A Deep Learning System for Differential Diagnosis of Melanoma

Hoang, T. M. T.

Anh, P.

Liao, Y.-P.

Tasneem, R.

Novoa, V.

Dongpakkij, P.

Department of Transdisciplinary Innovation, University of Technology Sydney
Ali Anaissi

November 6, 2024

ABSTRACT

With the advancement of technology, the diagnosis of medical issues has improved over time. Transforming the healthcare industry will lead to an approach to increasing diagnosis performance. The reason is that humans can make any error. It can be seen on the news that a doctor provides faulty advice to a patient, which unfortunately leads to losing a high amount of medical payment or a disease gone severe. However, to minimise this issue, one of the technologies that have become a crucial part of healthcare is machine learning. Machine learning is learning and improving by recognising patterns, data and feedback. A high performance of machine learning is mainly created by proper training and frequent tuning. The error can be gradually mitigated with high-performance machine learning to diagnose a patient.

This project is related to medical issues and the use of machine learning. We apply deep learning, which is the subset of machine learning. It has been recognised as an effective tool for tackling the problem in the healthcare world. The integration of deep learning can be applied by using neural networks that imitate human memory to recognise structures and patterns, such as categorising objects by image pattern. That is why we will use deep learning to diagnose one of the leading causes of skin diseases, melanoma. Melanoma is commonly a skin disease that Australians normally have unprotected exposure to the UV for quite a long time. Therefore, this can be a good example to be utilised by deep learning.

The expected result of this project is to see how accurate the trained deep learning will be after we have tested it by recognising the images of melanoma compared to the tested images. Secondly, this model will be able to be trusted by dermatologists for the accurate detection and differentiation of melanoma and other skin diseases. Lastly, in terms of accuracy and speed, well-trained artificial intelligence (AI) can outperform dermatologists' traditional diagnosis methods.

After what we have done, we received the effective result from the models we have been training since the project was begun. Eventually, we have concluded the result to answer all three research questions in this report.

Table of Contents

SECTION 1.
Introduction

It is known that Australia is a country where people have a high risk of being diagnosed with skin cancer. Around 2 in 3 of them will be diagnosed with skin cancer in their lifetime. People who have melanoma have a high rate of skin cancer. There has been a high number of cases for many years, and recently, in 2023, 18,257 cases happened and 1,314 deaths in Australia. 2023, Melanoma of the skin was the third estimated most diagnosed skin cancer, as shown in Figure 1.1 below.

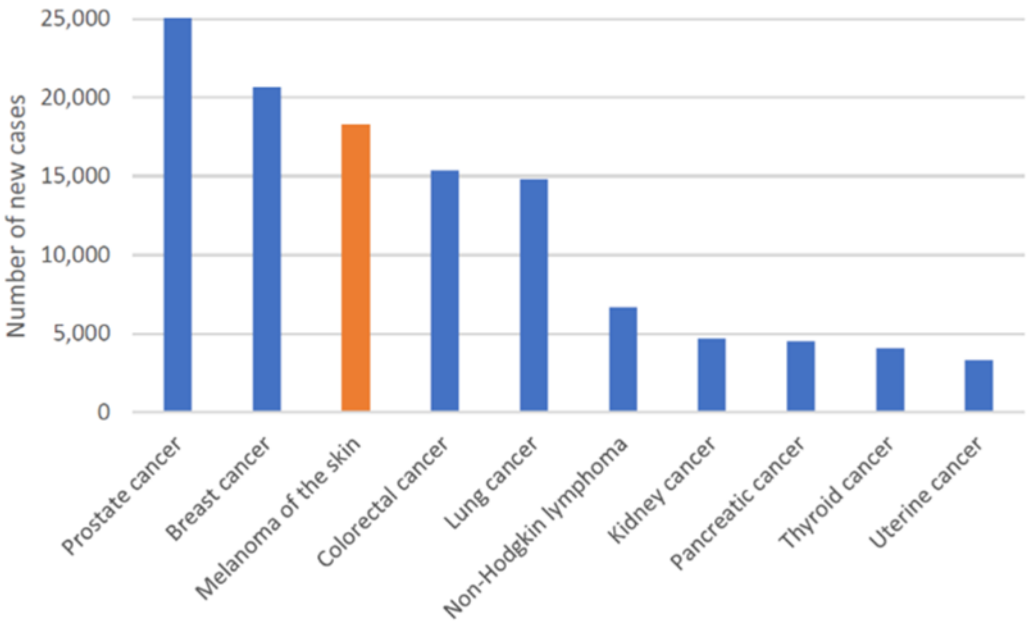


Figure 1.1: Estimated cancer incidence in Australia, 2023

Melanoma is a common type of skin cancer. It mainly develops in skin cells named melanocytes that mutate and divide uncontrollably, usually on the part of the body that has been exceedingly exposed to ultraviolet (UV) radiation without any protection. Melanoma can develop anywhere on the body. It is more likely to happen on the chest, back, legs, face or even in the eyes, as in Figure 1.2.

At the early stage, a patient can notice any changes in the skin, such as new spots or moles or colouring changes that are difficult to detect. In the next stage, the spot can become itchy, tender and bleeding. Later, the spot becomes a lump that looks ulcerated or crusty. After that, the spot can become a skin cancer and spread to any internal organ part such as the brain, lungs or liver.

Figure 1.2: Superficial spreading melanoma

Therefore, this proposal report aims to apply the new machine learning method, such as deep learning, to assist dermatologists in diagnosing melanoma, which can be the early stage of skin cancer and eliminate it before it develops to the next stage. This can be a tool for dermatologists to diagnose melanoma more accurately. This project based on 3 research questions to measure the performance of deep learning that will be able to diagnose melanoma firmly.

To provide the solution, we have created the application for detecting Melanoma through Streamlit Data Application and has the high accuracy model behind the detection process of melanoma by uploading skin images. By this, it will enable us to come to an approach to the practical solutions.

SECTION 2.

Related Literature

2.1 Overview of Related Works

Over the past few years, there has been a significant increase in the number of research papers applying transfer learning and deep learning models to detect skin diseases, specifically Melanoma, which is the deadliest type of skin cancer. This surge is primarily driven by advancements in medical imaging techniques and deep learning models which can automatically extract skin features based on skin lesion images to classify.

In this study, we applied a two-layer hierarchical search structure to effectively filter and identify relevant publications for our deep learning system aimed at Melanoma detection. The first-level keyword set involves broad techniques, such as convolutional neural network, deep learning, transfer learning, applying to a broad field (image classification). The second layer narrows the focus to DenseNet-121 and EfficientNet-B4 model. We employed the selected keywords across two prestigious platforms for accessing high-quality scientific and academic research: IEEE Xplore and ScienceDirect. Although IEEE Xplore primarily focuses on engineering and computer science papers, this website provides access to various machine learning studies. ScienceDirect is more superior when it comes to AI journal articles as this platform provides comprehensive collections of research papers in areas like life sciences, physical sciences, and health.

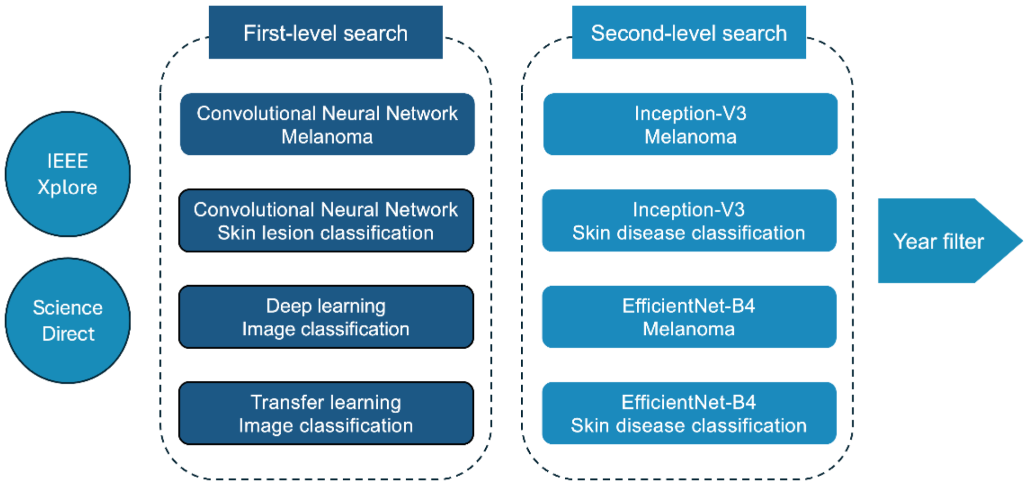


Figure 2.1: Search Strategy Flowchart

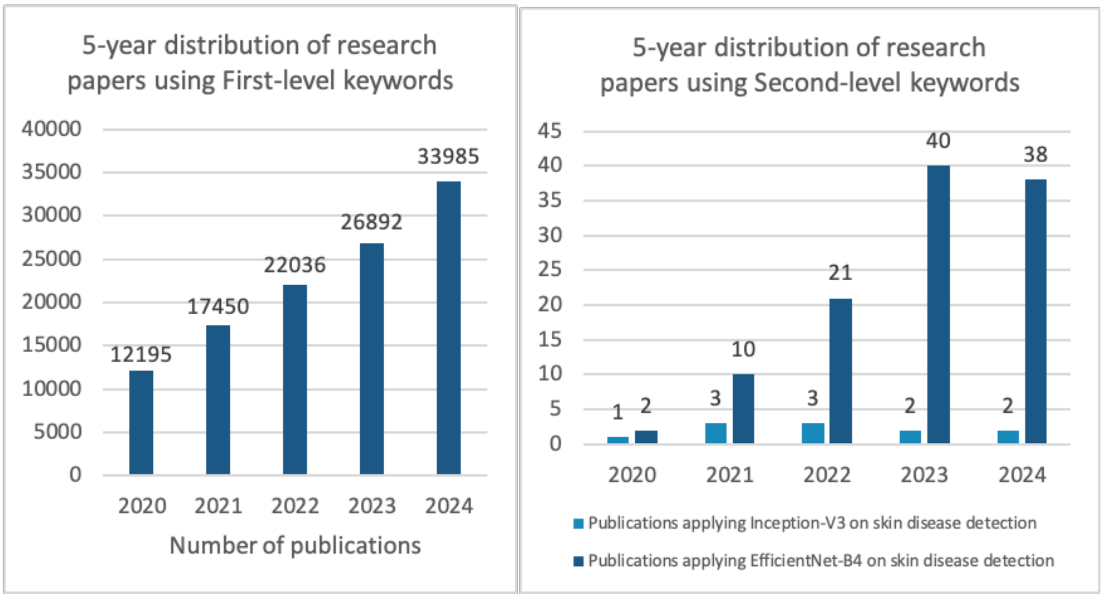


Figure 2.2: Distribution of related works

Figure 2.2 summarised the number of related works over the past 5 years by using two-layer hierarchical structure (Figure 2.1). Approximately 12,500 documents within 5-year period were retrieved through the first-layer keyword set. The number of research papers on applying Deep Learning/Convolutional Neural Network to skin disease identification has increased sharply by 1.8 times from 2020 to 2024. When keywords ["Inception-V3" and "skin disease classification"] or ["Inception-V3" and "Melanoma"] are filtered, 11 papers were screened out (Figure 2.2). Among these, 6 journal articles focus on applying Inception-V3 to detect Melanoma, indicating a significant interest in this architecture for Melanoma diagnosis. Compared to Inception-V3, EfficientNet-B4 has more publications in the context of Melanoma detection and skin disease classification with only 38 papers published in 2024. Particularly, 25 papers were identified using the keywords ["EfficientNet-B4" and "Skin disease classification"], while only 13 articles focus specifically on applying EfficientNet-B4 to Melanoma detection.

2.2 Literature Review

Melanoma is one of the most dangerous skin cancers in the world, with thousands of cases detected and about 55,500 cancer deaths annually (Schadendorf et al., 2015). Moreover, early detection is significant for this disease, but traditional diagnosis can be time-consuming. Hence, the application of deep learning is essential in this case because this technology has an advantage in image classification tasks. More specifically, the convolutional neural network (CNN) is a type of deep learning algorithm, that can understand the images effectively by extracting the image features such as shape or edges. Therefore, this algorithm is useful in analysing the medical images, including symptoms of Melanoma. This project will focus on the dangers of melanoma to public health; it will also review many studies to address the effectiveness of applying deep learning to detect this disease.

