrast limit of 0.2. The application of brightness and contrast adjustments has a 60% chance according to the value of p set at 0.6. The model becomes more resistant to diverse exposure levels because of this exposure adjustment simulation.

5. Random Crop (RandomCrop):

The function performs random cropping of images, maintaining a selected portion. This modification selects parts of the image that measure less than the actual dimensions by fifty pixels across length and width. The probability level stands at 0.4, which indicates that cropping of the image will happen in 40% of cases. This method allows the model to pay attention to different structural elements in various parts of the image through position manipulation.

6. Gaussian Blur (GaussianBlur):

The purpose of this code is to apply Gaussian blur for image-blurring effects that look like out-of-focus or noisy images. Blur limit: A random blur between 3 and 5. The probability (p) stands at 0.5 which indicates a blurring occurrence in half the cases. With this method, the model becomes more resilient to images of different sharpness levels or images taken with imperfect camera focus.

7. CLAHE (Contrast Limited Adaptive Histogram Equalization) (CLAHE):

The goal of this process generate improved image contrast through histogram equalization application. Clip limit: 2.0 — controls the contrast enhancement. Local contrast adjustment tiles measure (8, 8) represent the dimensions used for particular contrast modifications in the grid pattern. The probability (p) stands at 0.5 in the process which indicates a 50% chance for CLAHE enhancement application. The transformation process enhances image contrast distribution throughout various image areas particularly when dealing with lighting unevenness.

8. Random Gamma Correction (RandomGamma):

The procedure randomly adjusts the image brightness levels through gamma correction. Random selection determines the gamma value from 0.8 to 1.2. The probability of applying gamma correction stands at 0.5 which indicates a 50 percent chance for this process to occur. This adjustment operates to create lighting variations similar to real-life changes which affect particularly underexposed and overexposed sections of images.

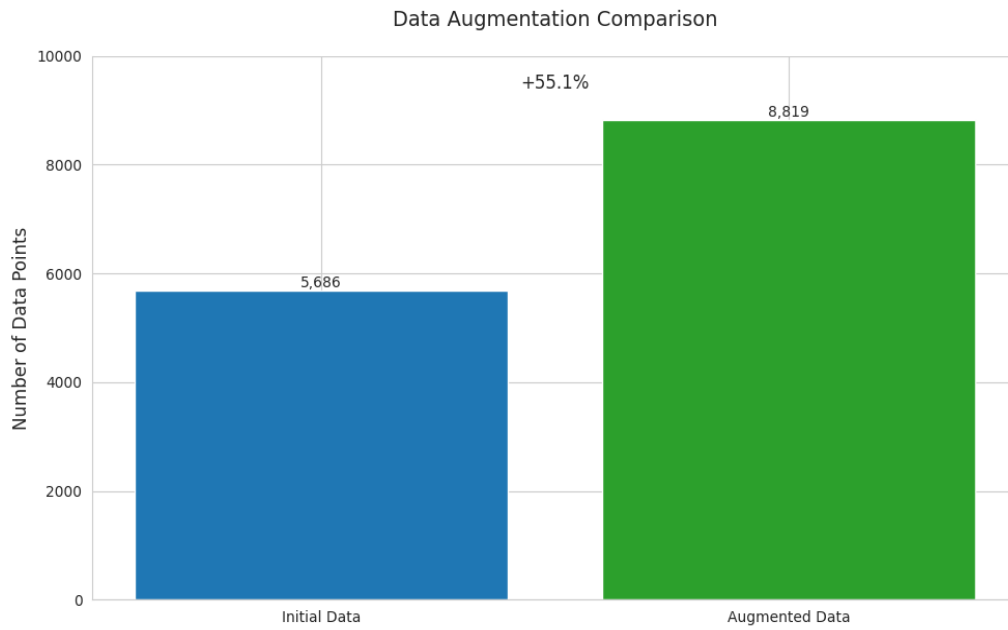


Figure 5.2: Data Comparison (After Augmentation)

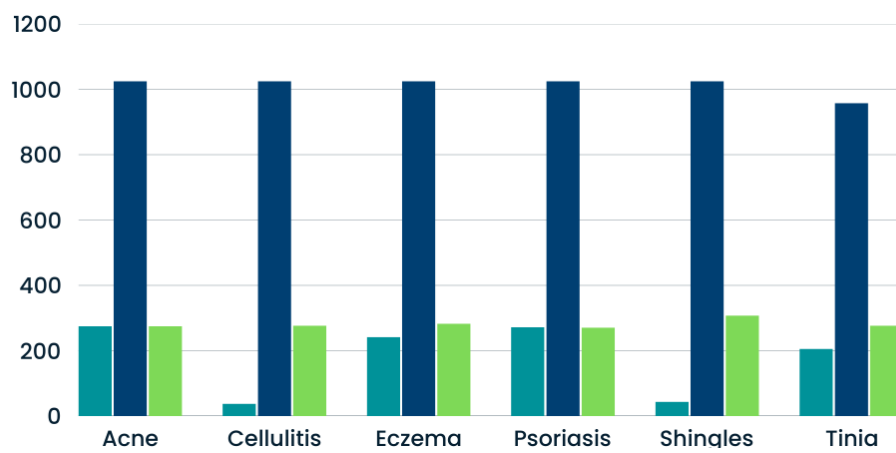


Figure 5.3: Data Splitting (After Augmentation)

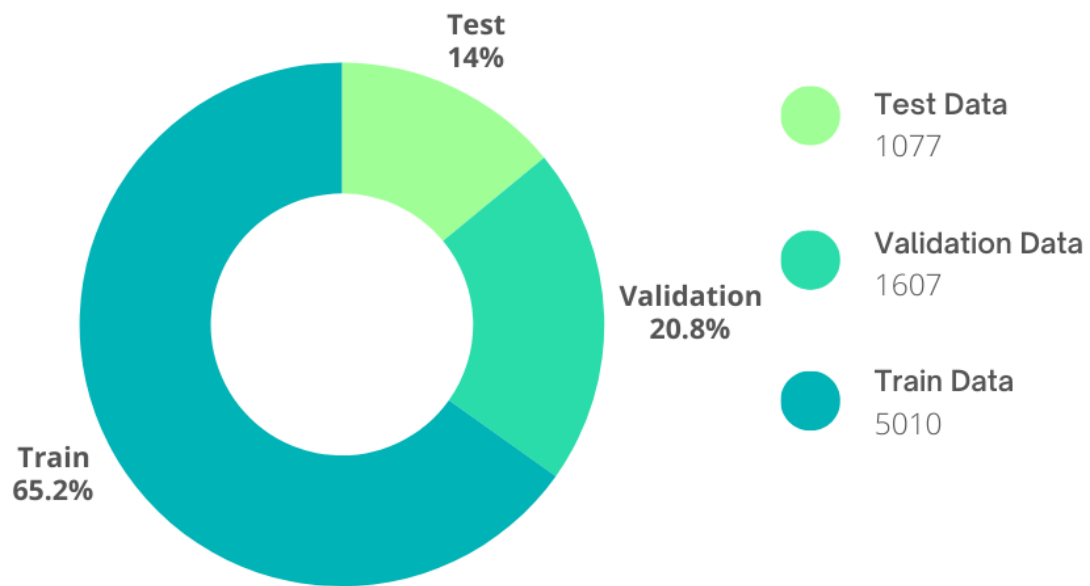


Figure 5.4: Data Splitting (All Classes Together Original)

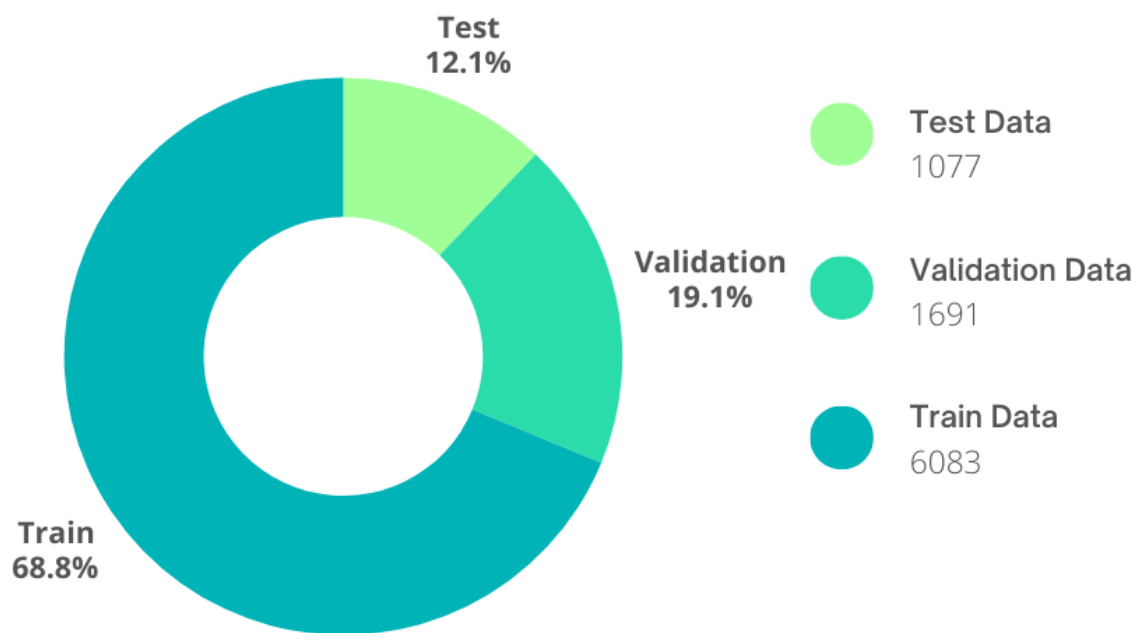


Figure 5.5: Data Splitting (All Classes Together Augmented)

	Acne	Cellulitis	Eczema	Psoriasis	Shingles	Tinea	Total
Original Data	1833	247	1609	1809	293	1369	7160
Train Original	1283	172	1126	1266	205	958	5010
Validation Original	275	31	241	271	44	205	1067
Test	275	38	242	272	44	206	1077
Validation Augment	275	277	283	271	308	277	1691
Train Augment	1025	1025	1025	1025	1025	958	6083
Augmented Data	1575	1348	1550	1568	1377	1441	8819
Increase	-258	1101	59	-241	1084	72	1817

Table 5.1: Dataset Distribution and Augmentation Summary

Chapter 6

Model Implementations And Analysis:

6.1 Deep Learning Models:

Convolutional Neural Networks (CNNs) use deep learning techniques to examine images through hierarchical convolutional layers which enable automatic detection of spatial features from edges up to complex patterns following filter and pooling applications and activation functions. Among prominent models, InceptionV3 optimizes multi-scale feature capture via parallel convolutions, ideal for detecting nuanced textures in medical images; ResNet50 uses residual blocks with skip connections to enable ultra-deep networks, mitigating vanishing gradients for robust hierarchical learning; DenseNet enhances feature reuse via dense inter-layer connections, improving efficiency in data-limited scenarios; MobileNet employs depth wise separable convolutions for lightweight, edge-device deployment without sacrificing diagnostic accuracy; Xception extends Inception with deeper separable convolutions, excelling in spatial and depthwise feature disentanglement; and VGG16, despite its simplicity and computational heft, offers reliable baseline performance via stacked 3x3 convolutions. The combination of CNNs demonstrates utility in disease discovery through their ability to pinpoint minute anatomical signs (e.g. tumors and lesions) in radiological imaging while optimizing their structural complexity for divergent clinical requirements and image data types.

The deep learning architecture Transformers started life focusing on sequential data such as text before developers adapted self-attention mechanisms to deal with visual input and help capture global context. The Vision Transformer (ViT) divides images into smaller grid sections after which it represents the tiles as sequences while applying pure attention capabilities to determine distant relationships so it becomes effective at recognizing complete graphical structures (such as expansive tissue damage). With its hierarchical shifted windows framework, Swin Transformer provides efficient coordination between local details and broad context understanding which works exceptionally well with medical image resolution matrices. The DeiT approach uses knowledge distillation techniques to train effectively with limited data since medical imaging has minimal labeled datasets available. The diagnostic accuracy of disease detection achieves new heights as transformers match CNNs through their ability to detect complex non-local patterns and scale-specific attributes and their adaptability to broad spatial relationships making them optimally suitable for professional diagnosis across radiology and pathology fields.

Model Name	Batch Size	Accuracy (%)	Epoch	Early Stop
MobileNET V3	08	74.63	100	N/A
Inception V3	08	71.31	100	N/A
DenseNET 121	08	58.40	100	6
ResNET 50	08	26.04	100	29
XceptionNet	08	53.64	100	6
VGG16	08	39.74	100	14
MobileNET V3	16	73.35	100	N/A
Inception V3	16	73.98	100	N/A
DenseNET 121	16	62.32	100	7
ResNET 50	16	26.04	100	26
XceptionNet	16	54.66	100	5
VGG16	16	48.23	100	14
MobileNET V3	32	76.42	100	N/A
Inception V3	32	70.75	100	N/A
DenseNET 121	32	59.47	100	7
ResNET 50	32	26.04	100	27
XceptionNet	32	56.82	100	6
VGG16	32	Did not Run	100	13

Table 6.1: CNN Models - Accuracy Comparison

Table 6.1 shows the comparison of the CNN models of MobileNet V3, Inception V3, DenseNet 121, ResNet 50, XceptionNet, and VGG16 on account of their performances during the training of the three different batch sizes of 8, 16, and 32 for 100 epochs. MobileNet V3 is singled out as the model exhibiting the best performance, that is, the model attaining higher accuracy over the three batch sizes with the highest being 76.42% for a batch size of 32. This indicates MobileNet V3 is a model well adapted for the task all probability due to its being lightweight and efficient, which probably makes it easy to train and also minimizes its chances of overfitting; this also makes it the best performer when making apps that rely on image detection. Inception V3 could be said to have performed well as it achieved accuracies in the range of 70.75% to 73.98%. Slightly lower than MobileNet V3, this concerns the fact that it enjoys capturing features at multiple

scales, thereby recognizing complex patterns in the data. DenseNet 121 has moderate performance in the range of 58.40% to 62.32%. It seems to benefit somehow from larger batch sizes, which is not too much to say. Since lower performance may be expected due to overfitting or slow training arising from the reuse of features in its dense connections, exceeding low performance is not worth expecting from it here, compared to MobileNet V3 or Inception V3. XceptionNet gives an accuracy between 53.64% and 56.82%, which was far from the expected value we had predicted. Maybe its intricate design is to blame, as it may need either delicate adjustment, more fine-tuning, or a more massive dataset to perform nicely. VGG16, a classic deep learning model, performs poorly, with accuracies ranging from 39.74% to 48.23%, and it fails to run at a batch size of 32. This is likely because VGG16 is a very deep and simple model, which might not be well-suited for this specific task or dataset. Finally, ResNet 50, which is known for its residual connections that help train very deep networks, surprisingly has the lowest accuracy 26.04% across all batch sizes. This could be due to overfitting, insufficient training, or a mismatch between the model's complexity and the dataset's size or complexity.