According to research from the University of Duisburg-Essen and the University Hospital Essen, they said the survivor rates in many countries in Europe are highly different because, after the first diagnosis, the patients in central Europe have five years of survival; however, this number is less than 50% and more than 90% in Eastern Europe and Northern Europe, respectively. Furthermore, Schadendorf et al. (2015) have mentioned that Australia and New Zealand are the two countries that have the highest rates of Melanoma; also, the occurrence of this skin cancer on white skin is significantly higher than in Asia and Africa. Figures (2.3), (2.4), and (2.5) below show the incidence rate and mortality rate of Melanoma worldwide; both figures address that there are three areas that have the highest rate of this disease, including North America, North Europe, Australia and New Zealand. In addition, Matthew et al. (2017) have mentioned that the number of Melanoma cases predicted in the United States was about 87 110 in 2017. Moreover, this country has spent 8.1 billion dollars for the treatment of all types of skin cancers, which 3.3 billion dollars is for Melanoma in the last 10 years.

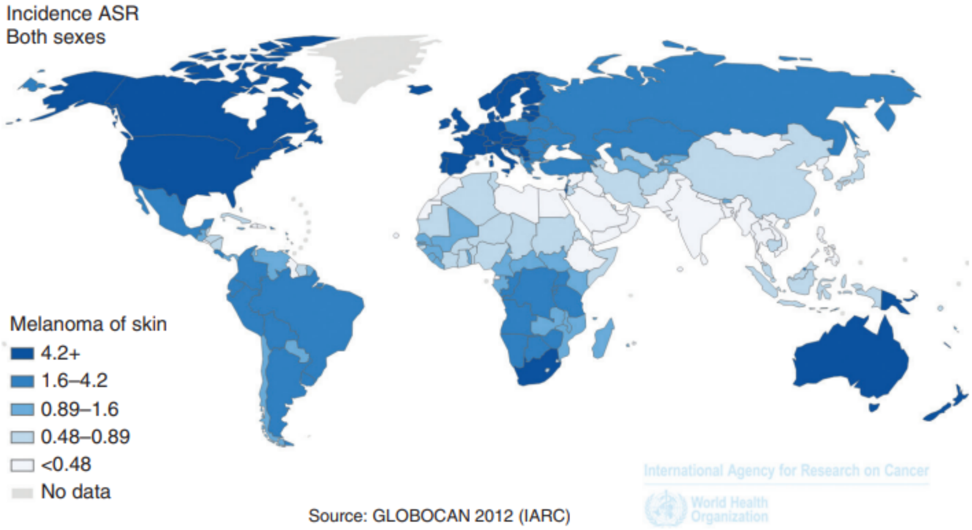


Figure 2.3: Incidence rate of Melanoma worldwide

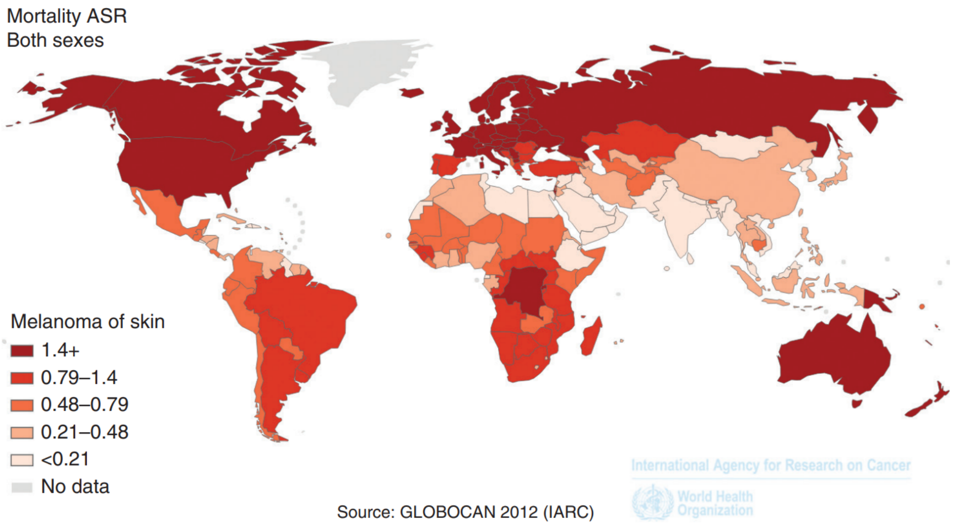


Figure 2.4: Mortality rate of Melanoma worldwide

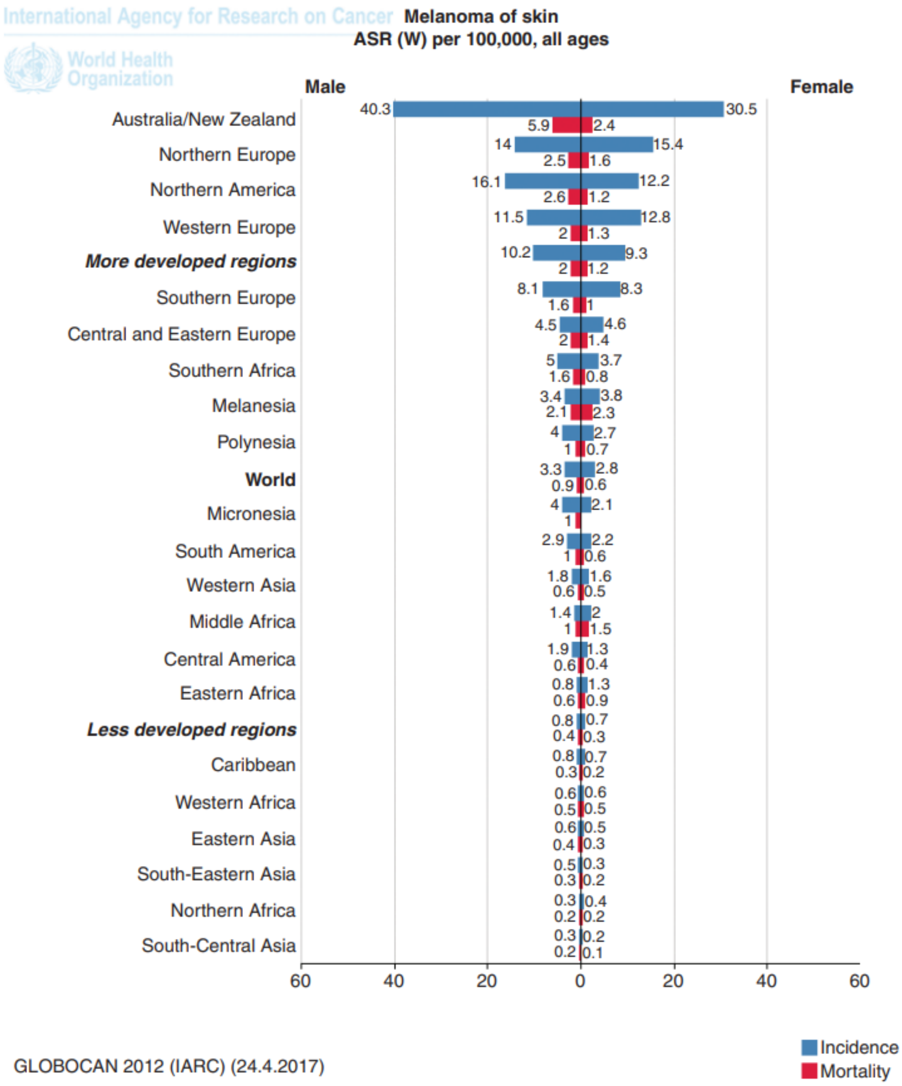


Figure 2.5: Incidence and Mortality of Melanoma in the world

The world has several treatment methods for Melanoma and Domingue et al. (2018) have mentioned them in the table below:

Melanoma Treatment	Description	Outcomes	Negative Side Effect
Chemotherapy	Primary option for Melanoma disease, and it plays a crucial role in the treatment including resistant, progressing, or have relapsed.	No improment in overall survival.	
Electrochemotherapy	Deliver drug into the cells byusing high-intensity electric pulses	Usefull method with 85% response in overall.	No major negative effects.
Dacarbazine	This is kind of a chemotherapy medication and it is using for metastatic melanoma.	2% to 6% of patients were achieved the 5-year survival. No major improvement in overall survival.	No major negative effects.
Temozolomide	This method is used for advanced melanoma.	Higher rate in progression-free survival compare to Dacarbazine. No major improvement in overall survival.	No major negative effects.
Biochemotherapy	Refers to the combination of immunotherapy and chemotherapy.	Show a highest rate of progression-free survivial. No major development in overall survival.	No major negative effects.
Immunotherapy	This technique mainly focus on reducing the melanoma recurrance after the surgical resection. For example: Interferon (IFN) a-2b, Interleukin-2 (IL-2), Treg Inhibition, or Oncolytic virus therapy.	This method has returned some positive outcomes especially from Adoptive T-cell therapy with the median survival is higher than 3 years, or Treg inhibition reported that the melanoma patients at stage IV has 5% stable disease, and partial responses up to 16.7%.	Several side effects occurs such as: headache, fatigue, nausea, or diarrhea.
Oncolytic virus therapy	This technique is use viruses to infect and destroy the cancer cells, and it is helpful in case of cancer metastatic.	Positive result with 28% of clinical response.	No serious side effects such as nausea, pyrexia, or fatigue.
Targeted therapy	The purpose of this method is inhibit the mutated molecules and slow down the development of cancer cells, while the healthy cells are still not affected. For example: BRAF inhibitors, MEK inhibitors, CKIT inhibitors or VEGF inhibitors.	Stop the melanoma cells increasing and spread the cancer to other parts of the body.	Several side effects occurs such as: rash, itching, headache, nausea, or swelling.

Figure 1.6: Table of treatment methods

Many research studies in recent years have focused on applying deep learning in medical image detection; CNN architectures such as ResNet-101, InceptionV3, and DenseNet-121 are also usually used for this purpose. Furthermore, those models have often been trained on some popular datasets, such as HAM10000 or PH2, applying hyperparameter tuning to handle image features, high sensitivity, and improving detection accuracy. However, feature extraction is still a challenge in this task due to the high similarities inter-class and diversity of intra-class in the skin lesions (Gajera et al., 2022). Convolution Neural Network has some limitations if the dataset size is small; therefore, some researchers found a solution that uses transfer learning to handle this issue. This method will improve the performance of the model by pre-training on large datasets and hyperparameter tuning on the medical datasets.

In recent years, models such as DenseNet and MobileNet have been the research train because they improve accuracy and focus on pre-trained architectures and ensemble models (Roy et al., 2024). Moreover, Patinge et al. (2024) mentioned that when working with a limited dataset, the normalisation or scaling methods should be used to improve the model; also, several research about the comparison between R-CNN and InceptionV3 to address the performance difference between classification model only and combination of classification and segmentation model. They found that, in some cases, the R-CNN model has significantly better results than the classification model (Likhitha & Baskar, 2022).

Convolutional Neural Networks and transfer learning bring many positive results in detecting Melanoma; however, several limitations still exist. Overfitting is the first issue because the public dataset sizes, such as PH2 or HAM10000, are small and not diverse enough (Ye et al., 2024). Secondly, Ye et al. (2024) continued to indicate that many public datasets are just binary classification, which means Melanoma or Non-melanoma; therefore, a larger dataset with more than two classes would handle the overfitting problem and significantly increase the accuracy of the models. In addition, AI has demonstrated its ability in medicine to the world; however, there is not much research about the challenges and impact of implementing AI daily, with all its advantages and disadvantages.

In conclusion, Melanoma is a dangerous skin cancer that alarmingly affects public health; moreover, Northern America, Northern Europe, Australia and New Zealand have the highest rates of this disease. In addition, deep learning is a potential solution to the problem of detecting melanoma in the early stages and increasing the survival rate of patients. On the other hand, the challenge of this solution is the limited dataset size, developing multi-class classification, and researching the advantages and disadvantages of implementing AI in daily life.

SECTION 3.
Project Problems

Melanoma, a form of severe skin cancer, is becoming more and more common every year, raising concerns about its impact on world health. It is particularly common in places like Australia and New Zealand that receive a lot of ultraviolet (UV) radiation. Although early melanoma diagnosis greatly improves patient outcomes, traditional diagnostic techniques are expensive and need specialized knowledge, which could delay treatment, especially in underprivileged areas.

Dermatological diagnoses are also frequently based on subjective evaluation, which increases the possibility of mistake and variability. The change becomes worse by the rising demand on healthcare systems brought on by the increase in skin cancer incidence. Therefore, this project aims in minimizing this diagnostic inefficiency and unpredictability by providing dermatologists with a reliable, effective tool through artificial intelligence (AI). This project aims to create a diagnostic model that improves melanoma detection's precision, rate, and accessibility through using deep learning.

3.1 Project Aims & Objectives

Aim:

To reliably identify melanoma from photos of skin lesions, this project intends to develop an AI-based tool for dermatologists that uses deep learning. Our goal is to minimize the time and resources needed to diagnose melanoma by implementing an automated diagnostic method, which will enable dermatologists to concentrate more on patient care. Furthermore, the tool will help standardize diagnostic precision, which will lessen the variability frequently found in manual evaluations.

Objectives:

1. **Develop an Effective Model:** Using deep learning, primarily convolutional neural networks (CNNs), trained on the selected dataset, to obtain high accuracy in diagnosing melanoma in skin lesion images. This model seeks to meet clinical requirements and deliver trustworthy findings by attaining an accuracy rate of at least eighty percent.
2. **Improve Diagnostic Efficiency:** Automating the image analysis procedure will reduce on the amount of time doctors need to identify melanoma. Dermatologists may be able to concentrate more on treating patients as a result, improving health outcomes.

3. **Establish an Accessible Interface:** Make the model available through an accessible Streamlit app so that dermatologists and other medical professionals can readily use it. By offering real-time melanoma detection, this application improves accessibility in distant and clinical settings.
4. **Enhance dermatological Research with AI:** We also intend for further study on dermatological uses of AI by investigating how deep learning may help identify melanoma early and could be modified for other skin conditions.

3.2 Project Questions

To achieve the project's objectives, the following research questions have been established:

Accuracy in Diagnosing Skin Diseases: ***How accurately can a deep learning model identify different skin diseases from images?*** This question is intended to evaluate the model's diagnostic performance, especially its ability to differentiate melanoma from other skin disorders with high accuracy.

Reliability in Clinical Use: ***Can dermatologists rely on a deep learning model to diagnose melanoma?*** This question focuses on the model's clinical dependability and if it can match or outperform human-level diagnostic accuracy.

Comparative Performance with Traditional Diagnostic Methods: ***Can deep learning models be more accurate and faster than traditional methods?*** This inquiry tries to establish whether the AI-based technique provides noticeable efficiency gains over traditional diagnostics.

3.3 Project Scope

The project's scope was deliberately chosen to focus on certain, manageable aspects of the larger problem of melanoma diagnosis using artificial intelligence, recognizing that a complete solution to this problem is still a challenging and developing objective in the technological and medical domains. By focusing, the study intends to provide a practical, attainable result that benefits dermatology while acknowledging the limitations of existing AI capabilities.

3.3.1 Emphasize on melanoma detection

Instead of trying to identify every kind of skin problem, this project focuses exclusively on melanoma, a particular and deadly form of skin cancer. Although a variety of skin conditions may be detected by AI-driven diagnostic tools, developing and validating such an entire system is beyond the scope of this study. However, the melanoma emphasis enables comprehensive model validation and training, resulting in more dependable and clinically meaningful results.

3.3.2 Using a single dataset

The study only includes HAM10000 dataset, a thoughtfully selected collection of dermatoscopic images for the categorization of skin lesions. By making this decision, the model's performance is guaranteed to be tailored for the type of pictures it has been trained on, enabling accurate melanoma diagnosis. The model's generality would probably be enhanced by adding more datasets, however owing to time and resource limitations, this project is now outside of its scope.

3.3.3 Implementation via an experimental application

A Streamlit application is used to launch a model prototype, allowing users to test melanoma detection in real time. Although the software is a demonstration of concept that shows practical applicability, it is not meant to be a fully functional commercial solution. Though more testing and development would be required for wider implementation, this prototype offers a foundation for assessing the model's usability and real-world functionality.

3.3.4 Performance measures

Reaching a minimum of 80% diagnostic accuracy for melanoma detection is the main performance target. Setting an achievable benchmark is in line with the project's confined scope and supports the model's completion within reasonable bounds, even though beyond this barrier will increase the model's usefulness.

3.3.5 Acknowledgement of limitations

The project's functional AI model for melanoma detection is one of its limitations, although total dependence on deep learning models in clinical settings is still a long-term objective rather than an immediate solution. Further research is needed to address issues including robustness against real-world image variabilities, accuracy across various skin conditions, and model generalization to a variety of skin types. Without stating to solve every issue with AI in dermatological diagnostics, the project defines a particular contribution to the field by recognizing these limits.

Therefore, our project scope is thoughtfully designed to support its results and completion by concentrating on a workable subset of the larger diagnostic issue in dermatology. This focused strategy guarantees that the initiative produces attainable benefits in melanoma diagnosis, providing up opportunities for further study and advancement in AI-powered medical applications.

SECTION 4.
Methodologies

4.1 Methods

The primary aim of this project is to create a robust deep learning model to accurately identify melanoma from dermoscopic images of skin lesions. Figure 4.1 illustrates the comprehensive workflow of the project, which follows a systematic approach to develop a deep learning system for the differential diagnosis of melanoma. It is structured into six main phases: data collection, data cleaning and pre-processing, model selection, model training, model evaluation, and deployment. The development of our Streamlit app commences during the second phase, with the purpose of providing a user-friendly tool for dermatologists to upload and classify skin lesion images in real time.

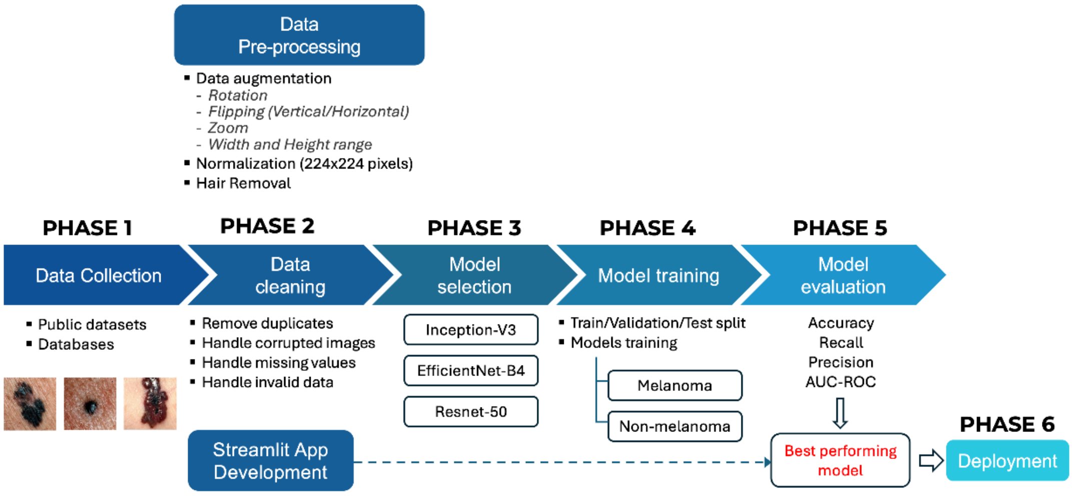


Figure 4.1: Project workflow

Data Collection

This phase involved gathering dataset for classification of melanoma and non-melanoma skin lesions. In this project, we utilise the HAM10000 dataset, which is one of the largest publicly available collections of dermatoscopic images. One of the reasons why HAM10000 was selected would be that it comprises a wide variety of skin lesions including melanoma, dermatofibroma, basal cell carcinoma, etc. Another reason is that this dataset is accompanied by detailed metadata including valuable patient information, such as lesion location, age, and gender, which enables deeper analysis by allowing the deep learning models to account for patient-specific factors that may affect lesion appearance.

Patients' detail regarding lesions' location, their age and gender allows the model to consider context-specific information that may influence lesion appearance. For example, lesions in sun-exposed areas might look different from those in covered areas, and aging skin may have different characteristics. Therefore, incorporating patients' information when developing models may improve the system's sensitivity to patient-specific characteristics, yet increase the accuracy of diagnosing melanoma.

Data Cleaning and Pre-processing

The integration of robust data cleaning and pre-processing methodologies is essential to develop reliable machine learning models. These techniques not only improve the model's ability to learn and generalize but also support ethical data practices, ensuring that the developed systems are both accurate and unbiased. The following techniques were applied to clean and preprocess the data, ensuring that the dataset was well-prepared for model training and analysis.

Data Cleaning: Initially, missing values in the dataset were addressed to minimize the potential impact on model training. Specifically, the age attribute was filled with the mean age of the dataset to maintain continuity, while the sex and localization attributes were filled with "unknown." This approach prevents the loss of valuable data while ensuring that the dataset remains comprehensive for model training. Duplicates were removed to eliminate redundancy, thereby ensuring that the model was trained on unique data points that accurately represent the variety of skin lesions.

Data pre-processing: In recent advancements in deep learning, model performance has greatly benefited from the availability of diverse and substantial datasets. However, obtaining large volumes of data is often resource-intensive in terms of both time and cost. To address this limitation, data augmentation techniques have been implemented in this project to artificially increase the dataset size by creating variations of the original dataset (Morid et al., 2021), mitigate the risks of overfitting, where a model performs well on training data but fails to generalize to unseen data (Shorten & Khoshgoftaar, 2019), thereby improving model robustness and generalization capacity. This is a well-established method in training complex models, where various image manipulation methods, including rotation, shifts, flips, and brightness adjustments—are applied to generate new, synthetic samples based on the original dataset. In this project, a range of augmentation techniques were employed to make the model more adaptive to unseen inputs, thereby improving its testing accuracy and minimizing the risks of overfitting. The parameters used for augmentation were carefully selected to create a diverse and comprehensive set of images. Specifically, images were subjected to random rotations of up to 20 degrees, width and height shifts of up to 20%, and shearing transformations with a 0.2 range. Furthermore, the images were scaled by up to 20% for zoom adjustments, horizontally flipped, and their brightness adjusted within a range of 0.8 to 1.2 to simulate various lighting conditions. The 'nearest' fill mode was applied to handle empty spaces created by these transformations, minimizing the risk of introducing artifacts.

Table 4.1 presents the data augmentation parameters used in this project to enhance the dataset's diversity and improve model generalization.

Table 4.1: Data augmentation parameters

Techniques	Range	Description
Resize	224x224	Resizes all images to 224x224 pixels to standardize input dimensions for the model
Rotation	20 degrees	Randomly rotates images by up to 20 degrees to add variability in orientation.
Width shift range	0.2	Shifts images horizontally by 20% of the image width to simulate movement or misalignment.
Height shift range	0.2	Shifts images vertically by 20% of the image height to simulate movement or misalignment.
Shear range	0.2	Applies shearing transformations with a range of 0.2 to slightly tilt the image.
Zoom	0.2	Zooms images by up to 20% in or out
Horizontal flip	True	Randomly flips images horizontally to enhance the model's robustness
Brightness range	[0.8, 1.2]	Adjusts brightness within the range of 0.8 to 1.2
Fill mode	nearest	Fills any empty areas created by transformations with the nearest pixel value to avoid artifacts

Model selection

In this project, we selected Convolutional Neural Network (CNNs) as the primary model for melanoma and non-melanoma classification due to its effectiveness in image analysis, particularly within the medical imaging domain (Esteva et al., 2017). CNNs excel at automatically learning hierarchies of features from images, which is important for high-precision tasks like melanoma detection from dermatoscopic images (Litjens et al., 2017). Its architecture has multiple layers, including convolutional layers, pooling layers, and fully connected layers, allows the system to capture intricate patterns and features in images.

We specifically select three CNN architectures: Inception-V3, EfficientNet-B4, and ResNet-50. Each of these architectures offers unique advantages that align with the objectives of our project.