Model Name	Batch Size	Accuracy (%)	Epoch	Early Stop
Vision Transformer	08	79.39	50	6
Swin Transformer	08	79.29	50	6
Data Efficient Image Transformer	08	76.42	50	8
Vision Transformer	16	79.02	50	6
Swin Transformer	16	80.04	50	6
Data Efficient Image Transformer	16	78.92	50	6
Vision Transformer	24	80.59	50	6
Swin Transformer	24	79.92	50	6
Data Efficient Image Transformer	24	81.15	50	6
Vision Transformer	32	80.78	50	6
Swin Transformer	32	81.06	50	7
Data Efficient Image Transformer	32	80.41	50	8
Vision Transformer	64	77.99	50	6
Swin Transformer	64	84.77	50	6
Data Efficient Image Transformer	64	83.10	50	6

Table 6.2: Transformer Models - Accuracy Comparison

Table 6.2 demonstrates how Vision Transformer (ViT) Swin Transformer and Data-Efficient Image Transformer (DeiT) perform under different batch sizes while showing their accuracy levels together with their epoch counts and early stopping epochs. At a batch size of 64 Swin Transformer delivers the optimal accuracy level of 84.77% while performing better than both ViT (77.99%) and DeiT (83.10%) under similar conditions. At batch 64 ViT shows a performance decrease from its previous peak while showing accuracy growth through small batch sizes below 32. This indicates possible platform errors emerging at extremely big batch numbers. DeiT maintains accuracy growth using larger batch sizes because of its data-efficient methods including knowledge distillation techniques that boost scalability. The hierarchical architecture of Swin Transformer uses local and global attention mechanisms that enable stable performance while achieving peak results with the maximum batch size because memory restrictions decrease and feature extraction improves.

Model	Acne	Eczema		Shingles		Accuracy		Macro	Weighted
		Cellulitis	Psoriasis	Tinea				Avg	Avg
DenseNet (8)	0.83	0.50	0.71	0.70	0.88	0.70	0.79	0.79	0.76
DenseNet (16)	0.81	0.50	0.69	0.68	0.86	0.69	0.78	0.78	0.76
DenseNet (32)	0.82	0.56	0.70	0.69	0.87	0.70	0.79	0.79	0.76
Inception (8)	0.91	0.71	0.74	0.73	0.92	0.72	0.80	0.80	0.77
Inception (16)	0.91	0.72	0.74	0.73	0.92	0.72	0.81	0.81	0.77
Inception (32)	0.91	0.74	0.75	0.74	0.93	0.73	0.81	0.81	0.77
ResNet (8)	0.48	0.59	0.59	0.62	0.72	0.55	0.76	0.76	0.75
ResNet (16)	0.48	0.59	0.58	0.61	0.71	0.54	0.75	0.75	0.75
ResNet (32)	0.10	0.10	0.20	0.15	0.18	0.12	0.30	0.30	0.30
MobileNet (8)	0.93	0.82	0.75	0.68	0.90	0.72	0.81	0.81	0.81
MobileNet (16)	0.93	0.82	0.75	0.68	0.90	0.72	0.81	0.81	0.81
MobileNet (32)	0.93	0.82	0.76	0.69	0.91	0.73	0.81	0.81	0.81
Xception (8)	0.78	0.58	0.68	0.65	0.87	0.64	0.79	0.79	0.79
Xception (16)	0.79	0.59	0.69	0.66	0.88	0.65	0.80	0.80	0.80
Xception (32)	0.80	0.60	0.70	0.67	0.89	0.66	0.80	0.80	0.80
VGG16 (8)	0.87	0.60	0.70	0.67	0.89	0.66	0.80	0.80	0.80
VGG16 (16)	0.70	0.60	0.70	0.67	0.89	0.66	0.80	0.80	0.80
VGG16 (32)	0.69	0.61	0.71	0.68	0.90	0.67	0.81	0.81	0.81
ViT (8)	0.95	0.79	0.69	0.71	0.95	0.66	0.79	0.79	0.76
ViT (16)	0.96	0.84	0.76	0.72	0.94	0.77	0.81	0.83	0.81
ViT (32)	0.95	0.82	0.73	0.73	0.95	0.77	0.80	0.82	0.80
ViT (64)	0.94	0.86	0.76	0.78	0.94	0.78	0.83	0.85	0.83
Swin (8)	0.91	0.82	0.74	0.68	0.95	0.66	0.76	0.79	0.76
Swin (16)	0.95	0.84	0.76	0.72	0.94	0.77	0.81	0.83	0.81
Swin (32)	0.95	0.82	0.73	0.73	0.95	0.77	0.80	0.82	0.80
Swin (64)	0.97	0.87	0.76	0.78	0.94	0.78	0.83	0.85	0.83
DeiT (8)	0.91	0.82	0.74	0.68	0.95	0.66	0.76	0.79	0.76
DeiT (24)	0.95	0.84	0.76	0.72	0.94	0.77	0.81	0.83	0.81
DeiT (32)	0.95	0.82	0.73	0.73	0.95	0.77	0.80	0.82	0.80
DeiT (64)	0.97	0.87	0.76	0.78	0.94	0.78	0.83	0.85	0.83

Table 6.3: F1 Scores Across Different Models

6.2 Comparative Analysis:

Deep learning model performance depends fundamentally on batch size, specific architectural choices, and proper regularization techniques per a comparative evaluation of transformer-based versus CNN-based models. Large batch processing benefits Swin Transformer because its hierarchical local-global attention system allows for better memory management and feature extraction capabilities, leading to superior scalability results (84.77% accuracy at batch 64). The accuracy of Vision Transformer (ViT) reaches its maximum of 80.78% at batch 32 but dips sharply when using batch 64 possibly due to limited regularization and gradient noise control in large experimental conditions. Data Efficient Image Transformer (DeiT) delivers smooth accuracy improvements with increasing batch size (81.15% accuracy at batch 24) as a result of its optimized training performance. Early stopping patterns further differentiate these models: The DeiT model demonstrates swift convergence during training (8 epochs at batch 8) probably because it is more susceptible to overfitting although ViT and Swin exhibit better early stability as evidence of their resilient architecture. The responsiveness of CNN models to different batch size configurations shows diversity when compared to other models. Despite their structure simplicity and anti-overfitting characteristics MobileNet V3 with Inception V3 operate robustly at peak accuracy levels throughout their training process. DenseNet 121 and XceptionNet show minor improvements with larger batches due to enhanced generalization but require early stopping (e.g., epoch 6 for DenseNet) to mitigate overfitting. Both ResNet 50 as well as VGG16 show limited progress at 26.04% accuracy for ResNet while causing memory consumption issues at batch 32 because they fail to efficiently utilize bigger batches without adapting their design patterns. These disparities underscore the necessity of model-specific tuning: Characteristic attention functioning benefits transformers Swin along with DeiT but MobileNet V3 CNN maintains its lead by being efficient with parameters. To maximize the benefits of larger batch processing, underperforming models ResNet and VGG16 require gradient noise reduction memory optimization, or advanced regularization techniques including dropout and weight decay. Ultimately, the interplay between architecture, batch dynamics, and training protocols dictates model success, emphasizing the need for customized optimization to get a more accurate result from both language models.

The superior architecture along with efficient training strategies coupled with large batch scalability made transformer models better than CNNs for skin disease detection in medical imaging applications. The hierarchical attention architecture of transformers like Swin implements local window-based attention features to detect fine textures yet also incorporates global context modeling functions to identify distribution patterns among lesions — capabilities required to diagnose diverse skin conditions across different lesion dimensions. The precision of VGG16 at 48.23% and ResNet 50 at 26.04% stands in contrast to Mixer-based Swin Transformers as the convolutional layers of those networks have rigid structures and fixed receptive fields that limit their ability to process irregular lesion boundaries and long-range dependencies. The optimized memory usage enabled by Swin’s shifted window mechanism delivered peak accuracy of 84.77% during batch 64 while DeiT’s knowledge distillation method improved accuracy to 83.10% by using pre-trained teachers and remaining more data efficient than XceptionNet (56.82%). When using transformers to process large batches they excel at gradient regularization while Swin Transformer’s hierarchical structure minimizes redundancy and ViT suffers accuracy loss at batch 64 (77.99%) because of inadequate regularization strategies. This challenge remains difficult for ResNet 50 to resolve because of its excessive parameteriza-

tion and vanishing gradient issues. CNNs like DenseNet 121 (62.32%) and MobileNet V3 (76.42%) show moderate performance but are limited by architectural constraints: The mobile application efficiency and the lightweight performance of MobileNet V3 also limit its ability to match transformer power and DenseNet’s connection architecture struggles with overfitting between epochs 6 and early-stop activation. Transformers demonstrate their ability to resist overfitting naturally since they maintain training stability without requiring aggressive early stopping methods. Transformers excel over CNNs because of their dual capabilities to scale operational size effectively (DeiT reaches 81.15% at batch 24) and efficiently consume memory while developing dynamic features that adapt better to medical tasks that require precision on heterogeneous data compared to CNNs which are restricted by inflexible structure and memory limitations (VGG16 stops operation at batch 32) and suboptimal batch utilization behaviors. New combinations of CNN and Transformer functionality to boost performance show evidence that Transformer systems yield superior results in challenging medical image processing tasks with limited available data.

6.3 Custom Model:

6.3.1 Introduction to TransNet-X

The innovative TransNet-X hybrid model unites Convolutional Neural Networks (CNNs) with Transformer-based architectures for performing skin disease detection. The classification of skin diseases requires achieving both optimal measurement precision and operational speed. MobileNet V3 demonstrates high operational efficiency along with light architecture but it fails to detect complex skin disease patterns and long-distance dependencies in medical diagnosis. The Transformer-based Swin Transformer succeeds at spotting complex patterns yet it needs large computational resources to operate.

The solution provided by TransNet-X solves the accuracy-complexity dilemma through knowledge distillation implementation. By using the Swin Transformer as its teacher model the system extracts advanced high-level features which monitor detailed image characteristics within a broad global network structure. The student network MobileNet V3 uses rich features extracted by the teacher to both reproduce teacher results and keep its compact design for efficient execution. The combination of CNN efficiency with Transformer model capabilities allows TransNet-X to reach outstanding results for skin disease detection.

6.3.2 Why choosing MobileNetV3 & Swin Transformer-

Convolutional Neural Networks (CNNs) - MobileNet V3	
Accuracy	76.42%
Overview:	MobileNet V3 is a lightweight CNN designed for mobile and embedded devices. It is known for its speed and efficiency, making it suitable for real-time applications. The detection system performs poorly when analyzing intricate and delicate disease patterns present in Psoriasis and Eczema because its capability to identify extended dependencies and refined features is restricted.
Strengths:	Efficiency: The fast inference operation combined with low memory requirements makes MobileNet V3 suitable for devices with restricted resources. Speed: It is quick, making it ideal for applications where fast processing is essential.
Limitations:	Limited Feature Extraction: The device cannot capture essential skin details needed for the precise identification of complicated skin diseases. Lack of Long-Range Dependency Modeling: Because of its limitations, MobileNet V3 tends to overlook essential spatial relationships between image areas that should be evaluated for diagnostic purposes.

Table 6.4: Performance Analysis of MobileNet V3

Transformers - Swin Transformer	
Accuracy	84.77%
Overview:	The Swin Transformer presents itself as a superior architecture designed for tasks needing deep image relationship understanding. Swin Transformer demonstrates superior performance at perceiving delicate image elements together with distant relationship patterns that establish crucial conditions for precise skin disease identification. The self-attention mechanism of Swin Transformer directs its attention toward significant image features that exist throughout different sections of the image.
Strengths:	Long-Range Dependencies: It is capable of capturing global relationships between distant regions of the image. Better Feature Extraction: The self-attention mechanism allows it to focus on the most relevant parts of the image, improving classification performance.
Limitations:	Computationally Expensive: The complexity of the model leads to reduced processing speed together with increased resource demands. Slower Inference: Due to its strong attention mechanism, the processing requires a longer time than what MobileNet V3 requires for lightweight processing.

Table 6.5: Performance Analysis of Swin Transformer

6.3.3 Details of TansNet-X (Hybrid Model):

Overview-

TransNet-X implements a teacher-student framework that unites CNNs and Transformers effectively. The Swin Transformer functions as the teacher system that delivers strong feature analysis together with long-range relationship handling abilities. Through knowledge distillation, MobileNet V3 retrieves information from Swin Transformer which allows it to conduct computational operations at reduced costs while preserving performance levels. TransNet-X uses this combination of efficiency-accuracy balancing to achieve its results. Accuracy: 80.5

Working Mechanism of TransNet-X Works:

- **Teacher-Student Setup:** The Swin Transformer functions as a teaching model to process images and obtain advanced complex features. The features of the Swin Transformer process fundamental information for both disease identification and classification.
- **Knowledge Distillation:** MobileNet V3 (student) adopts the teaching approach from its teacher model through pattern output analysis. MobileNet V3 can access the teacher's deep learning through this process while remaining free from executing the sophisticated operations of Swin Transformer.
- **Loss Function:** The training process uses two components:
 - **Cross-Entropy Loss:** This helps in correctly classifying skin diseases.
 - **Distillation Loss (KL Divergence):** This encourages the student model to align its predictions with the teacher's, ensuring that the student learns to replicate the teacher's decision-making process.

Inference: After training, the student model (MobileNet V3) is used for inference, ensuring that predictions are made quickly and efficiently while maintaining high accuracy.