Inception-V3 is well-known for effectively balancing model complexity with computational efficiency. This architecture employs a series of inception modules, which facilitate the simultaneous extraction of features at various scales. By applying convolutions of different kernel sizes, Inception-V3 can effectively capture both intricate details and patterns in medical images, enabling it to detect subtle variations in skin lesions effectively (Szegedy et al., 2017). The architecture also incorporates techniques such as batch normalization and dropout, which help mitigate overfitting - a common challenge in medical image classification tasks (Szegedy et al., 2017).

EfficientNet-B4 stands out for its optimization of network depth, width, and resolution, resulting in a highly efficient model that achieves state-of-the-art performance with fewer parameters than traditional architectures (Tan and Le, 2019). This efficiency is particularly beneficial in medical imaging, where the quality of data is crucial. EfficientNet-B4's compound scaling method ensures that each layer contributes optimally to the overall model performance, thereby improving the accuracy of classification tasks (Tan and Le, 2019). Additionally, its lightweight nature enables faster inference times, which is essential in clinical applications.

After the training is complete, validation set is used to evaluate the trained model performance.

Hair Removal

Hair removal techniques or preprocessing steps to isolate the lesion from surrounding hair can improve image clarity and analytic outcomes, ultimately enhancing model performance. In our project, hair removal from medical images is achieved through a series of systematic image processing techniques. The methodology begins with applying a BlackHat morphological operation, which highlights hair-like structures against the background. We define several critical parameters for this process:

- Edge detection thresholds: These are set to edge_low_threshold = 100 and edge_high_threshold = 220, which help to refine the detection of hair by focusing on edges that fall within this intensity range.
- Dark spot threshold: The threshold for identifying darker pixels is set at dark_spot_threshold = 150, effectively filtering out lighter noise and emphasizing hair structures.
- Line length threshold: Lines longer than linelength_threshold = 10 are prioritized, ensuring that only substantial hair features are considered.
- Divergence and patchiness thresholds: These are set at divergence_threshold = 0.25 and patchiness_threshold = 0.15, respectively, allowing for a better distinction between hair and skin by assessing the spatial distribution of detected lines.

Once potential hair regions are identified, we generate a mask and apply interpolation techniques to fill in the detected areas, ensuring the original image's structure remains intact. This approach is particularly suitable for our project as it allows for robust hair removal while maintaining the integrity of the skin lesions, which is critical for accurate diagnosis in medical imaging.

4.2 Data Collection

HAM10000 comprises a total of 10,015 dermatoscopic images for detecting pigmented skin lesions, including a representative collection of 7 diagnostic categories: Actinic keratoses and intraepithelial carcinoma (akiec), basal cell carcinoma (bcc), benign keratosis-like lesions (bkl), dermatofibroma (df), Melanoma (mel), melanocytic nevi (nv), and vascular lesions (vasc). However, these skin lesions are divided into 2 classes (melanoma and non-melanoma) by combining actinic keratoses and intraepithelial carcinoma, basal cell carcinoma, benign keratosis-like lesions, dermatofibroma, melanocytic nevi, and vascular lesions into one group.

4.3 Data Analysis

4.3.1 Distribution analysis of dataset attributes

This section shows an in-depth analysis of the dataset used in this study, with the objective of identifying key patterns and distributions of attributes.

Age is an important factor in the development and detection of melanoma. Research has shown that the incidence of melanoma increases with age, with older adults being at higher risk due to cumulative UV exposure (Australia, 2019). The dataset comprises age information for patients diagnosed with melanoma, ranging from a minimum age of 5 years to a maximum age of 85 years. The mean age of patients is 52 years, indicating that the majority of melanoma cases occur in middle-aged individuals.

The lesions are from various anatomical sites, mainly on humans back, lower extremity, trunk, upper extremity, and abdomen.

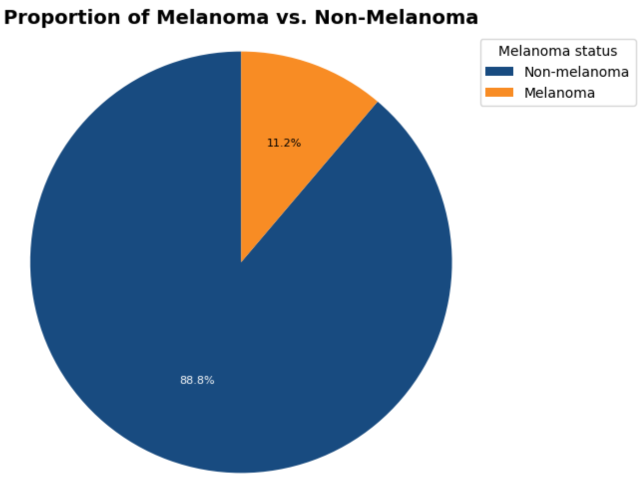


Figure 4.2: Proportion of melanoma and non-melanoma

Figure 4.2 displays the distribution of images classified as melanoma and non-melanoma. The data reveals a significant class imbalance, with only 11.2% of the images belonging to the melanoma class.

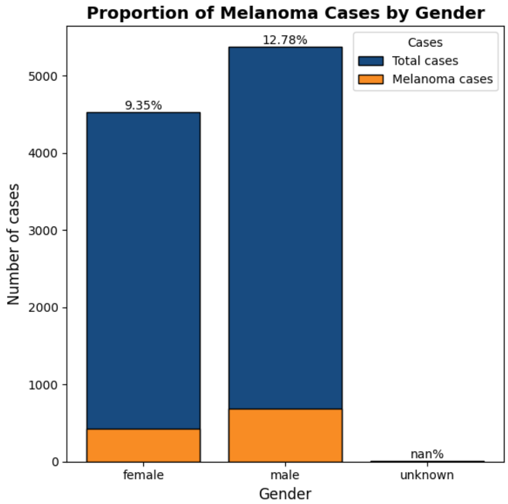


Figure 4.3: Proportion of melanoma cases by gender

Figure 4.3 reveals the gender-specific distribution of melanoma cases, where 9.35% of female patients and 12.78% of male patients are diagnosed with melanoma, aligns with broader global research findings indicating that melanoma incidence tends to be higher in men than in women (Bellenghi et al., 2020). This gender disparity in melanoma cases is seen worldwide and may reflect both gender-specific behaviors and intrinsic biological factors that influence melanoma development.

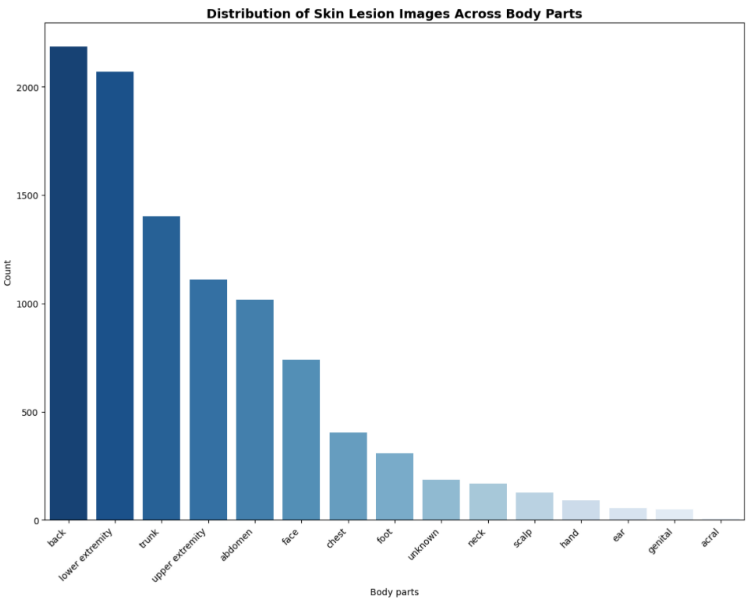


Figure 4.4: Distribution of skin lesion images across body parts

Figure 4.4 shows the distribution skin lesion images across various anatomical sites, with the majority of lesions located on the back, lower extremities, and trunk.

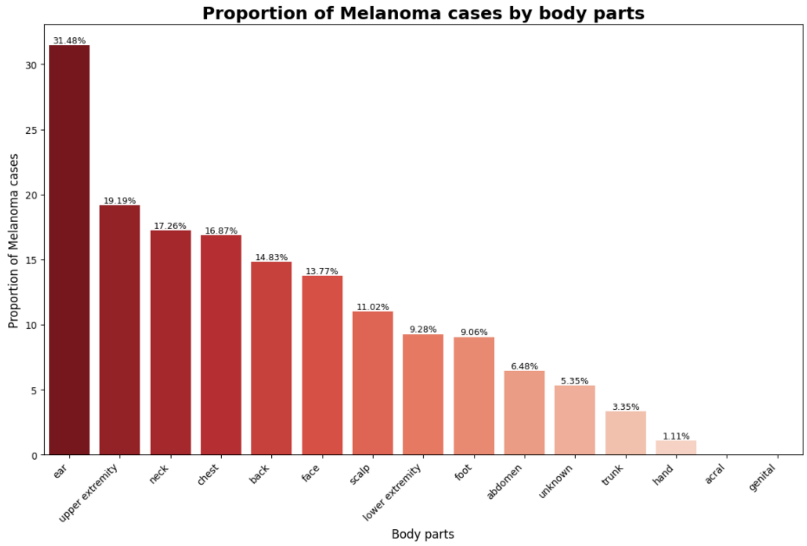


Figure 4.5: Proportion of melanoma cases by body parts

Figure 4.5 presents the proportion of melanoma cases by body part, with notable findings in regions such as the ear (31.48%), upper extremity (19.19%), and neck (17.26%). The high proportion of melanoma cases in these sun-exposed and vulnerable areas aligns with the findings of Wee, Wolfe, Mclean, Kelly, and Pan's study (2019), which show that locations like the male ear, upper back, and face are more likely to develop melanoma. Specifically, the research indicates that the ear and peri-auricular areas have higher rates of nodular melanoma, which is more invasive and has a tendency for thicker lesions when compared to other areas (Wee et al., 2019). This emphasizes how sun exposure contributes to melanoma risk particularly for areas like the ear and neck, which often receive sunlight yet are less likely to be protected by sunscreen or clothing.

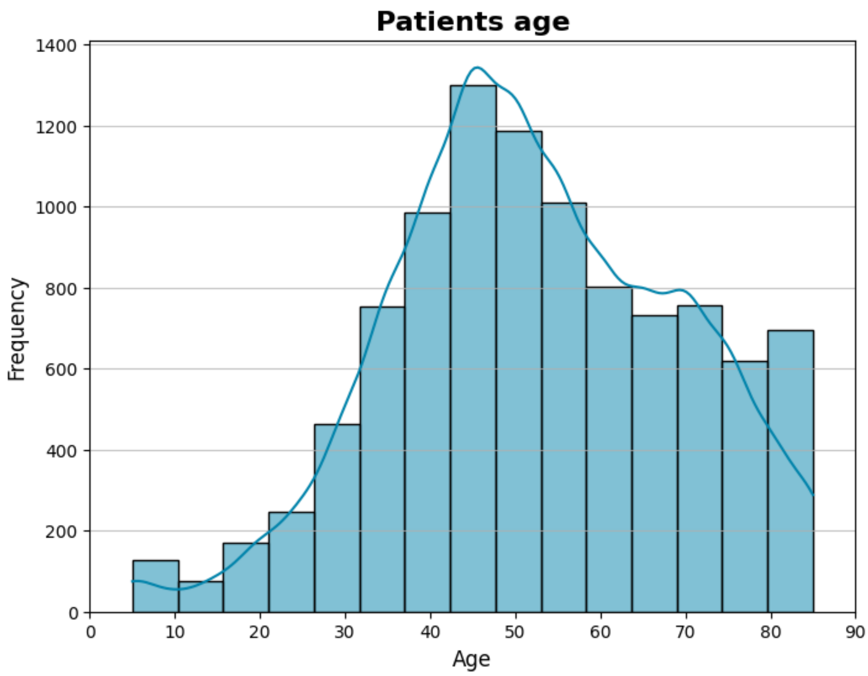


Figure 4.6: Distribution of patient's age

Figure 4.6 illustrates the age distribution of patients being surveyed in the study, which concentrates between 40 and 60 years old.