TransNET-X Architecture

1. Dataset Loading and Preprocessing

- **Transforms:**
 - **Resizing:** Images are resized to **224x224** pixels.
 - **Normalization:** The images undergo normalization using standard mean and standard deviation values which are commonly applied to ResNet and Swin Transformer pretrained models.
- **Dataset Loading:**
 - The `datasets.ImageFolder` function is used to load the dataset from the specified directories for **training**, **validation**, and **testing**.
 - **DataLoader:** This utility creates an iterable dataset from the input data by dividing it into batches and manages both shuffling and multi-threaded loading operations.

2. Model Architecture

Teacher Model: Swin Transformer The **Swin Transformer** (Teacher Model) is used to generate soft targets (predictions) that will guide the student model.

- **Patch Embedding:** The input image is split into patches, which are then embedded into a feature space (768-dimensional in this case).
- **Swin Transformer Blocks:**
 - The model uses multiple **Swin Transformer blocks**, which are based on **shifted window attention**. This allows the model to focus on local patches of the image while also maintaining the ability to capture global relationships between features.
 - Each block applies **multi-head self-attention** within local windows and **feed-forward networks (FFN)**.
- **Classification Head:** A linear layer (`nn.Linear(768, num_classes)`) transforms the last output to generate class prediction results.

Student Model: MobileNetV3 The MobileNetV3 (Student Model) delivers light processing with great efficiency thus making it ideal for all mobile and edge-based applications. This architecture receives enhancement through added layers together with batch normalization procedures which enhance overall performance and generalization capability.

- **Initial Convolution Layer:** A **3x3 convolution** layer with stride 2 reduces the image size from 224x224 to 112x112.
- **Bottleneck Blocks:**
 - The model uses **depth-wise separable convolutions**, Each of these stages begins with depthwise convolutions that operate on single input channels before uniting the produced results through pointwise convolutions.
 - Each block is followed by **Squeeze-and-Excitation (SE) blocks**, which help recalibrate the feature maps.
- **Global Average Pooling:** After the bottleneck blocks, the model uses **Global Average Pooling**, which reduces each channel to a single value.
- **Fully Connected Layers:**
 - The features learned by the network are refined through two added fully connected layers. After the layers come ReLU activations which are paired with dropout layers to reduce overfitting.
 - The final layer (`nn.Linear(512, num_classes)`) outputs the class predictions.

3. Knowledge Distillation

The core of this hybrid model is **Knowledge Distillation**. The student model learns from authentic labels as well as relaxed targets which come from the teacher model. During learning the student attempts to match the probabilities the teacher produces as it learns to correctly classify items

Teacher Predictions: The teacher model (Swin Transformer) generates soft targets using softmax probabilities.

Student Predictions: The student model (MobileNetV3) generates its logits (raw class scores).

The **distillation loss** is calculated by combining:

- **KL Divergence Loss:** Measures the difference between the softmax outputs of the teacher and the student, encouraging the student to mimic the teacher’s predictions.
- **Cross-Entropy Loss:** Standard classification loss that compares the student’s predictions to the true labels.

The total loss is:

$$\text{Loss} = \alpha \times \text{KL Divergence Loss} + (1 - \alpha) \times \text{Cross-Entropy Loss}$$

Where T is the temperature (controls the softness of the targets), and α is the weight that balances the two loss components.

4. Training Process

The model is trained using the following steps:

1. Optimization:

- **SGD** optimizer is used with a learning rate of 0.01, momentum of 0.9, and weight decay for regularization.
- **Cosine Annealing Learning Rate Scheduler** adjusts the learning rate during training.

2. Mixed Precision Training: `torch.cuda.amp.GradScaler` is used to enable **mixed precision training**, which speeds up training and reduces memory usage without compromising model accuracy.

3. Loss and Accuracy Calculation:

- During each epoch, the loss and accuracy are calculated for both the training and validation datasets.
 - **Backpropagation:** The loss is backpropagated through the student model, and the weights are updated using the optimizer.
- ### 4. Early Stopping: `EarlyStopping` monitors the validation loss and stops training early if the validation loss does not improve for a given number of epochs (patience=10).

5. Model Evaluation

After training, the model’s performance is evaluated:

- **Evaluate Loss:** Through the `evaluate_loss()` function the loss gets calculated by comparing predicted outputs against true labels using Cross-Entropy Loss on the validation dataset.

- **Evaluate Accuracy:** Accuracy computation through the `evaluate_model()` function depends on matching predicted class with actual class labels. It calculates accuracy by totaling correct predictions while splitting the result by total samples.

6. Summary of the Architecture

- **Swin Transformer (Teacher):** A pretrained transformer-based model that generates **soft targets** for training the student model.
- **MobileNetV3 (Student):** A lightweight convolutional model trained with **knowledge distillation** to match the teacher's soft targets while learning the true labels.
- **Knowledge Distillation:** The student model learns from both the teacher's soft predictions (via KL Divergence) and the true labels (via Cross-Entropy Loss).
- **Optimization and Training:** The model is trained with the **SGD optimizer**, using **mixed precision** for efficiency and **early stopping** to prevent overfitting.

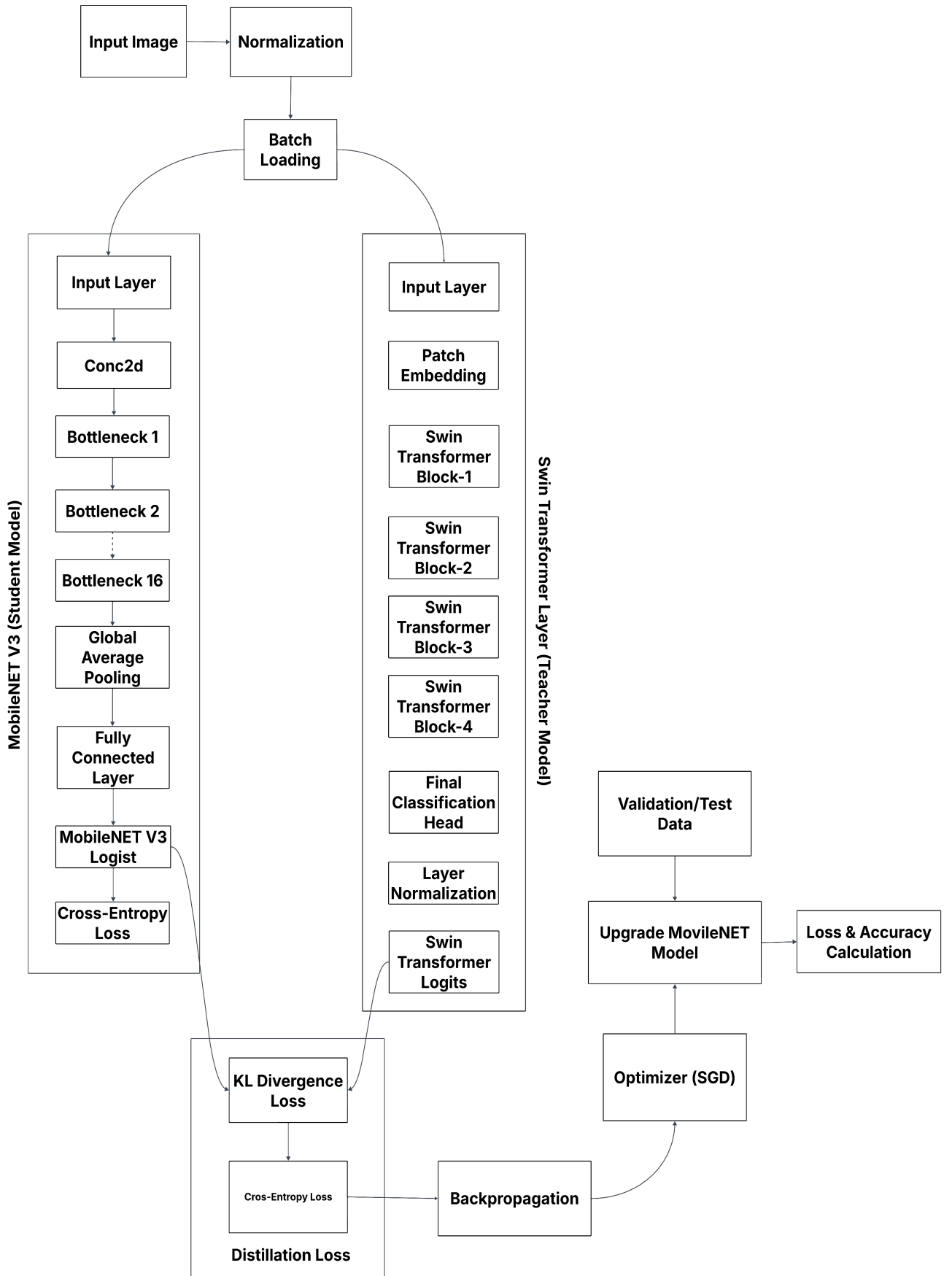


Figure 6.1: Model Architecture of TransNet-X

Chapter 7

Results:

7.1 CNN Models Result:

7.1.1 DenseNet

DenseNet Results Description: A batch size of 8 achieves the highest accuracy but also the highest loss, indicating potential overfitting even in the presence of an early stop function. Batch sizes 16 and 32 show lower accuracy and loss, with 32 having the lowest loss but also the lowest accuracy, suggesting a trade-off between stability and performance.

- DenseNet Batch-32

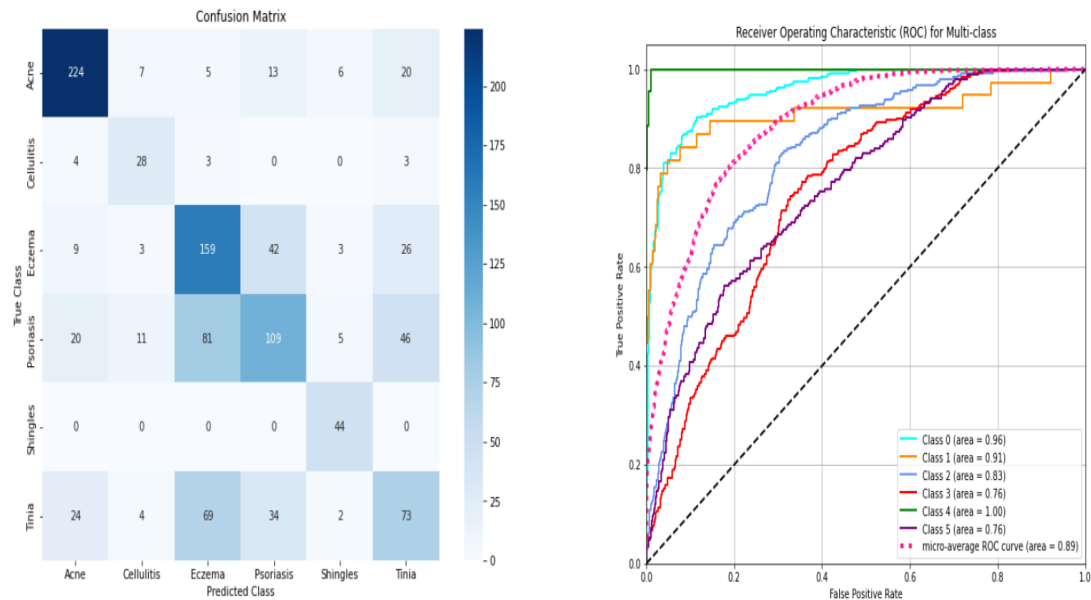


Figure 7.1: DenseNet Batch-32 Confusion Matrix (Left) and ROC Curve (Right)

7.1.2 InceptionV3

InceptionV3 Results Description: Research shows that Inception V3 consistently demonstrates classification accuracy regardless of the batch sizes tested (8, 16, 32) because its AUC values exceed 0.85 while Shingles reaches 1.00. The accuracy achieved with batch size 8 shows some superiority for particular classes, yet training efficiency improves most effectively when using batch sizes 16 or 32, which produce nearly equivalent prediction outcomes. Batch size 8's accuracy performance is slightly better, yet batch sizes 16 and 32 provide a balance of performance quality with computational efficiency.

- InceptionV3 Batch-32

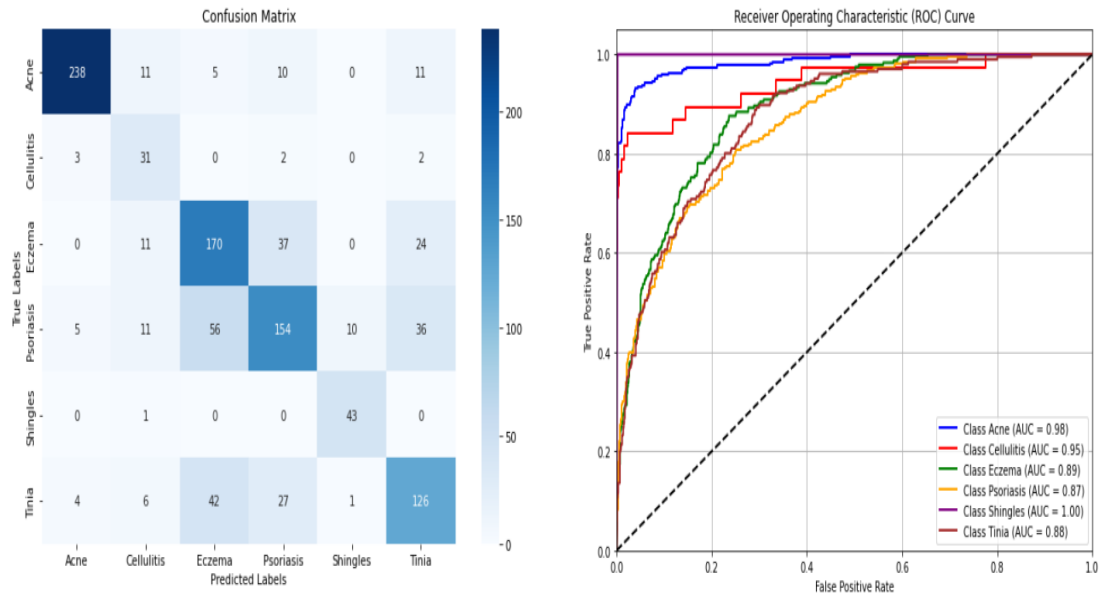


Figure 7.2: InceptionV3 Batch-32 Confusion Matrix (Left) and ROC Curve (Right)

7.1.3 MobileNet

MobileNet Results Description: The MobileNet models operate efficiently across various batch sizes from 8 to 32 to deliver AUC results higher than or equal to 0.87 for every class. The AUC values indicate Shingles provide perfect classification while Psoriasis exhibits the lowest performance (AUC=0.87–0.89). Smaller batch sizes including 8 improved diagnostic outcomes in some specific classes such as Acne and Tinia while the diagnostic performance average (micro-average AUC=0.96) remained unaffected by such changes.

- MobileNet Batch-32

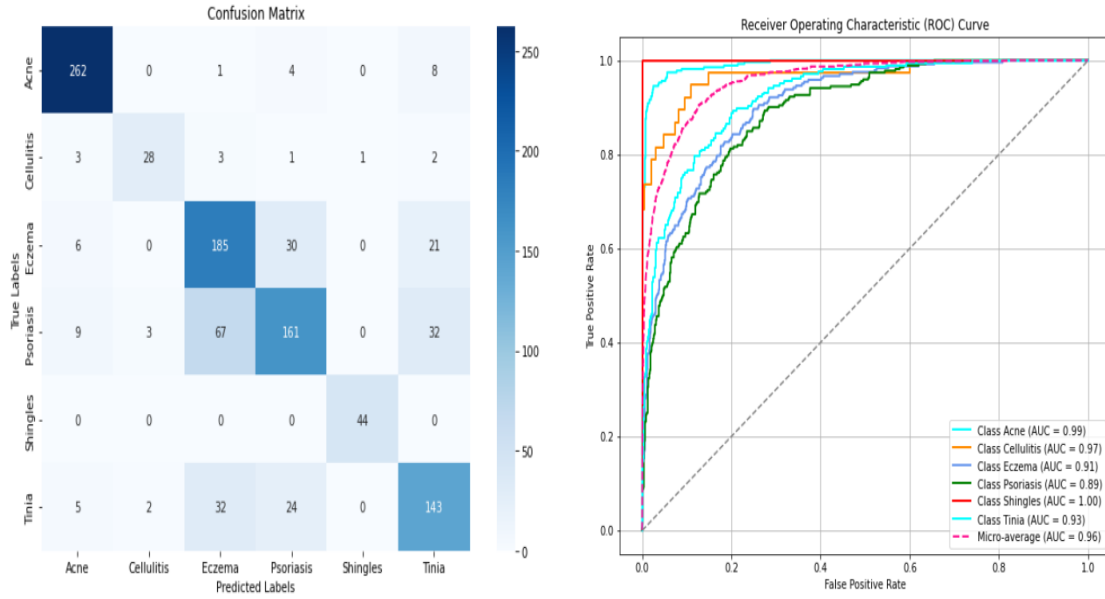


Figure 7.3: MobileNet Batch-32 Confusion Matrix (Left) and ROC Curve (Right)

7.1.4 Xception

Xception Results Description: The Xception model delivers steadfast high-performance identification for most of the classes through all batch size configurations (AUC=0.95–0.99). Moreover, it shows perfect classification results (AUC=0.99 at batches 16 and 32) for psoriasis. The AUC score increases step by step from eczema: 0.71→0.72→0.75 and Shingles. Model robustness indicators show an upward trend in the micro-average AUC score which begins at 0.86 at batch 8 and reaches its peak of 0.88 at batch 32. Data from Class 2 demonstrates no variation in reliability (AUC=0.80 across all batches) and the model’s stability grows slightly better with bigger batch sizes leading to optimal performance between class-specific accuracy and systematic consistency.

- Xception Batch-32

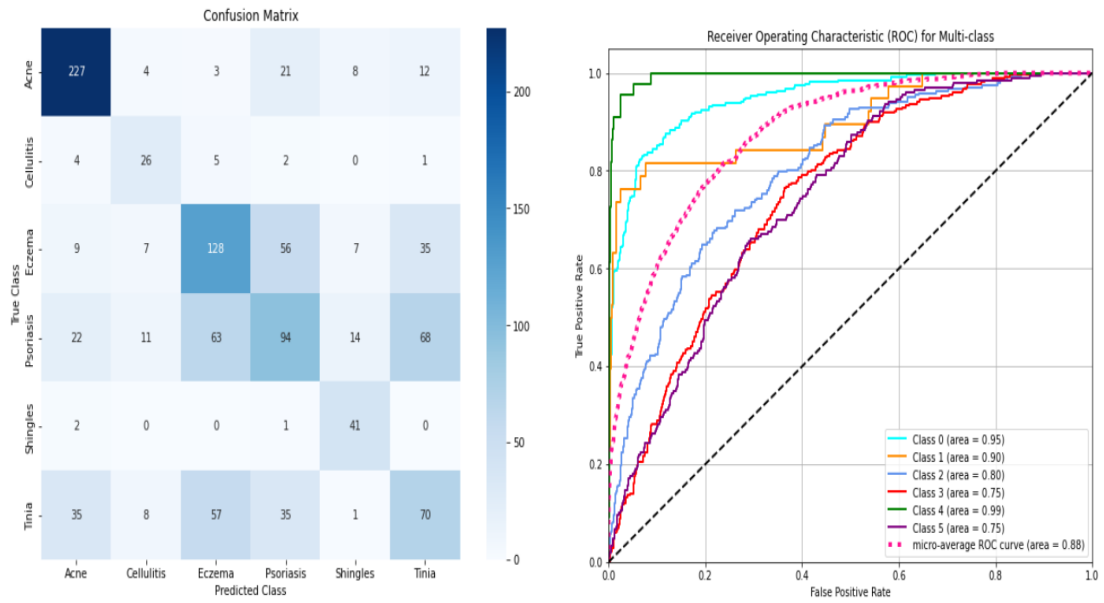


Figure 7.4: Xception Batch-32 Confusion Matrix (Left) and ROC Curve (Right)

7.1.5 VGG16

VGG16 Results Description: The VGG16 model displays variable class results when operating with different batch sizes between 8 and 16. ACNE and SHINGLES scores improve under Batch 16 through better precision along with recall enhancement (both F1 scores rise from 0.58 to 0.69 and from 0.50 to 0.72) but Tinia completely vanishes (F1 score drops from 0.44 to 0.00) due to class imbalance. F1 scores of both Psoriasis and Eczema show variable results throughout the experiment (F1=0.04→0.21 for Psoriasis and F1=0.28→0.50 for Eczema). Data distribution and model applicability appear to cause performance challenges because metric values show only marginal growth (0.36 F1 to 0.42 F1 and 0.39 accuracy to 0.46). The system’s performance may benefit from dealing with class-specific distribution issues while studying different network designs.

- VGG16 Batch-16

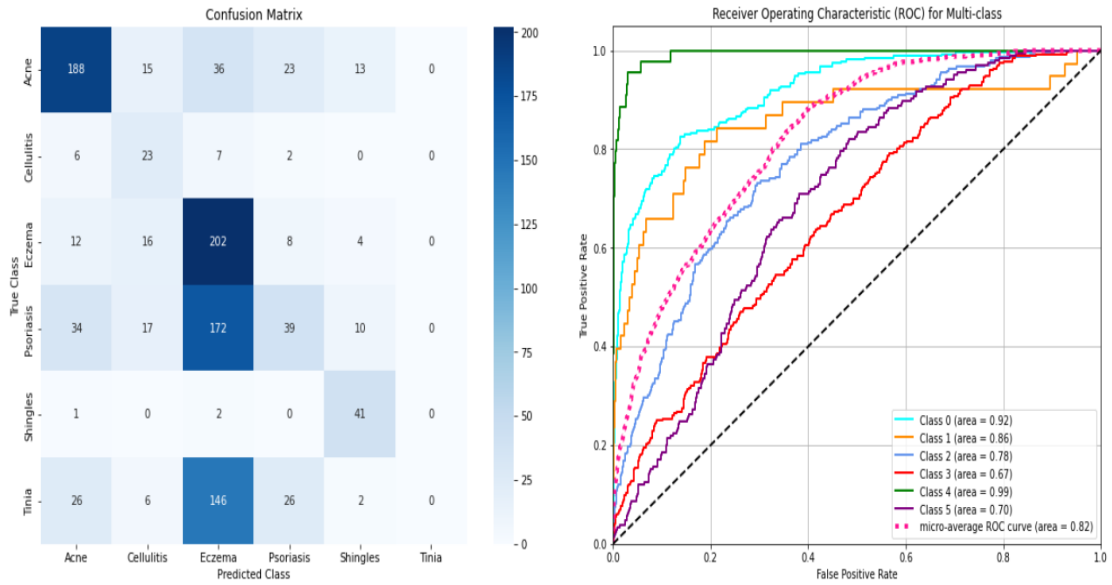


Figure 7.5: VGG16 Batch-16 Confusion Matrix (Left) and ROC Curve (Right)

7.1.6 ResNet50

ResNet50 Results Description: The results of resnet50 appear to be horrendous as it is only able to properly detect one class and fumbles for the rest, batch size seems to have no significant effect on the results.

- ResNet Batch-16

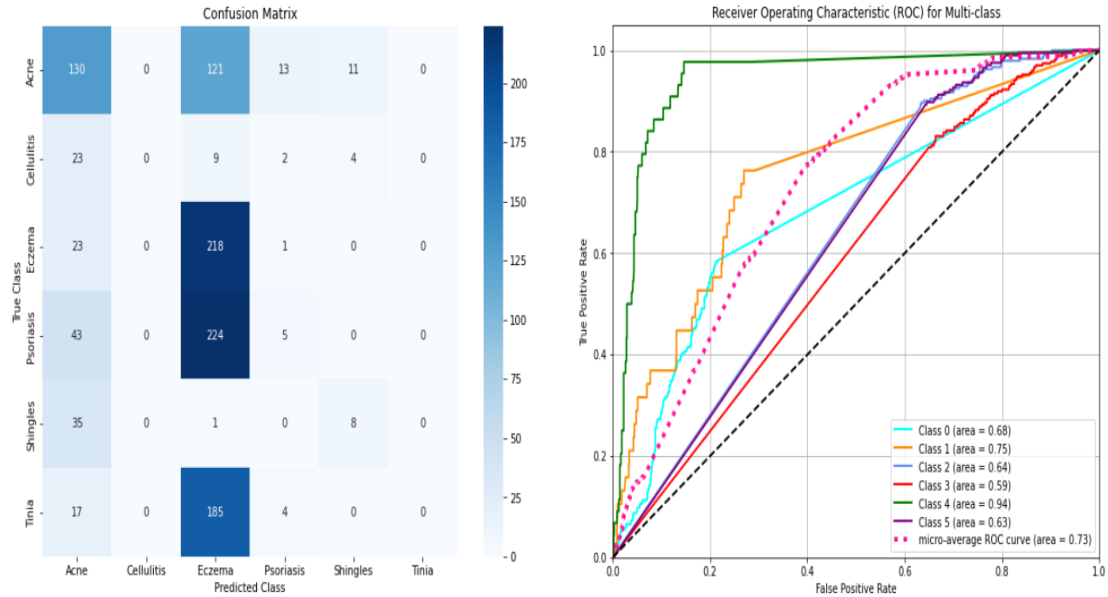


Figure 7.6: ResNet Batch-16 Confusion Matrix (Left) and ROC Curve (Right)

7.2 Transformer Models Result:

7.2.1 Deit

Deit Results Description: Deit performs very well across all of the batch sizes, and it seems to provide the best output when the batch size is 32. However, it struggles to differentiate between Eczema and Psoriasis.

- Deit Batch-32

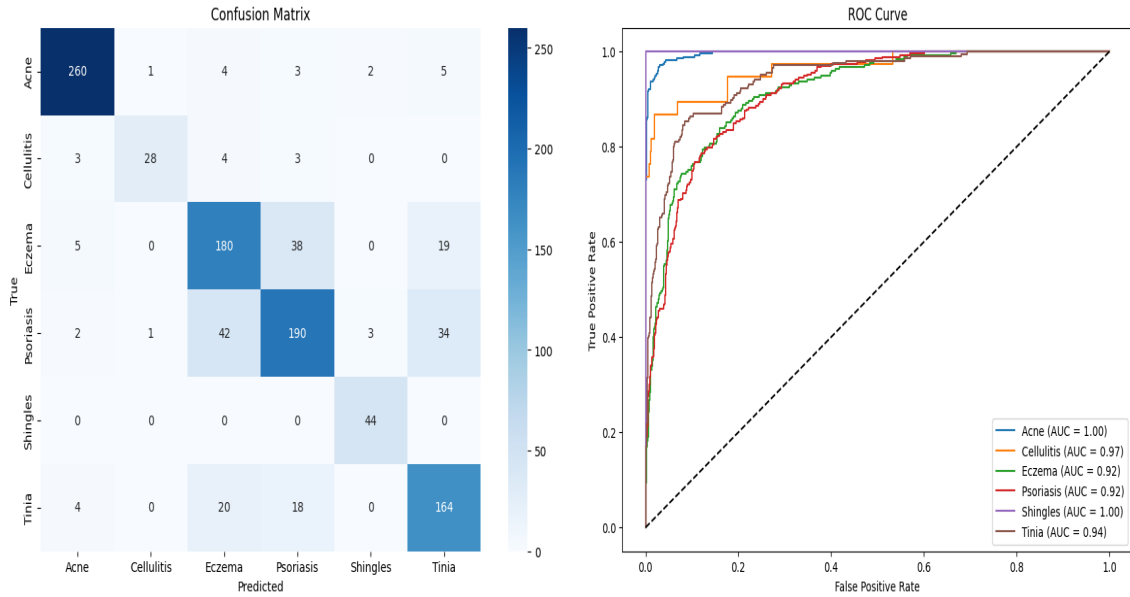


Figure 7.7: Deit Batch-32 Confusion Matrix (Left) and ROC Curve (Right)

- Deit Batch-64

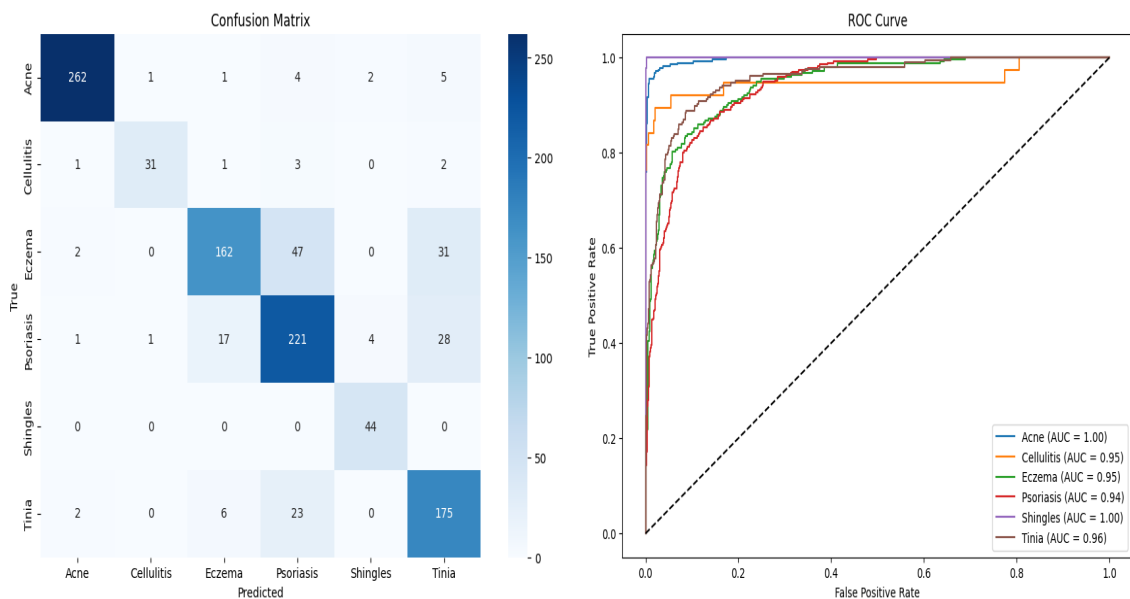


Figure 7.8: Deit Batch-64 Confusion Matrix (Left) and ROC Curve (Right)

7.2.2 Swin

Swin Results Description: Swin similar to Deit performs well across all batch sizes and gives the beat output when batch size is the greatest, however, it has a better output than Deit as Swin gives a max accuracy of 85% while triggering the early stop function.