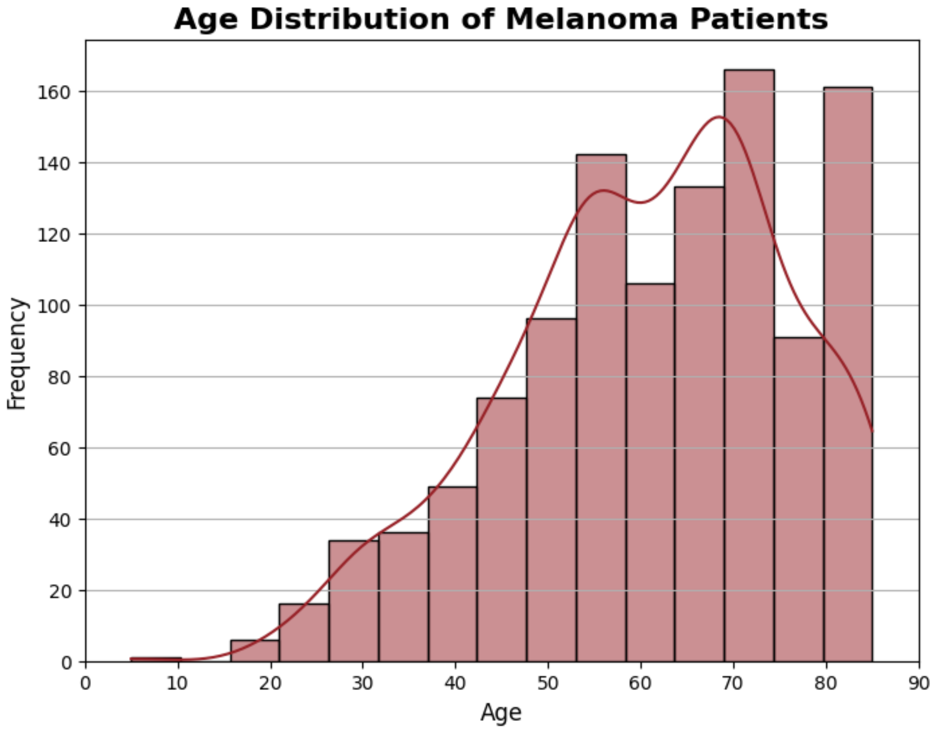


Figure 4.7: Age distribution of Melanoma patients

According to table 4.7, patients aged 55 and above are more likely to be diagnosed with melanoma. This is understandable as aging population are more vulnerable to skin cancer, as they tend to have accumulated sun damage and other risk factors over years. According to Ribero et al. (2018), people who have a high number of nevi after the age of 50 are at greater risk of developing melanoma, as their lesions tend to persist rather than diminish with age. The finding that melanoma significantly affects aging population is crucial for this study's predictive model development, as it highlights the need to accurately detect melanoma in the 55+ age group.

4.3.2 Image analysis

Class sizes

Table 4.2: Size properties lesions

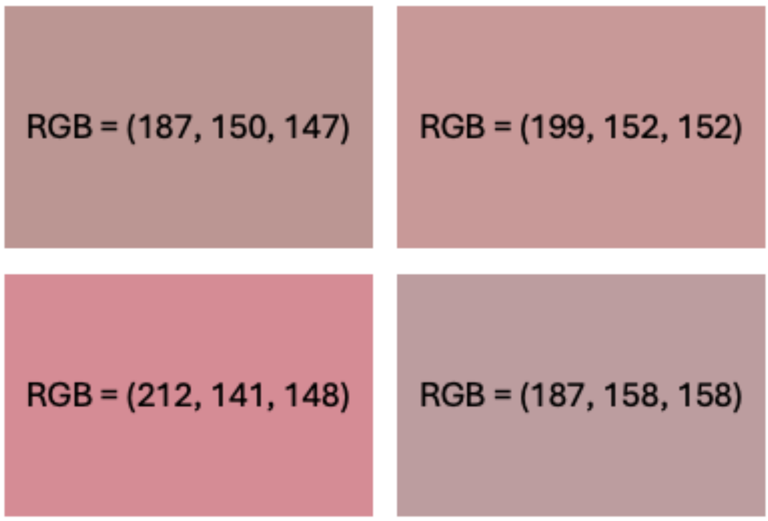
Class	Count	Avg area	Max area	Min area
Melanoma	1112	32.38%	98.86%	0.56%
Melanocytic nevi	6705	24.34%	99.15%	0.39%
Benign keratosis-like lesions	1098	31.9%	93.99%	0.85%
Basal cell carcinoma	514	26.01%	93.61%	0.58%

Actinic Keratoses	327	40.3%	97.04%	2.18%
Vascular lesions	142	15.2%	90.2%	0.96%
Dermatofibroma	115	31.19%	89.98%	2.67%

(Tschandl et al., 2021)

Skin tone analysis

Based on a study of Ahmed in 2022, red channel frequencies concentrated between 180-225 and green/blue channels between 135-165, the unimodal histogram shape indicates limited diversity in skin tones, skewed towards lighter tones. Figure 4.8 illustrates sample skin tones from the dataset, providing a visual overview of the range of skin colours represented. This finding is understandable given that images were sourced from patients in Vienna and Australia, two regions with relatively homogenous populations in terms of skin tone (Ahmed, 2022).



(Ahmed, 2022)

Figure 4.8: Sample skin tones

Presence of hair in Medical Imaging

In medical imaging of dermatological conditions, artifact such as hair is one of the most significant challenges to automatic segmentation of pigmented skin lesion images for diagnosis systems. Hair may cover parts of a lesion, interfere with edge detection, thereby mislead the deep learning models into learning patterns unrelated to the actual clinical features, reduce the accuracy. Therefore, it is necessary to address this issue, especially in datasets aimed at algorithm training for lesion detection.

SECTION 5.
Resources

The software that will be applied to this project has been listed below.

- Google Colab: Used for experimentation, data analysis, data preprocessing, and model training.
- Kaggle: Utilised for data collection.
- Github: Used for version control, collaboration, and hosting the application codebase, ensuring efficient development and project management.
- Streamlit: Used to develop an interactive user interface by integrating libraries and model outputs, enabling a user-friendly experience for application users.
- draw.io mind map: Utilised for task distribution, brainstorming, and note-taking in every meeting.

5.1 Materials

- Dermnet
- Cancer Council Victoria
- Cancer Council Australia
- Government Cancer Australia

5.2 Roles & Responsibilities

Table 5.1: Team roles and responsibilities

Number	Members	Roles	Responsibilities
1	Tran Minh Toan Hoang	Organiser Presenter Researcher Report Writer Data Scientist Project Reporter	• As an organiser, I always plan ahead for our project, team working process, task distribution and upcoming meetings or reports. I mainly built a mind map, which is a helpful tool that assists our team in understanding the project or assessments. • As a presenter, I prepare for slides, practice speaking and wrap up the ideas and results from our team to support the presentation. • As a researcher, I am responsible for the Literature Review, reading articles and books about Melanoma and Deep Learning that assist our team in understanding the problems or looking for new solutions. • As a report writer, I am responsible for the Literature Review, formatting the final report, checking grammar mistakes and submitting the report. • As a data scientist, my task is to collaborate with other data scientists to develop deep learning models and evaluate the results. • As a project reporter, I am the main one who reports to the mentor and prepares and updates the project progress in the meeting.
3	Ya-Ping Liao	Researcher Report writer Streamlit App Developer	• As a researcher, I focused on deep learning techniques, particularly their application in detecting skin cancer melanoma. Finding out the gap that our project can address in developing our application. • As a report writer, I am currently working on the milestone/schedule and literature review sections. In the future, I will write about the relevant parts of the project that I contributed to. • As a Streamlit app developer, I am creating a user-friendly interface that connects seamlessly to our model, making it easy for users to interact with and utilise the tool effectively.

2	Phuong Anh Pham	Librarian Data Collector Data Scientist Researcher Report writer	• As the librarian, I am responsible for setting up and organising document storage on Microsoft Teams, including uploading datasets, and documents and submitting assignment files. • As a data collector, I am responsible for collecting appropriate datasets and upload to Google Drive for use. • As a Data Scientist, my responsibility includes cooperating with other data scientists to develop various deep learning models and fine-tune the models. • As a researcher, my responsibility involves researching existing paperwork about Melanoma and other skin lesions and identifying gaps in current methodologies that our deep learning system can address. • As a report writer, I am responsible for the problem statement section, including defining the project's objective, scope, and research questions.
4	Rafia Tasneem	Data Collection Data Analysis Methodology	Developed the methodology section and selected the Convolutional Neural Network (CNN) model, specifically the EfficientNetB0 architecture. Oversaw data collection processing by dividing the SIIM-ISIC Melanoma Classification dataset into training and validation sets. Constructed the data preparation pipeline using augmentation techniques and the Adam optimizer to enhance the model's performance for binary classification.
5	Vicente Novoa	Researcher Report Writer Data Scientist	• General research to define the project aims and general objectives. • Writing the Expected Outcomes section. • Developing and testing a classification CNN to detect Melanoma in the selected dataset.
6	Peeranont Dongpakkij	Librarian Researcher Data Analyst	• Research the Melanoma skin disease and find the resources from the websites. • Compose the abstract for the project proposal. • Provide suggestions for this project and test the application before launching. • Provide the comment to make the project improvement.

SECTION 6.
Milestones / Schedule

Table 6.1: Project milestones

Milestone	Tasks	Date	Detail
Week-1	Project Selection Setup Workplace for the project	07/08/2024	Setup Meeting on Zoom to discuss the project topic
Week-2	Project Understanding Project Brainstorming First meeting with mentor William	14/08/2024	Define the project problems, target stakeholders, methodology and product delivery.
Week-3	Literature Review Data Collection Plan the direction to build the models	21/08/2024	This week, team DLSK started collecting data from many sources, such as Kaggle or Hugging Face, and discussed planning to build the models.
Week-4	Data Collection Literature Review	28/08/2024	This week our team continue working on data collection and literature review. Kicking off Proposal Report due in the week after.
Week-5	Group Proposal Report Due	8/09/2024	Finish the proposal report and submit; also, discuss with mentor William about the target delivery and dataset size.
Week-6	Develop and Improve Models Develop Streamlit App	11/09/2024	Begin developing the Streamlit app following William's recommendations.
Week-7	Develop and Improve Models Develop Streamlit App	18/09/2024	Continue working on Models and developed, deployed Streamlit App.
Week-8	Presentation of the demo Streamlit App and Models	25/09/2024	Consider the suggestions provided by alumni and make corresponding changes, thinking about how to improve our project.
Week-9	Progress Report Due Complete Dataset for models Streamlit App deployment	13/10/2024	Finish the progress report and submit. Discuss model results and the app with William. The app has been successfully deployed, with all functionalities working as expected, and is accessible via a customed link.
Week-10	Continue developing the UI of	16/10/2024	Complete the models and start training on the final dataset.

	Streamlit Training models on the final dataset Prepare slides for the presentation in week 11		Enhance and develop the UI of Streamlit App by customising the theme colour, modifying the page options using a navigation method, adding group member information, creating a news page using the News API, and replacing dropdown menus with buttons to simplify user input.
Week-11	Group Oral Presentation Finalise the project final report	23/10/2024	Prerecording presentation and distribute tasks in the final report. Also, prepare the Showcase slides and demo.
Week-12	Group Final Report Due	6/11/2024	Wrapped up all material and achievements and finalised the final report.

SECTION 7.
Results

In this part, we aim to deliver the outcome we achieved in this project. We discuss two parts of our results. On one hand, we illustrated the result of our Deep Learning Model. On the other hand, we expounded the outcome of our product - Streamlit App.