- Swin Batch-32

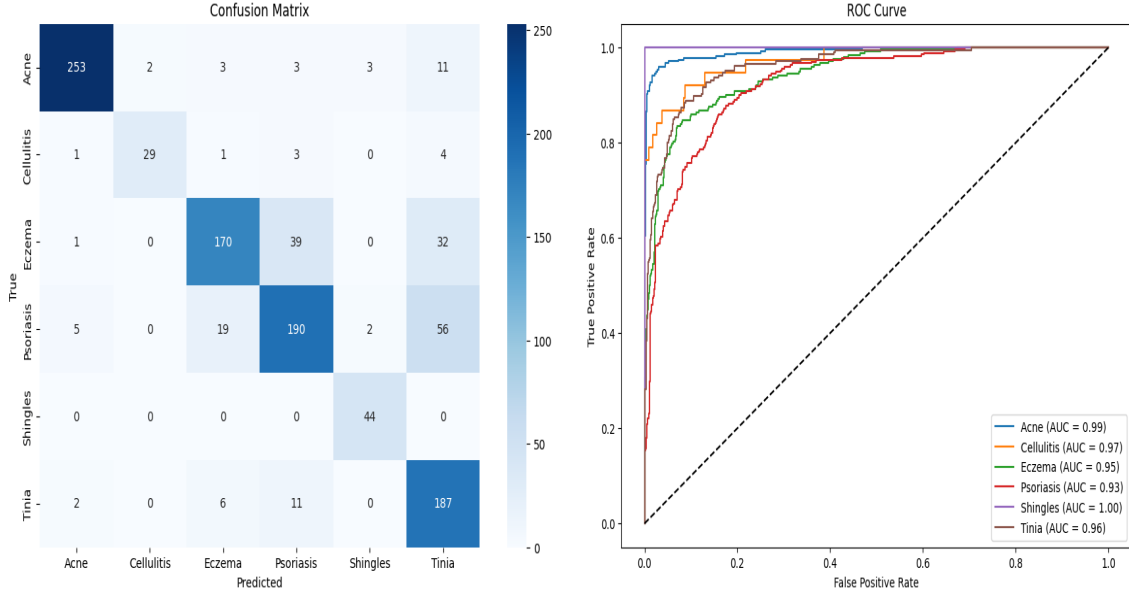


Figure 7.9: Swin Batch-32 Confusion Matrix (Left) and ROC Curve (Right)

- Swin Batch-64

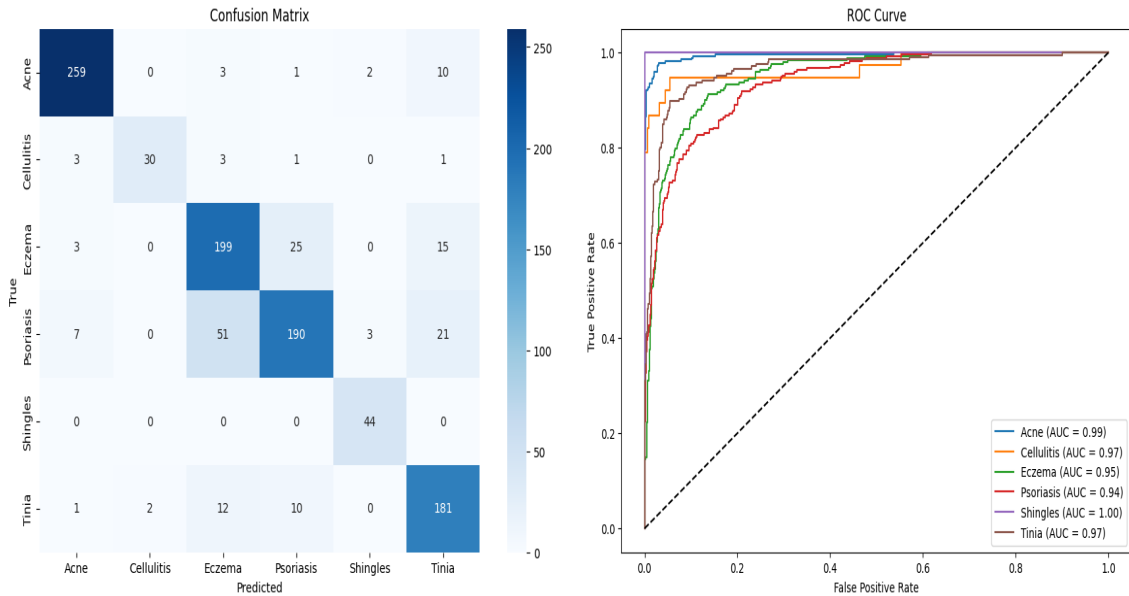


Figure 7.10: Swin Batch-64 Confusion Matrix (Left) and ROC Curve (Right)

7.2.3 ViT

ViT Results Description: the performance of ViT is on par with DeiT with not much difference in their results across the batch sizes or in their outputs as per the ROC curve in F1 scores. It also triggers the early stop function relatively early like Swin.

- ViT Batch-32

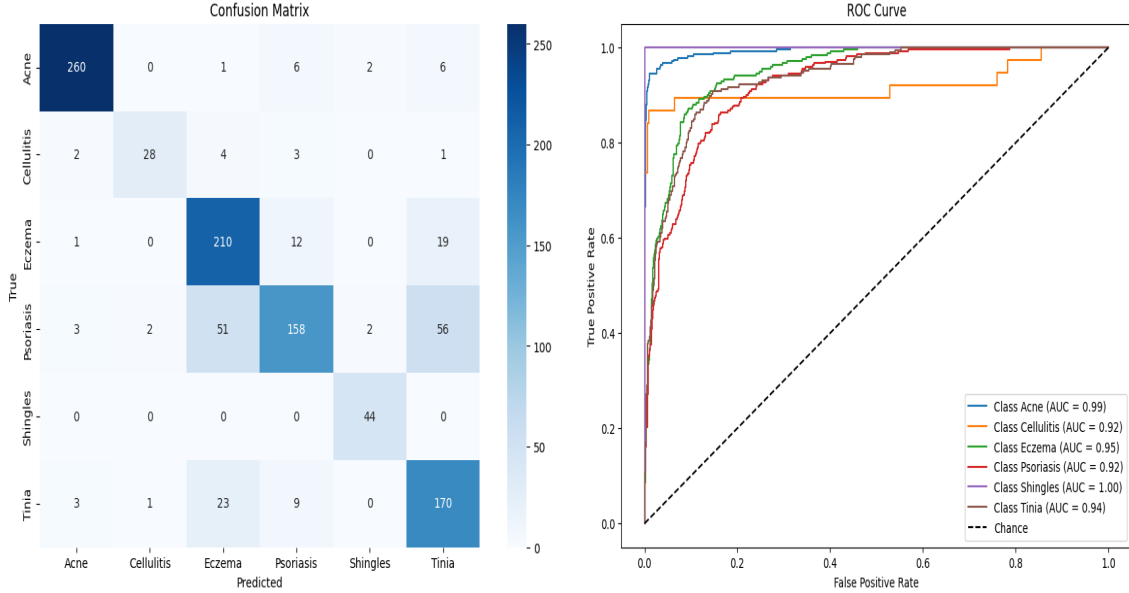


Figure 7.11: ViT Batch-32 Confusion Matrix (Left) and ROC Curve (Right)

- ViT Batch-64

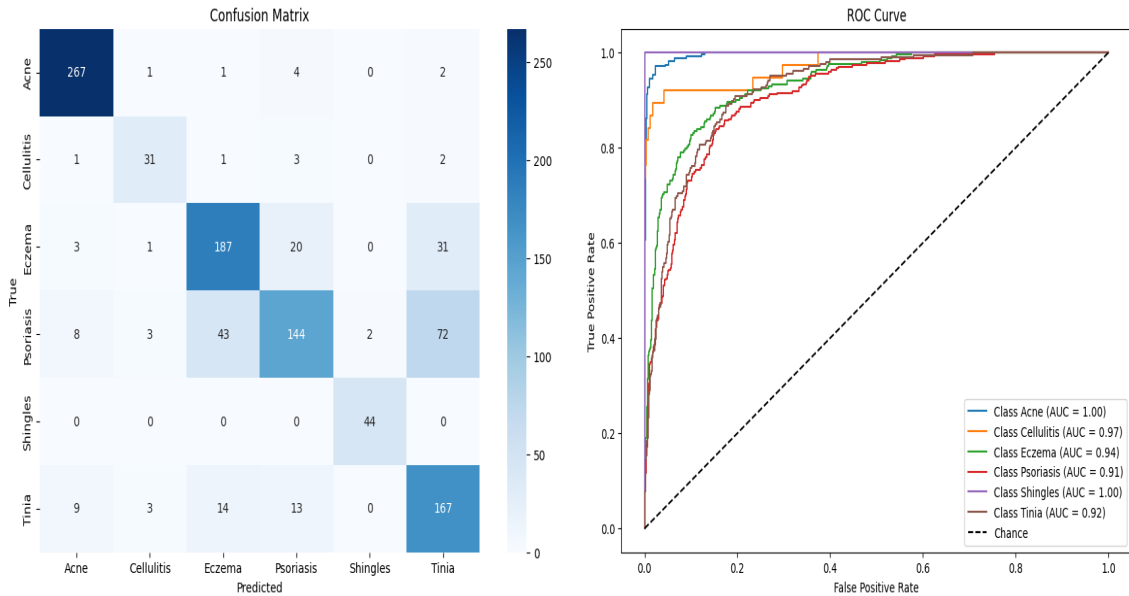


Figure 7.12: ViT Batch-64 Confusion Matrix (Left) and ROC Curve (Right)

7.3 Custom Model Results:

7.3.1 The Experimental report of TransNet-X is given below:

- Confusion Matrix -

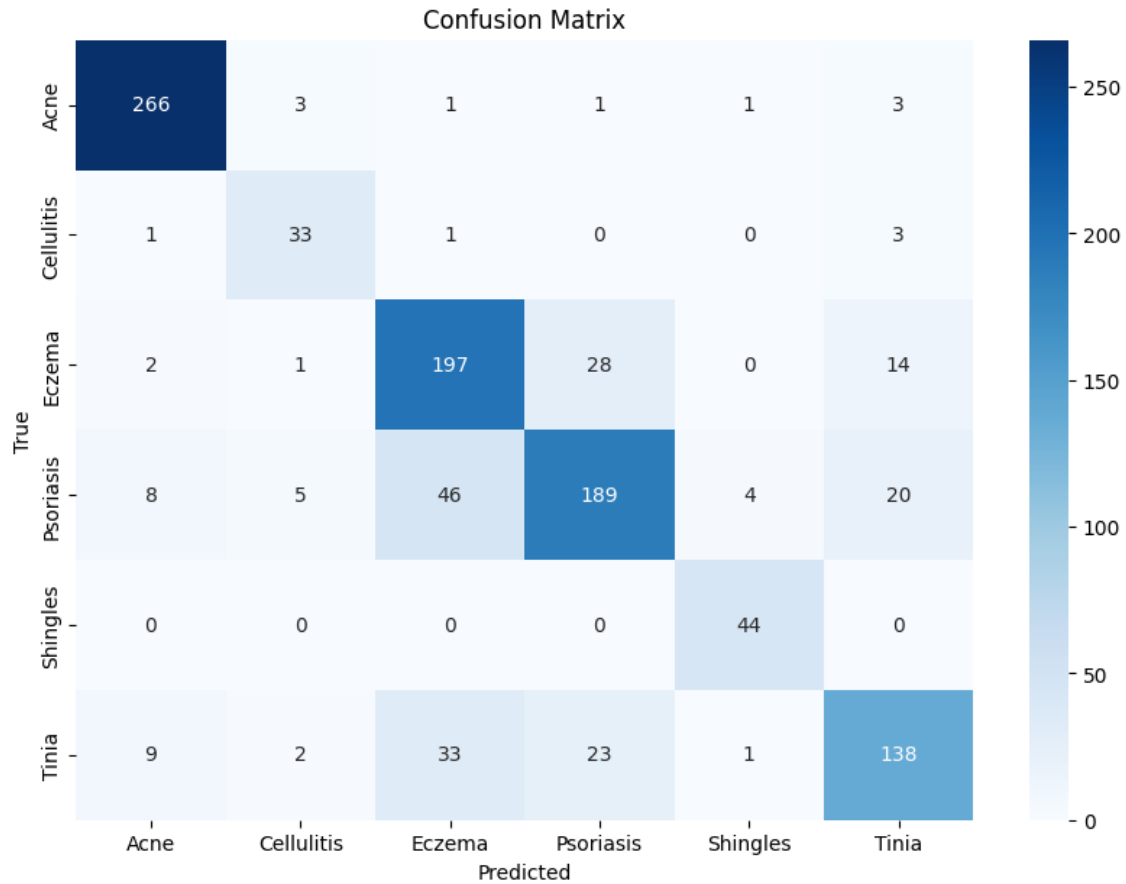


Figure 7.13: **TransNet-X (Custom Model) Confusion Matrix**

The confusion matrix demonstrates performance measurement of six skin types including Acne, Cellulitis, Eczema, Psoriasis, Shingles, and Tinea. The model shows strong performance through diagonal values mainly in Acne (266 true positives) Eczema (197) and Shingles (44). The model displays both accurate and inaccurate results in its classification process: Acne receives confusion from four alternative diagnoses including Cellulitis and three other conditions (9 total errors) whereas Cellulitis (33 correct) shows only three minor misdiagnoses to Acne, Eczema and Tinea (5 errors). Eczema (197 correct) faces notable confusion with Psoriasis (28) and Tinea (14). Both Psoriasis (189 correct) and Tinea (138 correct) experience major misidentification by healthcare staff because they repeatedly confuse Eczema (46 or 33) or Psoriasis (23). Shingles (44 correct) show minimal errors, with only 4 misclassifications to Psoriasis. Acne and Cellulitis and Shingles classifications remain suitable while the model displays limitations when attempting to classify Psoriasis and Tinea primarily because of erroneous assignments between Eczema and Psoriasis as well as Eczema and Tinea classification errors. These results indicate the need to improve the distinction between these skin class categories.

- ROC Curve -

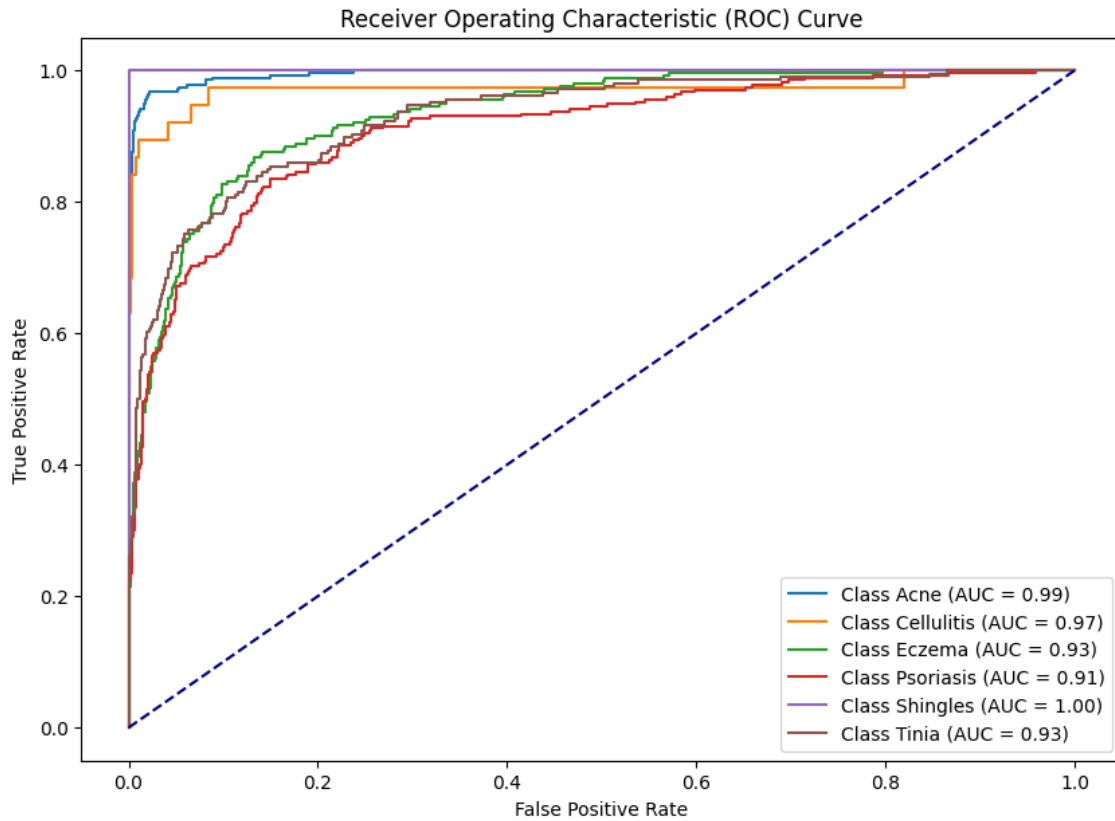


Figure 7.14: **TransNet-X (Custom Model) ROC Curve**

The ROC Curve above shows TransNet-X's performance in classifying different skin diseases. The x-axis represents the False Positive Rate (FPR), and the y-axis represents the True Positive Rate (TPR).

- Class Acne (AUC = 0.99) and Class Cellulitis (AUC = 0.97) have high AUC values, indicating excellent performance in distinguishing these diseases from others.
- Class Shingles (AUC = 1.00) has a perfect score, meaning the model can perfectly classify Shingles.
- Class Psoriasis (AUC = 0.91), Class Eczema (AUC = 0.93), and Class Tinea (AUC = 0.93) also show good performance, though with slightly lower AUCs compared to the top-performing classes.

The model generally performs well across all classes, with AUC values close to 1.0, demonstrating high accuracy in skin disease detection.

- Classification Report-

Condition	Precision	Recall	F1-Score	Support
Acne	0.93	0.97	0.95	275
Cellulitis	0.75	0.87	0.80	38
Eczema	0.71	0.81	0.76	242
Psoriasis	0.78	0.69	0.74	272
Shingles	0.88	1.00	0.94	44
Tinea	0.78	0.67	0.72	286
Accuracy			0.81	1077
Macro Avg	0.80	0.84	0.82	1077
Weighted Avg	0.81	0.81	0.80	1077

Table 7.1: TransNet-X F1 Score Table

The Classification Report for TransNet-X shows the following:

- Acne is detected with high accuracy (precision: 0.93, recall: 0.97, F1-score: 0.95).
- Shingles has perfect recall (1.00), but slightly lower precision (0.88), leading to an F1-score of 0.94.
- Cellulitis, Psoriasis, and Tinea show decent performance, but with some room for improvement, particularly in recall for Tinea (0.67).
- Eczema has the lowest precision (0.71) and recall (0.81), showing moderate performance.
- Overall Accuracy is 81%, with a macro average F1-score of 0.82 and weighted average F1-score of 0.80, indicating balanced performance across classes.

In summary, TransNet-X performs well overall but could be improved, especially for Eczema and Tinea.

7.3.2 Advantages and Disadvantages of TransNet-X:

Advantage
Improved Accuracy: The accuracy rate of TransNet-X reaches 80.5%, surpassing MobileNet V3 (76.42%) and standing close to Swin Transformer (84.77%). Our model demonstrates continuous performance enhancement while handling a wide range of skin conditions, particularly in the diagnosis of advanced illnesses such as Psoriasis and Eczema.
Efficiency: The TransNet-X model achieves both performance excellence through Swin Transformer and computing efficiency via MobileNet V3. The Swin Transformer functions as an educational figure that analyzes high-level patterns across large data ranges, while MobileNet V3 acts as a student through information distillation. The combination of MobileNet V3 with Swin Transformer allows TransNet-X to reach high accuracy levels while consuming minimal computational resources, thus enabling real-time mobile use. The system achieves high performance without compromising its functional capabilities.
Scalability: TransNet-X benefits from the Swin Transformer's ability to scale with larger datasets while still maintaining the efficiency of MobileNet V3 for deployment on resource-constrained devices.

Disadvantage
Implementation Complexity: Combining two different architectures (Swin Transformer and MobileNet V3) and using knowledge distillation adds complexity to the model. This requires advanced skills in model implementation and tuning.
Training Time: Although MobileNet V3 is a lightweight model, the distillation process can still increase training time due to the need for both models to interact during the learning process.
More Complex to Train: The process of knowledge distillation requires additional time and effort to implement and fine-tune the model. Training the teacher and student models together also requires more computational resources during the training phase.

Why TransNet-X is Better than MobileNet V3 and Swin Transformer?

TransNet-X provides a high overall accuracy of 80% by integrating the computational efficiency of MobileNet V3 with Swin Transformer's accuracy capabilities (84.77%) thus achieving better performance than MobileNet V3 (76.42%) while maintaining significantly reduced computation expenses. A key advantage of this model consists of Acne (0.99 AUC) and Cellulitis (0.97 AUC) and Shingles (1.00 AUC) alongside enhanced AUC detection of complex conditions such as Eczema (0.93) and Psoriasis (0.91). Precision and recall metrics for Acne reached 0.93 precision alongside 0.97 recall yet Shingles obtained 1.00 recall while Eczema demonstrated balanced yet moderate results with 0.71 precision and 0.81 recall. The number of misclassifications in TransNet-X for Psoriasis and Tinea remains lower than what was observed in MobileNet V3. The integration of Swin Transformer's feature extraction depth into MobileNet V3's compact architecture creates TransNet-X which provides swift operation in mobile and real-time conditions while retaining excellent performance.

Chapter 8

Approach

An interactive interface of the app is designed which can give a real-time experience of how the hybrid model is going to be incorporated into skin disease detection. It is also going to show how the test cases will be stored in a dataset and the admin can access the dataset.

Now a dummy overview is given of how the app will be operated –

8.1 User Login

- Sign Up: Choose your industry's name.
- User Login: Open the app and log in with your username and Department name.

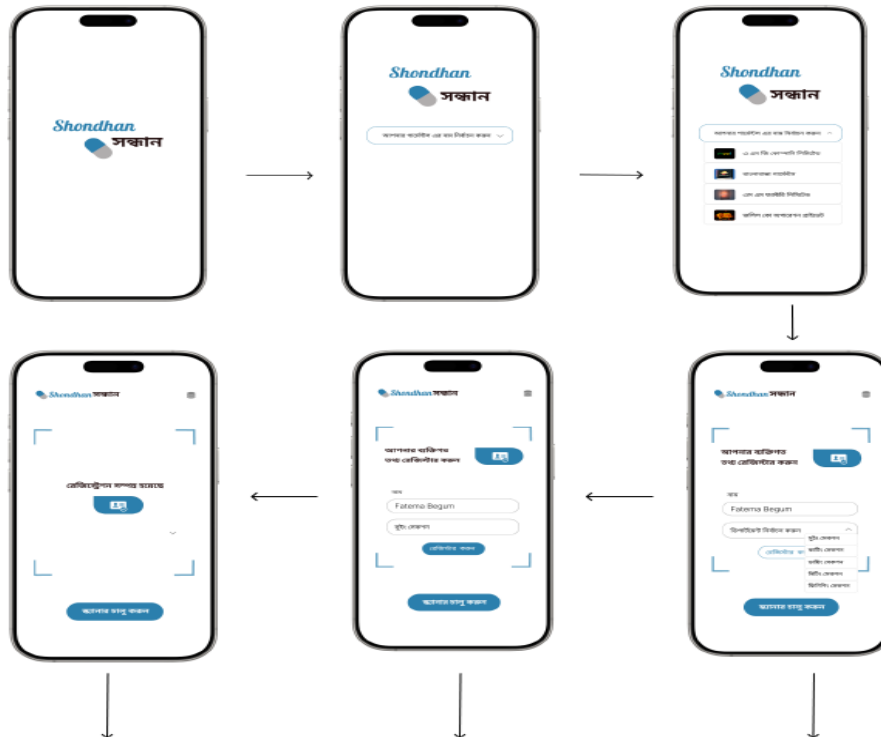


Figure 8.1: Workers' Sign Up

Admin Login

Admins can log in with their admin credentials (Admin ID and password). Access is granted to the backend features for managing the system.

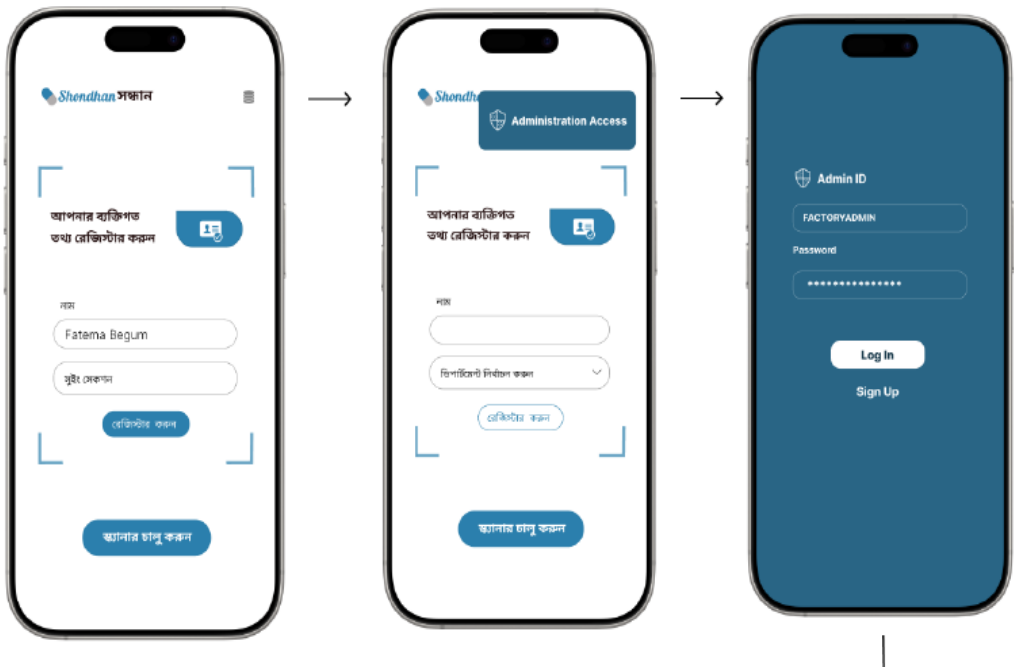


Figure 8.2: Admin Login

8.2 Skin Analysis (User Functionality)

Scanning Process

Once logged in as a user, you will be directed to the camera interface to begin scanning your skin.