Model

We experimented with different deep learning alternatives like EfficientNetB4, ResNet50 and DenseNet. After careful analysis and many model iterations with different parameters, we decided on a certain model architecture and parameters:

Layer (type)	Output Shape	Param #	Connected to
image_input (InputLayer)	(None, 224, 224, 3)	0	-
conv2d_9 (Conv2D)	(None, 222, 222, 32)	896	image_input[0][0]
max_pooling2d_9 (MaxPooling2D)	(None, 111, 111, 32)	0	conv2d_9[0][0]
batch_normalization_9 (BatchNormalization)	(None, 111, 111, 32)	128	max_pooling2d_9[0][0]
conv2d_10 (Conv2D)	(None, 109, 109, 64)	18,496	batch_normalization_9…
max_pooling2d_10 (MaxPooling2D)	(None, 54, 54, 64)	0	conv2d_10[0][0]
batch_normalization_10 (BatchNormalization)	(None, 54, 54, 64)	256	max_pooling2d_10[0][0]
conv2d_11 (Conv2D)	(None, 52, 52, 128)	73,856	batch_normalization_1…
max_pooling2d_11 (MaxPooling2D)	(None, 26, 26, 128)	0	conv2d_11[0][0]
attribute_input (InputLayer)	(None, 17)	0	-
batch_normalization_11 (BatchNormalization)	(None, 26, 26, 128)	512	max_pooling2d_11[0][0]
dense_9 (Dense)	(None, 16)	288	attribute_input[0][0]
flatten_3 (Flatten)	(None, 86528)	0	batch_normalization_1…
dropout_6 (Dropout)	(None, 16)	0	dense_9[0][0]
concatenate_3 (Concatenate)	(None, 86544)	0	flatten_3[0][0], dropout_6[0][0]
dense_10 (Dense)	(None, 32)	2,769,440	concatenate_3[0][0]
dropout_7 (Dropout)	(None, 32)	0	dense_10[0][0]
dense_11 (Dense)	(None, 2)	66	dropout_7[0][0]

Figure 7.1: The model architecture and parameters

With it, we achieved the following results:

7.1 Accuracy over epochs

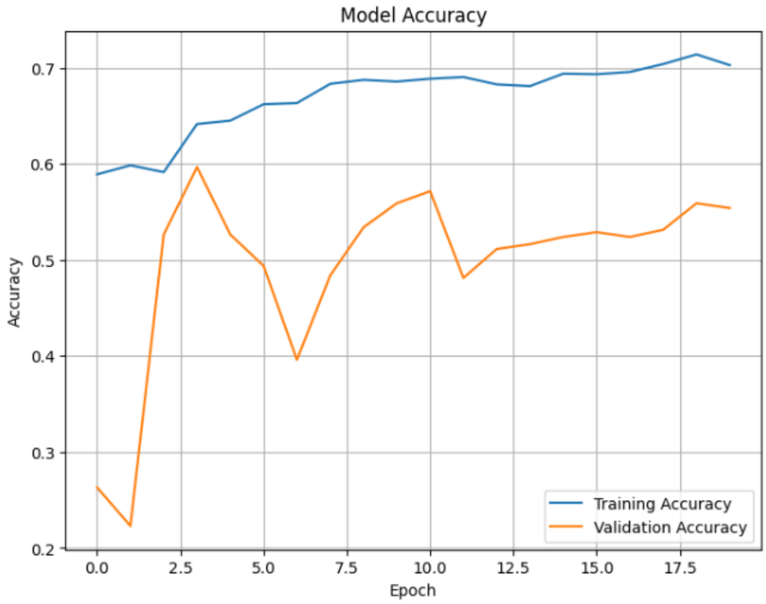


Figure 7.2

As Figure 7.2 shows, both accuracies grow consistently higher along the epochs, which is what one would expect in a good working model. In previous attempts, with other structures and parameters this wasn't happening. In many previous iterations the validation accuracy dropped straight down in the latest epochs, suggesting overfitting.

7.2 Classification Report 1

	precision	recall	f1-score	support
class_0	1.00	0.48	0.65	343
class_1	0.24	1.00	0.39	56
accuracy			0.55	399
macro avg	0.62	0.74	0.52	399
weighted avg	0.89	0.55	0.61	399

Figure 7.3

As Figure 7.3 shows, our model achieved an accuracy of 55% and a recall of 100% on Class 1 which are the metrics on which we must focus more on this project. A recall of 100% is of course, an amazing result, but a 55% accuracy is far from ideal. Thus, we adjusted the prediction threshold so that the model is less likely to predict Class 1 instances. In other words, what this does, is that it makes the model only predict Class 1 when it is very certain that the instance is Class 1.

7.3 Classification Report 2

	precision	recall	f1-score	support
class_0	0.91	0.83	0.87	343
class_1	0.33	0.52	0.41	56
accuracy			0.79	399
macro avg	0.62	0.67	0.64	399
weighted avg	0.83	0.79	0.81	399

Figure 7.4

Therefore, as Figure 7.4 shows, we achieved more balanced results. Our final model achieved a 79% accuracy and a 52% recall on Class 1, which are overall very good results. The only problem now is that a 52% recall may not be that good and it might be particularly harmful since we are dealing with a very serious disease.

7.4 Final approach

That is why we decided to take a different approach to this problem and go from predicting two classes, "Melanoma"/"No Melanoma", to predicting three classes, "Low Risk", "Medium Risk", and "High Risk".

Streamlit App

Our app aims to provide suggestions, raise awareness of cancer, and bridge the gap between dermatologists and patients.

There are three main functions and five pages in the DLSD Streamlit app, which are Melanoma Detection, Book and Appointment, Cancer News, About This Project and Group Member.

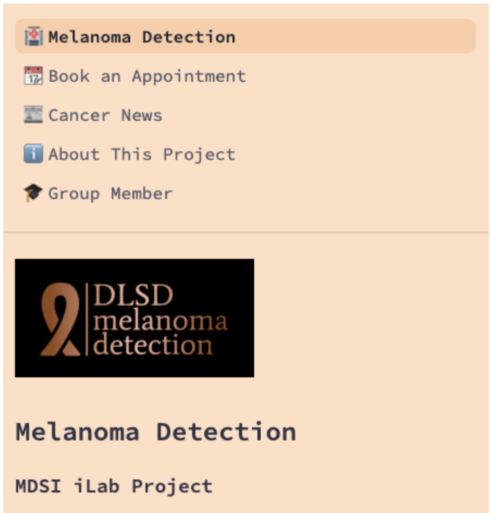


Figure 7.5: App Page Navigation UI

The first feature is associating the deep learning model to differentiate Melanoma. This feature can let our user input their age and where the potential Melanoma is located. Since medical terminology is complicated for users to understand, we added explanations to improve user experience. Furthermore, we added instructions on what example photos to upload or take a picture of.

Figure 7.6: App features

Afterwards, the user can upload an image or take a photo of their skin to have the potential low, medium, and high risk of Melanoma. We will transform the result into three categories - Low, Medium, and High risk of Melanoma. If the user unfortunately got the result for the high-risk Melanoma, we provided a button that can directly go to the booking page to accelerate the early treatment.

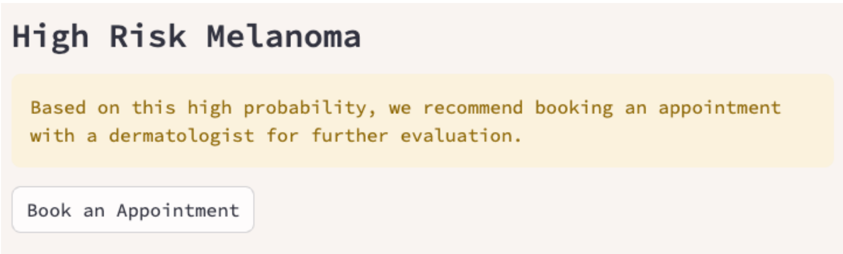


Figure 7.7: High risk result in the UI

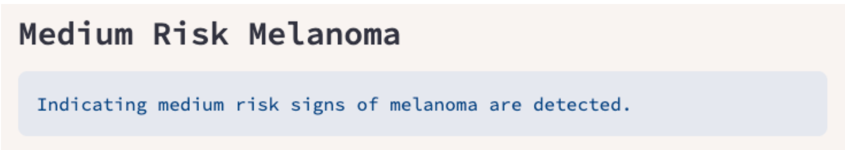


Figure 7.8: Medium risk result in the UI

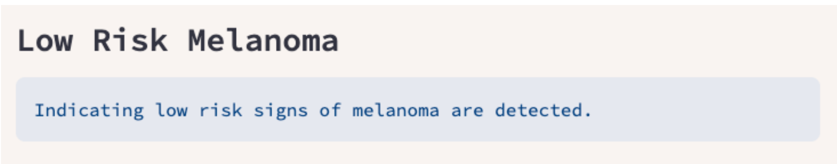


Figure 7.9: Low risk result in the UI

The second page**,** The Book an Appointment page, the main feature provides a Proof of Concept (POC) to build a booking system for the convenience of patients side to book an appointment quickly for the clinic side; they can easily manage the booking and also can get the patient information, photo in advance can do the management and diagnose in advance.

The consent form is needed on this page to inform users that their data will be used solely for medical purposes within the clinic, ensuring transparency, protecting privacy, and obtaining their permission before data collection in the next step.

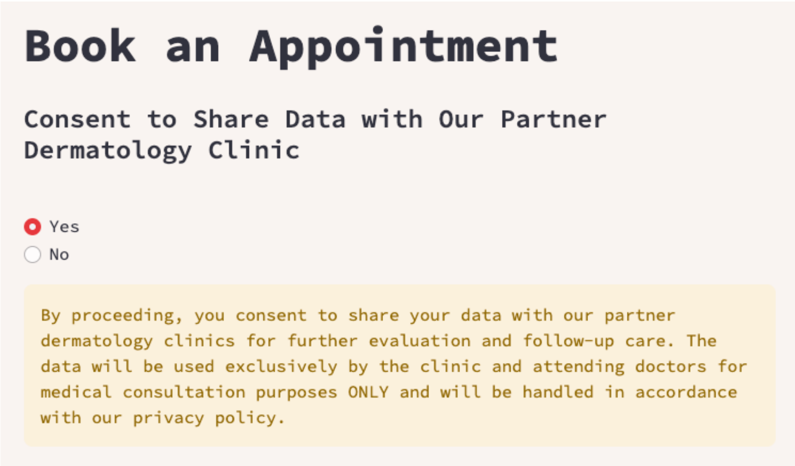


Figure 7.10

The second column is for the user to fill in their email address; we consider Personally Identifiable Information PII information. Instead of using the name, we use email to book an appointment and send booking confirmation. On the clinic side, they can only see the booking ID to avoid directly exposing their complete identity.

Figure 7.11

The remaining questions are as follows: gender, with an option for the user to either upload a photo or take one using the camera. This photo will be included in the PDF and sent to the appointed dermatologist in advance for diagnosis.

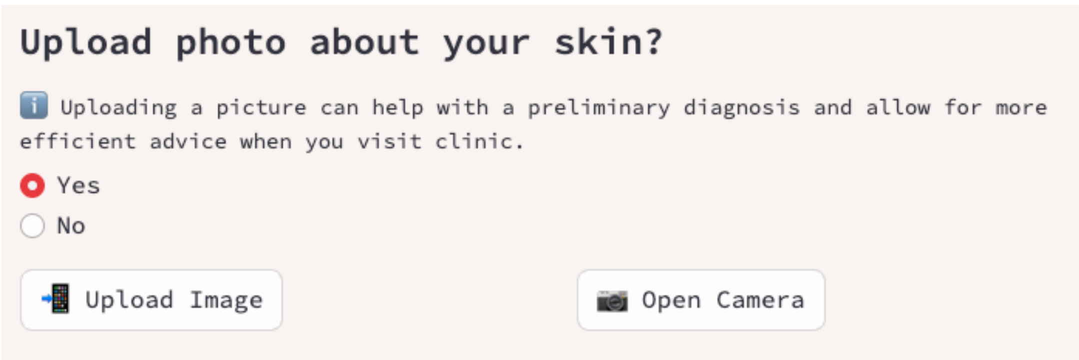


Figure 7.12

Another section asks for symptom details, such as the location of the potential melanoma, the duration of the symptom, and whether the user experiences any pain.

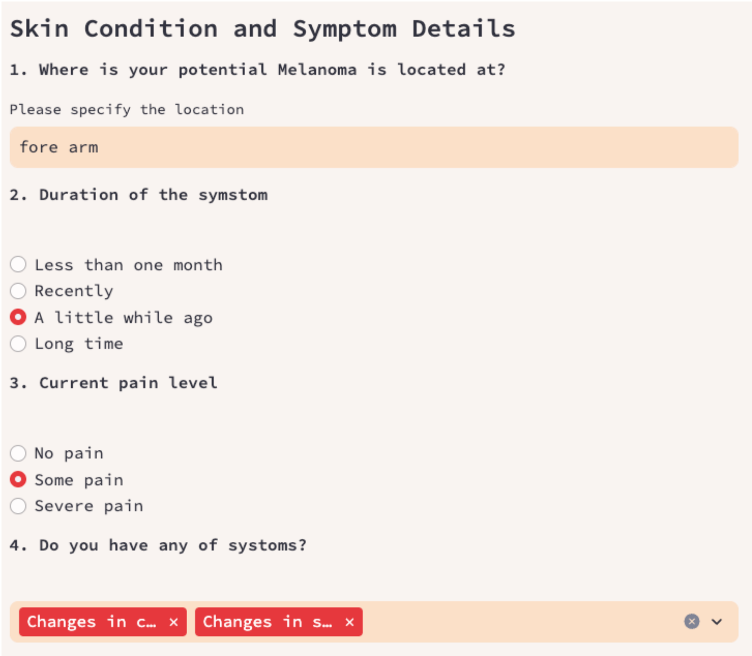


Figure 7.14

The final section allows the user to choose a partner clinic, select a specific dermatologist, and pick an available time slot.

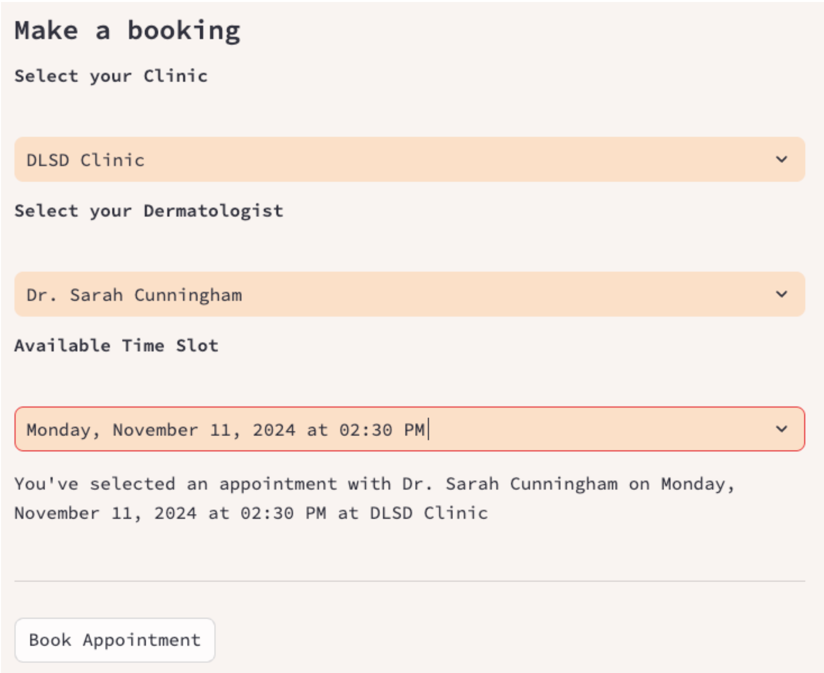


Figure 7.14

Afterward, the user will receive an email confirming the appointment time, while the clinic will receive an email with the booking details and an attached PDF containing the patient's information.

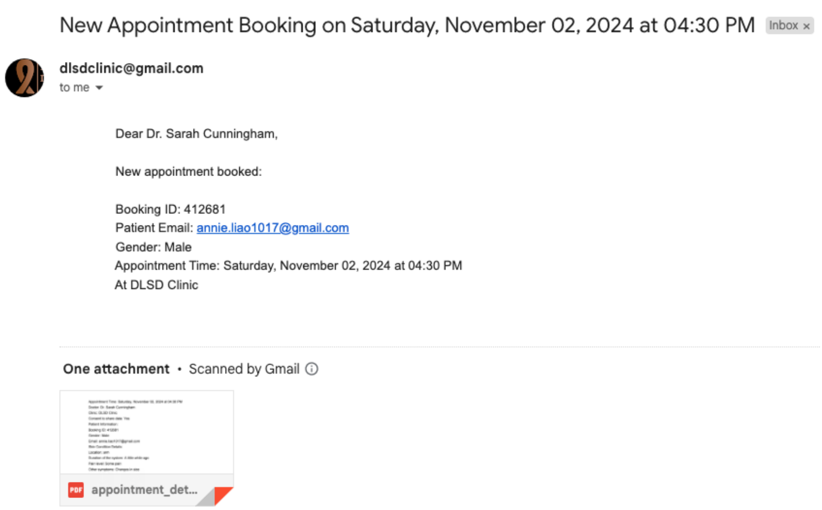


Figure 7.15

On the third page, our app integrates with the News API to retrieve cancer-related news updates from sources such as BBC. This feature aims to enhance cancer awareness by providing users with relevant, up-to-date information. By including a dedicated news section, we hope to keep users informed on the latest research, prevention strategies, and advancements in cancer treatment, supporting a broader understanding of health topics that may be beneficial for early detection and overall well-being.

The fourth page presents our motivation for undertaking this project. As noted, Australia faces a critical challenge with skin cancer, particularly melanoma. Our goal is to integrate a deep learning model that provides users with tools to increase their awareness and understanding of melanoma risks.

Finally, the last page introduces our project team members, including photos and brief biographies to share our backgrounds and roles in the project.

In conclusion of the result, our project addresses the growing challenge of melanoma by developing an AI-driven app to assist in early detection and support diagnostic efforts. The app, consisting of five fully functional pages, provide a user-friendly interface that lets users interact with essential features intended to speed up the diagnosis procedure. Users can input symptom details, upload or capture photos for analysis, and book appointments with partner clinics and dermatologists. Central to the app's functionality is a deep learning model that analyses uploaded images to assess potential melanoma risk. This model aims to reduce diagnostic delays by providing a preliminary analysis directly through the app, giving dermatologists a reliable tool to support their evaluations. Additionally, the app enhances awareness by offering users access to curated cancer-related news, which helps in promoting knowledge and early vigilance against skin cancer. By integrating this AI model within a comprehensive app, our project provides an accessible, efficient tool that aids both patients and healthcare providers. It not only helps in early detection by supplying an initial diagnostic indicator but also streamlines appointment scheduling and record-keeping, ultimately supporting timely, data-driven decisions for melanoma care.

SECTION 8.
Discussion

Model

After trying different models and iterating with many different structures and parameters, we ended up with a Convolutional Neural Network model that met our needs. As shown in figure 7.3, the problem is that this model had a very high False Positive Rate. Therefore, we had to adjust the prediction threshold so that it less likely to predict positive instances.

Once we did that, we achieved a much higher accuracy of 79%. Then, the problem was that the recall of 52% was indicating that we were only predicting half of the actual melanomas, which in the case of a serious disease like melanoma, is very dangerous. Thus, we found an approach that could help resolve this problem. Instead of predicting just two classes: “Melanoma”/ “No Melanoma”, we found that it was better to predict three classes: “Low Risk”, “Medium Risk” and “High Risk”. This way, we achieved the following results:

Bins	Total	Melanoma	Rate %
Low Risk	166	0	0%
Medium Risk	204	42	21%
High Risk	29	14	48%
Grand Total	**399**	**56**	**14%**

Figure 8.1

Thus, if a customer gets the “Low Risk” response, it means that the probability of having a melanoma is very close to 0%. On the other hand, if a customer gets the “High Risk” result, it means that the probability of having a melanoma is almost 50%, so that customer is strongly recommended to do something about it. Finally, if a customer gets the “Medium Risk” result, it means that the probability of having a melanoma is 20%, which is high, but not that high, so he probably should do something about but not as urgently as the “High Risk” group. In other words, the approach of risk groups allows us to warn everybody that needs to be warned, but also give them a sense of urgency that corresponds with their situation.

Streamlit App

On the first page of the app, users can access real-time detection of their skin symptoms, allowing them to monitor their skin health proactively. This feature aims to increase user awareness of potential skin issues and provide an automated preliminary assessment, thereby bridging the gap in early symptom recognition. While our project primarily focuses on melanoma detection, we have incorporated additional functionalities to enhance the app's overall effectiveness and user experience.

The integrated booking system is designed to facilitate communication between dermatologists and patients, streamlining the appointment process. To support this functionality, we implemented the Simple Mail Transfer Protocol (SMTP), which enables the app to send real-time email notifications to both the clinic and the patient. This ensures that both parties receive timely updates regarding appointment details, thereby improving coordination and communication. The incorporation of SMTP not only enhances the app's reliability but also addresses the critical gap of efficient information exchange in the healthcare process. By facilitating smoother interactions between users and healthcare providers, our app contributes significantly to the overarching goal of improving melanoma diagnosis and patient care.

In addition to these features, the app integrates with the News API to provide users with access to current news articles about cancer awareness and prevention. This feature informs users about the latest research findings, treatment advancements, and preventive measures related to melanoma and other skin cancers by curating relevant information from reputable sources, such as BBC. This fosters a proactive approach to health management enables consumers to make knowledgeable choices regarding the health of their skin. By enhancing the user experience through timely information, the News API functionality complements the app's core aim of improving melanoma diagnosis. It encourages users to engage more actively in their health journeys.

SECTION 9.

Limitations and Future Works

Limitations

Limitations - Data

The choice of dataset for training a diagnostic model directly influences its accuracy, diversity, and ethical reliability; hence, evaluating the implications of selecting HAM10000 is essential for understanding the scope and limitations of the data. One of the weaknesses of the HAM10000 dataset pertains to its representation of diverse skin tones among patients, as fewer than 5% of the images came from dark-skinned patients (Morales-Forero et al., 2024). Compared to White patients, Black patients are more likely to be diagnosed with advanced-stage melanomas that are deeper and have a higher likelihood of being spread to distant parts of the body (Shao & Feng, 2022). This issue might raise concerns about the model's ability to generalize across different demographics. Another limitation pertains to the metadata, as it contains poor patient information limited to only gender, age, and lesion localization, lacking essential details like ethnicity, skin type, and clinical history. Moreover, if the metadata quality is inconsistent, the model may unintentionally learn biased associations, such as linking specific lesion types with genders or age groups disproportionately, which could negatively impact diagnostic fairness (Singh et al., 2024). Additionally, although the dataset includes diverse lesion types and patient demographics, relying on a single dataset may lead to dataset bias because each dataset represents only a subset of the real-world conditions, limiting the model's robustness and adaptability across diverse applications. One fundamental consideration would be the privacy concerns, when the use of data in healthcare poses significantly privacy challenges. Although the data is anonymized, there is still a risk of re-identification. Therefore, it is necessary to implement privacy-preserving methods to avoid unauthorized access to personal health information (Singh et al., 2024).

Limitations - Model

The primary limitation of our model stems from the restricted computational resources available through the free version of Google Colab. This constraint affects the model in three key ways: first, it limits the volume of data we can use for training, which can reduce model performance due to the smaller dataset. Second, it necessitates a more conservative model architecture to prevent excessive RAM usage that could otherwise terminate the Colab session. Lastly, using limited training data increases the model's susceptibility to overfitting, requiring us to carefully adjust the architecture to achieve an optimal balance between underfitting and overfitting.

Limitations - Streamlit App

There are four points about our application limitations, including limited scope of detection, inefficiencies in the booking system, news categorisation issues with NewsAPI, and lack of comprehensive educational content.

Limited Scope of Detection: Currently, our model is designed exclusively to detect melanoma, which limits its utility for users who may be concerned about other forms of skin cancer. As a result, the app can only provide suggestions for melanoma, and users may not receive comprehensive diagnostic support for other skin conditions. This limitation impacts the effectiveness of the app in addressing a broader spectrum of skin health concerns.

Inefficiencies in the Booking System: Although the booking system successfully sends email notifications via SMTP, it could be further refined to enhance efficiency. The current system requires manual oversight by clinic staff to manage appointments, which may lead to delays and increased administrative burden. Implementing a more sophisticated, automated booking system that factors in geographic location could provide users with convenient options for selecting clinics, hospitals, and dermatologists, ultimately improving the user experience and reducing management time.