- **Open the Camera:** The app will prompt you to focus the camera on the area of your skin you want to analyze. Ensure the skin is well-lit for best results.
- **Take a Photo:** Tap the camera button to capture a photo of the skin. The app will automatically focus and capture the area.

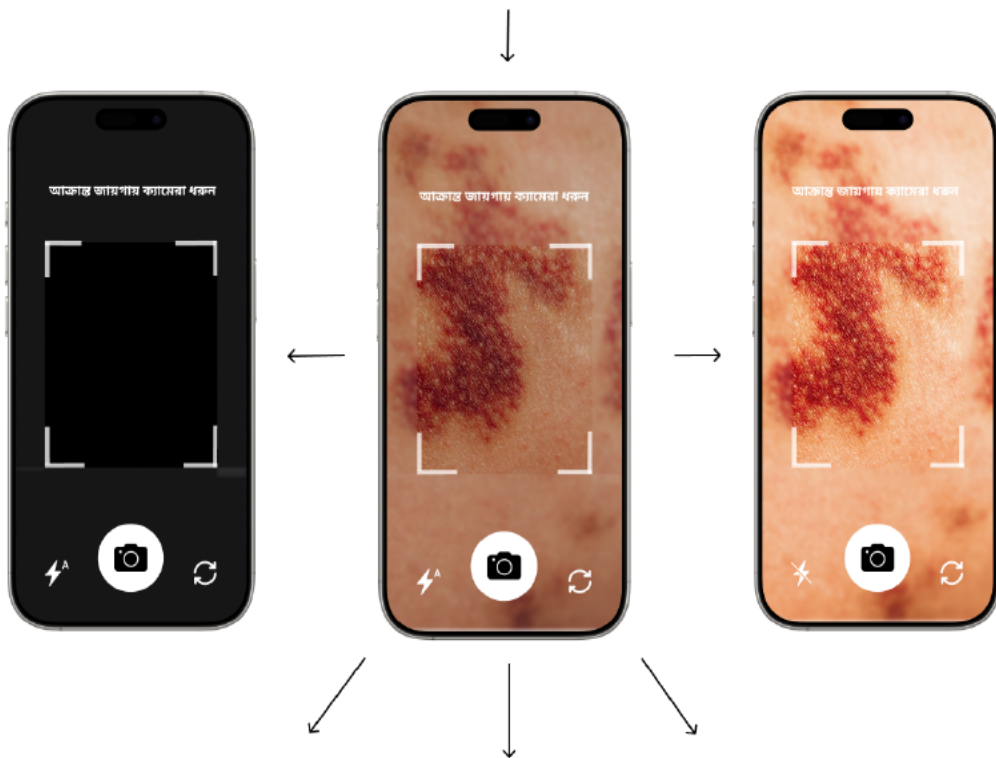


Figure 8.3: Infected Skin Capture Through Mobile Camera

8.3 Viewing Results

Once the scan is completed, the app analyzes the skin and displays the results.

- **Color-Coded Indicators:** Red indicates a potential problem area, green means healthy and yellow might suggest further monitoring.
- **Open the Camera:** The app may provide a percentage (e.g., “87% Likely to be Acne”) and a description of the skin condition.

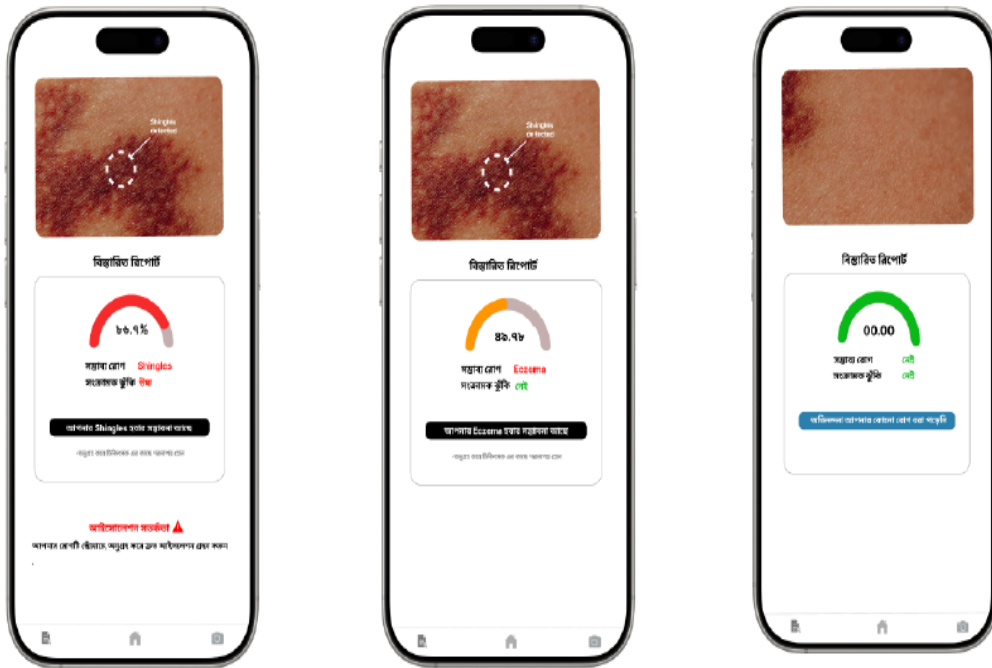


Figure 8.4: Infected Skin Capture Through Mobile Camera

8.4 Admin Features

Admin Login

Admins can log in to access and manage the system. Upon logging in, the app provides access to the admin dashboard for overseeing various tasks.

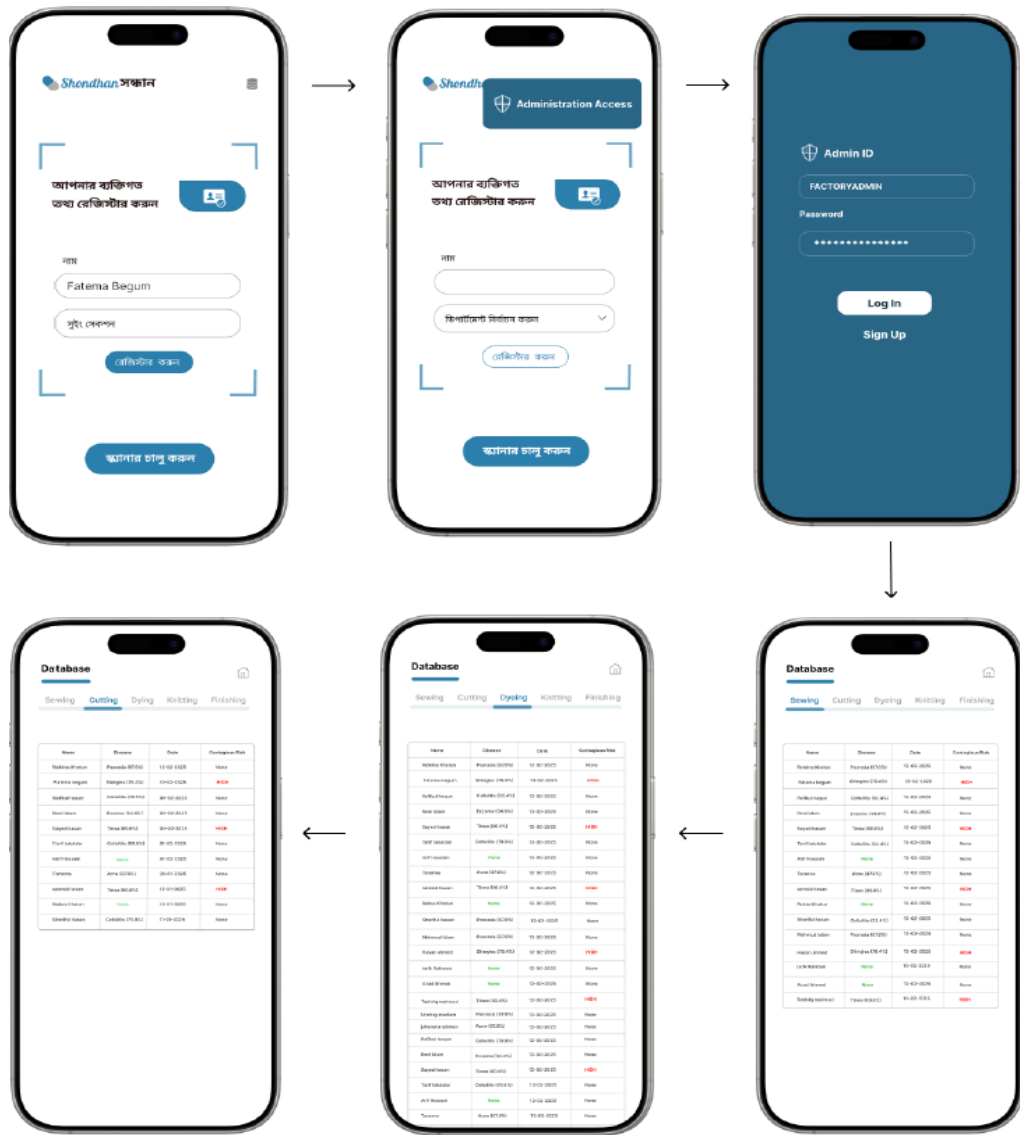


Figure 8.5

8.5 Managing Databases

The admin dashboard allows admins to access and manage databases, such as sewing, dyeing, knitting, cutting, and finishing.

- **Task Monitoring:** The app displays the real-time progress of tasks in each department.
- **Color-Coded Updates:** Each department shows status updates, indicating whether tasks are

8.6 Sitemap of the APP -

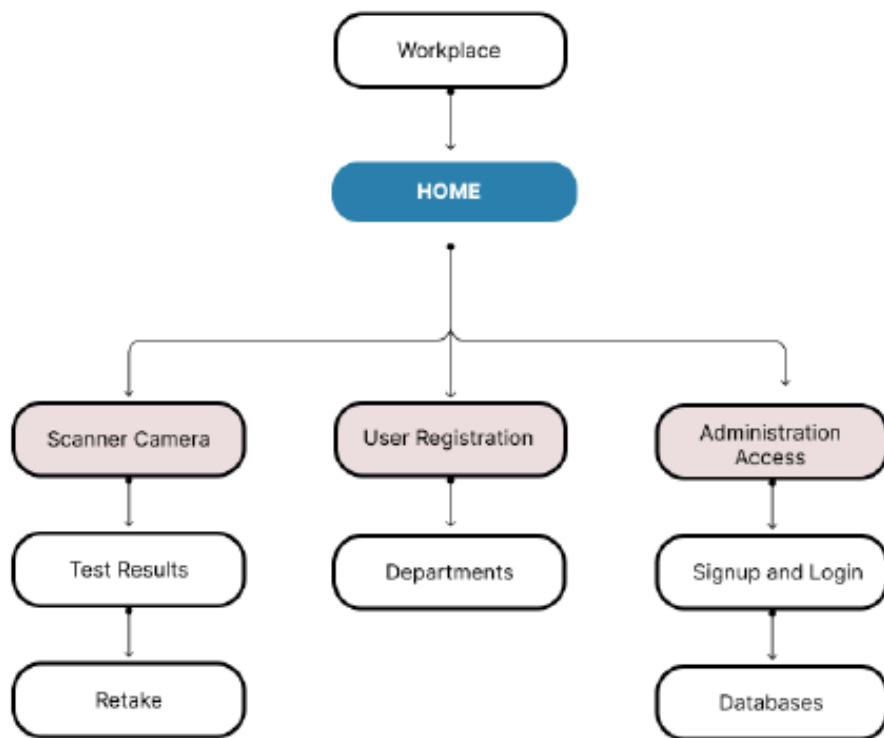


Figure 8.6

8.7 User Flow for Worker and Admin

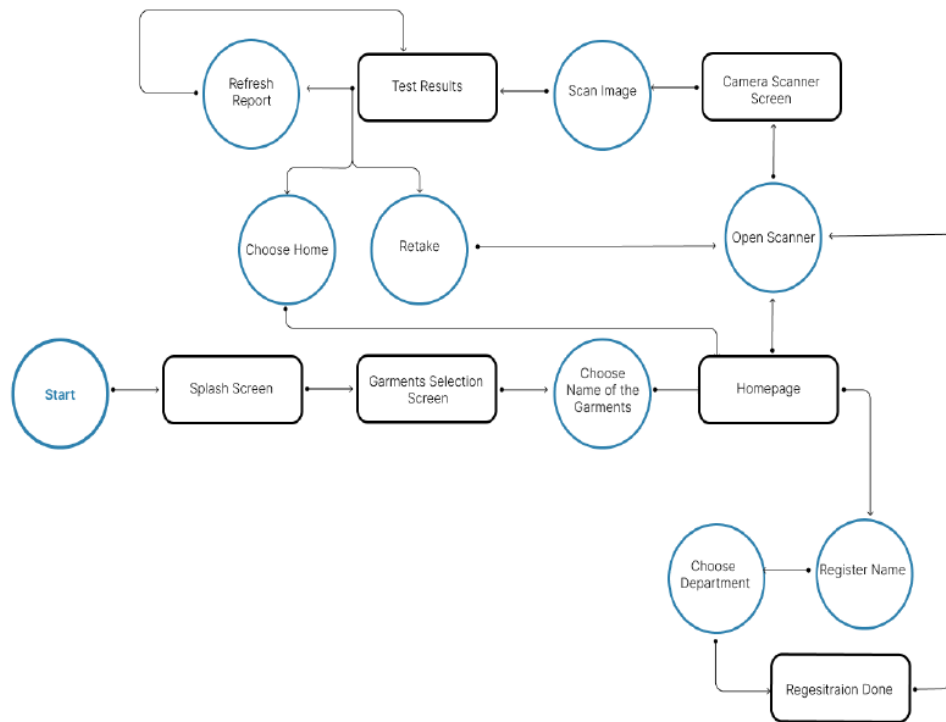


Figure 8.7: User flow of Worker

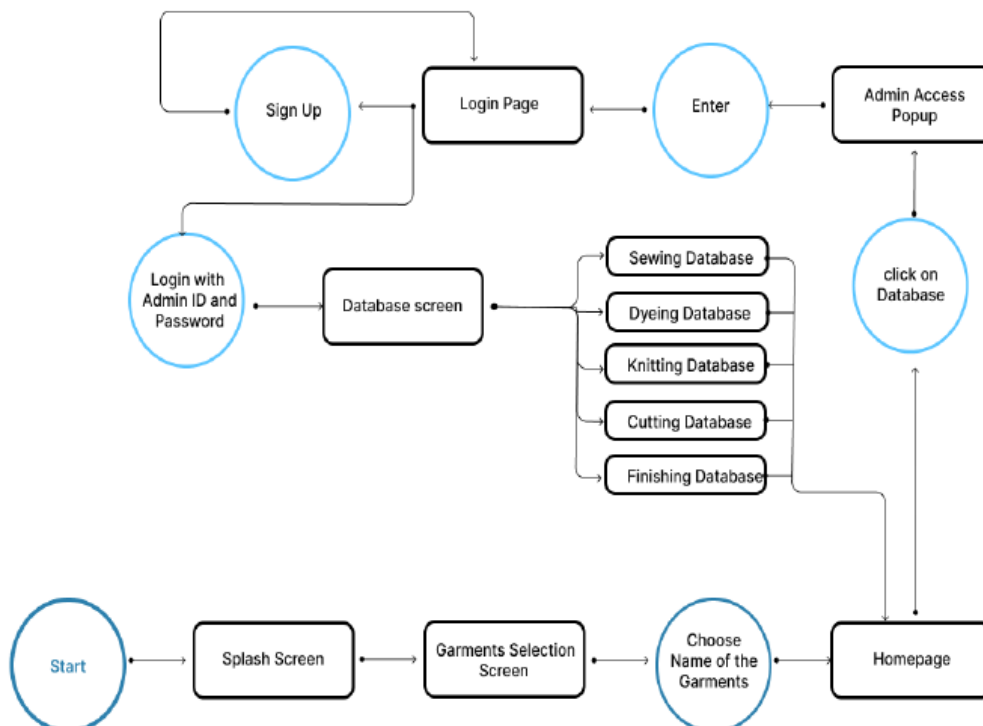


Figure 8.8: User flow of Administration

Chapter 9

Conclusion

This study has pursued health challenges faced by garment workers, with a key focus on skin diseases resulting from chemical exposure within the factory environment. We consulted with dermatologists and identified six common skin that usually occur with garment workers. We also surveyed 2 garment factories so that we could have some one-on-one interaction with the workers and get to know what health issues they suffer from and what sort of health facilities they are currently receiving, we also consulted the doctor of the garment factory so that we know what issues his patients are mainly suffering from. We then individually found out the datasets related to each disease and constructed our dataset. We then split that data into the training set, test set, and validation set in a ratio of 70%, 10%, and 20%, respectively, for effective model training. For expanding our dataset, we use a data augmentation on only the train and validation class to balance them out. The CNN models for predicting skin diseases are DenseNet, ResNet50, VGG16, MobileNet, InceptionV3 and Xception, along with 3 Transformer models that include ViT, Swin and DeiT. We then took the best performing CNN and Transformers which were MobileNet and Swin Transformer and then made a hybrid model that was better performing than all the CNN models and was more efficient and lightweight than the Transformers, this ensured that the model would work very well in mobile-based applications due to it being very light.

Bibliography

- [1] Yifokire Tefera Zele, Abera Kumie, Wakgari Deressa, Magne Bråtveit, and Bente E Moen. Registered health problems and demographic profile of integrated textile factory workers in ethiopia: a cross-sectional study. *BMC Public Health*, 21:1–15, 2021.
- [2] Yuta Yanagisawa, Kosuke Shido, Kaname Kojima, and Kenshi Yamasaki. Convolutional neural network-based skin image segmentation model to improve classification of skin diseases in conventional and non-standardized picture images. *Journal of Dermatological Science*, 109(1):30–36, 2023.
- [3] Tanweer Islam. Health concerns of textile workers and associated community. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, 59:00469580221088626, 2022.
- [4] A Haque and Estique Bari. A survey report on the garment workers of bangladesh 2020. *Dhaka: Asian Center for Development*, 2021.
- [5] Marc Combalia, Noel Codella, Veronica Rotemberg, Cristina Carrera, Stephen Dusza, David Gutman, Brian Helba, Harald Kittler, Nicholas R Kurtansky, Konstantinos Liopyris, et al. Validation of artificial intelligence prediction models for skin cancer diagnosis using dermoscopy images: the 2019 international skin imaging collaboration grand challenge. *The Lancet Digital Health*, 4(5):e330–e339, 2022.
- [6] Md Tanvir Hasan, Amina Amin, Md Uzzal Chowdhury, Kuhel Faizul Islam, Sk Mojibul Huq, Rezvina Parvin, S Jobayear Ahmed, Md Liakath Ali, and Sabina Faiz Rashid. Health sufferings, healthcare seeking behavior, awareness about health insurance, and health related rights of ready made garments workers in bangladesh: Findings from a cross-sectional study. *journal of Occupational Health*, 63(1):e12256, 2021.
- [7] Mostafiz Ahammed, Md Al Mamun, and Mohammad Shorif Uddin. A machine learning approach for skin disease detection and classification using image segmentation. *Healthcare Analytics*, 2:100122, 2022.
- [8] Bhuvaneshwari Shetty, Roshan Fernandes, Anisha P Rodrigues, Rajeswari Chengoden, Sweta Bhattacharya, and Kuruva Lakshmanna. Skin lesion classification of dermoscopic images using machine learning and convolutional neural network. *Scientific Reports*, 12(1):18134, 2022.
- [9] Evgin Goceri. Diagnosis of skin diseases in the era of deep learning and mobile technology. *Computers in Biology and Medicine*, 134:104458, 2021.
- [10] KHAN NR, DIPTI TR, FERDOUSI SK, HOSSAIN MZ, SONY SA, ISLAM MS, et al. Occupational health hazards among workers of garment factories in dhaka city, bangladesh. *Journal of Dhaka Medical College*, 24(1), 2015.

- [11] Khadija Khanom, Md Tariqul Islam, Abdulla Al-Towfiq Hasan, Shah Mahmud Sumon, and Mohammad Rakibul Islam Bhuiyan. Worker satisfaction in health, hygiene and safety measures undertaken by the readymade garments industry of bangladesh: A case study on gazipur. *Journal of Business Studies Pabna University of Science and Technology ISSN*, pages 2410–8170, 2022.
- [12] Savitha Rani BB, Vinay KS, Nikhil Dev, Krishna Iyengar, et al. Assessment of health status and qol of the womens working in garment factory in tumkur: A cross-sectional study. *Indian Journal of Public Health Research & Development*, 11(5), 2020.
- [13] Ponnusamy Venkataramanan, Paulraj Prathap, Palanisamy Sivaprakash, and Sivaprakash Kanchana. Assessment of azo dye in textile wastewater and examination of the fire safety measures in garment industries in tirupur district, india. *Polish Journal of Environmental Studies*, 32(4), 2023.
- [14] Jafaridarabjerdi Mahin, Xuezhu Xu, Ling Li, and Cong Zhang. Diagnosing skin cancer using social spider optimization (sso) and error correcting output codes (ecoc) with weighted hamming distance. *Scientific Reports*, 14(1):22997, 2024.
- [15] Ibrahim Abunadi and Ebrahim Mohammed Senan. Deep learning and machine learning techniques of diagnosis dermoscopy images for early detection of skin diseases. *Electronics*, 10(24):3158, 2021.
- [16] Philipp Tschandl, Cliff Rosendahl, and Harald Kittler. The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific data*, 5(1):1–9, 2018.
- [17] Syed Inthiyaz, Baraa Riyadh Altahan, Sk Hasane Ahammad, V Rajesh, Ruth Ramya Kalangi, Lassaad K Smirani, Md Amzad Hossain, and Ahmed Nabih Zaki Rashed. Skin disease detection using deep learning. *Advances in Engineering Software*, 175:103361, 2023.
- [18] G Schaefer, R Tait, Arcangelo Merla, Shao Ying Zhu, et al. Automated overlay of infrared and visual medical images using skin detection and image registration. *International Journal of Computational Vision and Biomechanics*, 3(1):81–87, 2017.
- [19] S Bharathi, J Premchand, AS Nivedeethaa, M Kousiya, and V Ajay Kumar. Identification of melanoma from nevus images. In *Journal of Physics: Conference Series*, volume 1917. IOP Publishing, 2021.
- [20] Sander Feringa, Sietse Ludger Geertsema, M Biehl, and M Wilkinson. Automatic separation of skin lesions from healthy human skin. Technical report, tech. rep., University of Groningen, 2012.
- [21] Adil Lateef Albukhnefis. An efficient algorithm for automated skin lesion detection: A non-machine learning approach. *Journal of Al-Qadisiyah for Computer Science and Mathematics*, 16(3):Page–1, 2024.
- [22] Supreet Bindra, AGK Sinha, and AI Benjamin. Questionnaire for low back pain in the garment industry workers. *Indian Journal of Occupational and Environmental Medicine*, 17(2):48–57, 2013.
- [23] Sweta Bhattacharya, Praveen Kumar Reddy Maddikunta, Quoc-Viet Pham, Thippa Reddy Gadekallu, Chiranji Lal Chowdhary, Mamoun Alazab, Md Jalil Piran, et al. Deep learning and medical image processing for coronavirus (covid-19) pandemic: A survey. *Sustainable cities and society*, 65:102589, 2021.

- [24] Arslan Javaid, Muhammad Sadiq, and Faraz Akram. Skin cancer classification using image processing and machine learning. In *2021 international Bhurban conference on applied sciences and technologies (IBCAST)*, pages 439–444. IEEE, 2021.
- [25] Meghavi Rana and Megha Bhushan. Machine learning and deep learning approach for medical image analysis: diagnosis to detection. *Multimedia Tools and Applications*, 82(17):26731–26769, 2023.
- [26] Sen Yang and Liejun Wang. Hmt-net: Transformer and mlp hybrid encoder for skin disease segmentation. *Sensors*, 23(3067), 2023.
- [27] Anna Escalé-Besa, Oriol Yélamos, Josep Vidal-Alaball, Aïna Fuster-Casanovas, Queralt Miró Catalina, Alexander Börve, Ricardo Ander-Egg Aguilar, Xavier Fustà-Novell, Xavier Cubiró, Mireia Esquius Rafat, et al. Use of artificial intelligence as a diagnostic support tool for skin lesions in primary care: feasibility study in clinical practice. *nature*, 2023.
- [28] Saad Rasool, Ali Haider Khan, Qaisar Rasool, Shahzad Abbas, Sayyid Kamran Hussain, et al. A systematic analysis for the detection of skin disease using deep learning methodologies. *Journal of Computing & Biomedical Informatics*, 4(02):66–75, 2023.
- [29] Ursula Schmidt-Erfurth, Amir Sadeghipour, Bianca S Gerendas, Sebastian M Waldstein, and Hrvoje Bogunović. Artificial intelligence in retina. *Progress in retinal and eye research*, 67:1–29, 2018.
- [30] Sen Yang and Liejun Wang. Enhancing skin disease classification leveraging transformer-based deep learning architectures and explainable ai. *Sensors*, 23(3067), 2023.
- [31] Yali Nie, Paolo Sommella, Marco Carratù, Mattias O’Nils, and Jan Lundgren. A deep cnn transformer hybrid model for skin lesion classification of dermoscopic images using focal loss. *Diagnostics*, 13(1):72, 2023.
- [32] Vlad-Constantin Lungu-Stan, Dumitru-Clementin Cercel, and Florin Pop. Skindistilvit: Lightweight vision transformer for skin lesion classification. *arXiv preprint arXiv:2308.08669*, 2023.
- [33] Santham Lillypet, Timsi Jain, and Bobby Joseph. Health problems among garment factory workers: A narrative literature review. *Journal of Occupational health and epidemiology*, 6(2):114–121, 2017.
- [34] Yaya Sudarya Triana, Mohd Azam Osman, Adji Pratomo, Muhammad Fermi Pasha, Deris Stiawan, and Rahmat Budiarto. Neural network models selection scheme for health mobile app development. *IAES International Journal of Artificial Intelligence*, 12(3):1191–1203, 2023.
- [35] Hong Qing Yu and Stephan Reiff-Marganiec. Targeted ensemble machine classification approach for supporting iot enabled skin disease detection. *IEEE Access*, 9:50244–50252, 2021.
- [36] Maad M Mijwil. Skin cancer disease images classification using deep learning solutions. *Multimedia Tools and Applications*, 80(17):26255–26271, 2021.
- [37] Siddharth Gupta, Avnish Panwar, and Kishor Mishra. Skin disease classification using dermoscopy images through deep feature learning models and machine learning classifiers. In *IEEE EUROCON 2021-19th International Conference on Smart Technologies*, pages 170–174. IEEE, 2021.

- [38] Parul Sharma, Yash Paul Singh Berwal, and Wiqas Ghai. Performance analysis of deep learning cnn models for disease detection in plants using image segmentation. *Information Processing in Agriculture*, 7(4):566–574, 2020.
- [39] Lawrence C Ngugi, Moataz Abelwahab, and Mohammed Abo-Zahhad. Recent advances in image processing techniques for automated leaf pest and disease recognition—a review. *Information processing in agriculture*, 8(1):27–51, 2021.
- [40] M Yuvaraja, G Aravindh, R Priya, and V Suresh Babu. Image processing-driven deep learning model for plant disease detection to enhance irrigation efficiency in smart agriculture. *ICTACT JOURNAL ON IMAGE AND VIDEO PROCESSING*, 15, 2024.
- [41] Mohammad Shahin, F Frank Chen, Ali Hosseinzadeh, Hamid Khodadadi Koodiani, Awni Shahin, and Omar Ali Nafi. A smartphone-based application for an early skin disease prognosis: Towards a lean healthcare system via computer-based vision. *Advanced Engineering Informatics*, 57:102036, 2023.
- [42] Nayan Chaure. Acne dataset, 2021. Accessed: 2025-02-02.
- [43] Pallapuraj Kumar. Psoriasis skin dataset, 2021. Accessed: 2025-02-02.
- [44] Subir Biswas. Skin disease dataset, 2021. Accessed: 2025-02-02.