News Categorisation Issues with News API: The integration of the News API presents challenges related to the categorization of news articles. Currently, the API sometimes returns unrelated news items, which can dilute the relevance of the information provided to users. Unfortunately, since we do not have control over the categorization features of the News API, it limits our ability to ensure that users receive only pertinent and informative content related to melanoma awareness and prevention.

Lack of Comprehensive Educational Content: While the app raises awareness about melanoma, it could benefit from more robust educational features. Currently, there is a lack of in-depth information on melanoma prevention, risk factors, and early detection strategies. Expanding the app's content to include additional resources and features focused on melanoma education could enhance user knowledge and promote proactive health management.

Future works

Future works for data

To overcome the current data limitations, future work should focus on expanding, enhancing, and securing the dataset used in training.

One essential step is to incorporate additional datasets beyond HAM10000 to acquire correlations and patterns that may remain obscure within a single dataset (Kaplan, 2024), thereby minimize dataset bias and enhance model robustness. By combining multiple datasets with diverse patient populations and lesion types, the model can potentially capture subtle differences across skin types, thereby generalize to broader populations, enhance diagnostic accuracy, as cross-validation across datasets helps to validate and verify results, reducing the likelihood of overfitting to specific population groups (Kaplan, 2024). To implement this, techniques such as concatenation, joining, and merging are suitable and commonly used. Concatenation allows for stacking datasets with similar structure; while joining and merging enable the integration of datasets with partially overlapping information, leveraging common attributes to create a unified dataset. This strategy not only improves model accuracy but also provides a competitive edge by developing a diagnostic tool that is responsive to diverse population needs (Kaplan, 2024).

Regarding the issue of skin tone imbalance within the dataset, we implemented a simple solution which was to adjust the image brightness with a purpose of mimicking diversity in skin tone. Although this approach would be useful to a certain extent, it may not accurately reflect the true characteristics of skin lesions on dark-skinned individuals. Therefore, to address the lack of dark-skinned images in the dataset to ensure that the models are equipped to recognized and diagnosed melanoma across all skin tones, targeted data collection efforts should be implemented. One possible strategy could be collaboration with healthcare facilities that serve diverse populations to collect additional sample of dermatoscopic images.

Additionally, future studies should aim to broaden the dataset's metadata to include a more comprehensive range of patient characteristics, such as ethnicity, skin type, and relevant clinical histories, which would enhance diagnostic accuracy across diverse groups, reduce bias in model training, and contribute to improving personalisation of care by providing recommendations or screenings tailored to individual risk profiles.

When using healthcare data, privacy is utmost importance due to the high risk of exposing sensitive patient information. Even when data is anonymized, re-identification risks persist, necessitating stringent privacy measures. Techniques like differential privacy, which introduces controlled noise to protect individual data points, and federated learning, maintains data on local devices, can mitigate unauthorized access concerns. Adopting such privacy-preserving methods not only reduces re-identification risks but also boost patient trust by ensuring compliance with ethical and regulatory standards in healthcare research.

Future works for the model

To enhance model performance and overcome the limitations posed by computational constraints, future work should focus on improving access to computational resources. Using a paid or more powerful cloud computing platform, for example, would enable the use of larger datasets, resulting in a more robust and generalizable model. With additional resources, more complex model architectures could also be explored, potentially leading to higher accuracy and nuanced learning without the risk of memory-related session interruptions. These advancements would allow us to optimize both model complexity and training data volume for more reliable predictions.

Future works for the Streamlit App

As for our application limitation, there are four points of view we can improve on and have the potential implementation and research in the application.

Integration of additional models for differential diagnosis: One of the primary goals for future development is to incorporate additional machine learning models to detect various skin cancers. By expanding the model's capabilities beyond melanoma, we can provide users with a more comprehensive diagnostic tool that addresses a broader spectrum of skin conditions. This approach will improve user satisfaction and enhance the app's overall effectiveness in promoting skin health.

Enhancement of the booking system: To streamline the appointment scheduling process, future work will involve developing a more sophisticated booking system. This system should automate various aspects of appointment management, reducing the manual oversight currently required by clinic staff. Implementing geographic location-based clinic recommendations and available time slots can significantly enhance user convenience and satisfaction.

Establishment of a comprehensive database: Building a robust database will allow dermatologists to access historical patient data easily, improving diagnostic accuracy and treatment planning. This database should include patient profiles, previous consultations, and treatment history, enabling healthcare providers to make more informed decisions. Additionally, the clinic administration should have access to this database to manage booking details and patient records efficiently.

Collaboration with healthcare experts: Collaborating with healthcare professionals, including dermatologists and oncologists, is essential. Working with experts in the field will help us to integrate best practices and ensure that the application delivers correct and sufficient information. This partnership can also guide the development of educational resources within the app, empowering users with knowledge about skin cancer prevention, early detection, and treatment options.

SECTION 10.
References

Agrawal, N., & Aurelia, S. (2021). Corroboration of skin Diseases: Melanoma, Vitiligo & Vascular Tumor using Transfer Learning. https://doi.org/10.1109/icees51510.2021.9383682

Awadh Alanazi, S. (2022). Melanoma Identification Through X-ray Modality Using Inception-v3 Based Convolutional Neural Network. Computers, Materials & Continua, 72(1), 37–55. https://doi.org/10.32604/cmc.2022.020118

Bellenghi, M., Puglisi, R., Pontecorvi, G., De Feo, A., Carè, A., & Mattia, G. (2020). Sex and Gender Disparities in Melanoma. Cancers, 12(7). https://doi.org/10.3390/cancers12071819

C. Kavitha, S. Priyanka, M. Praveen Kumar, & Kusuma, V. (2024). Skin Cancer Detection and Classification using Deep Learning Techniques. Procedia Computer Science, 235, 2793–2802. https://doi.org/10.1016/j.procs.2024.04.264

Demir, A., Yilmaz, F., & Kose, O. (2019). Early detection of skin cancer using deep learning architectures: resnet-101 and inception-v3. 2019 Medical Technologies Congress (TIPTEKNO). https://doi.org/10.1109/tiptekno47231.2019.8972045

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level Classification of Skin Cancer with Deep Neural Networks. Nature, 542(7639), 115–118. https://doi.org/10.1038/nature21056

Gajera, H. K., Nayak, D. R., & Zaveri, M. A. (2023). A comprehensive analysis of dermoscopy images for melanoma detection via deep CNN features. Biomedical Signal Processing and Control, 79, 104186. https://doi.org/10.1016/j.bspc.2022.104186

Ghosh, S., Dhar, S., Raktim Yoddha, Kumar, S., Abhinav Kumar Thakur, & Nanda Dulal Jana. (2024). Melanoma Skin Cancer Detection Using Ensemble of Machine Learning Models Considering Deep Feature Embeddings. Procedia Computer Science, 235, 3007–3015. https://doi.org/10.1016/j.procs.2024.04.284

Huang, A., Kwan, S.-Y., Chang, W.-Y., Liu, M.-Y., Chi, M.-H., & Chen, G.-S. (2013). A robust hair segmentation and removal approach for clinical images of skin lesions. Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference, 2013, 3315–3318. https://doi.org/10.1109/EMBC.2013.6610250

Jain, S., & Agrawal, K. (2023). An Efficient Diagnosis of Melanoma Skin Disease Using DenseNet-121. https://doi.org/10.1109/ictacs59847.2023.10390147

Kaur, G., Sharma, N., Chauhan, R., Rawat, D., & Gupta, R. (2023). Early Detection of Melanoma using EfficientNet B4 Transfer Learning Model. https://doi.org/10.1109/gcitc60406.2023.10426141

Kaushik, P., Rathore, R., Kumar, A., None Kanishka, Ganpat Goshi, & Sharma, P. (2024). Identifying Melanoma Skin Disease Using Convolutional Neural Network DenseNet-121. https://doi.org/10.1109/iatmsi60426.2024.10502880

Li, C., Qiao, Z., Wang, K., & Jiang Hongxing. (2021). Improved EfficientNet-B4 for Melanoma Detection. https://doi.org/10.1109/icbaie52039.2021.9389915

Li, L.-F., Wang, X., Hu, W.-J., Xiong, N. N., Du, Y.-X., & Li, B.-S. (2020). Deep Learning in Skin Disease Image Recognition: A Review. IEEE Access, 8, 208264–208280. https://doi.org/10.1109/access.2020.3037258

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A Survey on Deep Learning in Medical Image Analysis. Medical Image Analysis, 42(1), 60–88. https://doi.org/10.1016/j.media.2017.07.005

Morales-Forero, A., Rueda Jaime, L., Sebastian Ramiro Gil-Quiñones, Barrera Montañez, M. Y., Bassetto, S., & Coatanea, E. (2024). An insight into racial bias in dermoscopy repositories: A HAM10000 data set analysis. JEADV Clinical Practice. https://doi.org/10.1002/jvc2.477

Morid, M. A., Borjali, A., & Del Fiol, G. (2021). A scoping review of transfer learning research on medical image analysis using ImageNet. Computers in Biology and Medicine, 128, 104115. https://doi.org/10.1016/j.compbiomed.2020.104115

Patinge, S., Ponde, K., Pokar, J., Gadgile, V., Raut, M., & Philips, N. (2024). Exploring Pre-Trained CNN Models for Classification of Melanoma. 2024 International Conference on Expert Clouds and Applications (ICOECA), 1000–1005. https://doi.org/10.1109/icoeca62351.2024.00174

Ribero, S., Stucci, L. S., Marra, E., Marconcini, R., Spagnolo, F., Orgiano, L., Picasso, V., Queirolo, P., Palmieri, G., Quaglino, P., & Bataille, V. (2018). Effect of Age on Melanoma Risk, Prognosis and Treatment Response. Acta Dermato-Venereologica, 98(7), 624–629. https://doi.org/10.2340/00015555-2944

Roy, S. S., Kardan, R., & Neubert, J. (2024). A Deep Learning-Based Model for Melanoma Detection in Both Dermoscopic and Digital Images. 2024 IEEE International Conference on Electro Information Technology (EIT), 668–673. https://doi.org/10.1109/eit60633.2024.10609889

Shao, K., & Feng, H. (2022). Racial and Ethnic Healthcare Disparities in Skin Cancer in the United States: A Review of Existing Inequities, Contributing Factors, and Potential Solutions. The Journal of Clinical and Aesthetic Dermatology, 15(7), 16. https://pmc.ncbi.nlm.nih.gov/articles/PMC9345197/

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. Journal of Big Data, 6(1). https://journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0197-0

Singh, M., Sharma, K., & Baghel, S. (2024). Ethical and Privacy Considerations in Medical Big Data: Balancing Innovation and Safeguarding Patient. International Journal of Research Publication and Reviews, 5(5), 10368–10375. https://doi.org/10.55248/gengpi.5.0524.1387

Suleman, M., Ullah, F., Ghadah Aldehim, Shah, D., Abrar, M., Irshad, A., & Sarra Ayouni. (2023). Smart MobiNet: A Deep Learning Approach for Accurate Skin Cancer Diagnosis. Computers, Materials & Continua/Computers, Materials & Continua (Print), 77(3), 3533–3549. https://doi.org/10.32604/cmc.2023.042365

Sultana, J., Saha, B., Khan, S., T.M Sanjida, Hasan, M., & Mohammad Monirujjaman Khan. (2022). Identification and Classification of Melanoma Using Deep Learning Algorithm. 2022 IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE). https://doi.org/10.1109/icdcece53908.2022.9792698

Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. (2017). Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. Proceedings of the AAAI Conference on Artificial Intelligence, 31(1). https://doi.org/10.1609/aaai.v31i1.11231

Tan, M., & Le, Q. (2020). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. https://arxiv.org/pdf/1905.11946

Tschandl, P., Rosendahl, C., & Kittler, H. (2018). The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Scientific Data, 5(1). https://doi.org/10.1038/sdata.2018.161

Tschandl, P., Rosendahl, C., Kittler, H., & et al. (2021). Skin cancer dataset (HAM10000). DatasetNinja. https://datasetninja.com/skin-cancer-ham10000

V. Vipin, Malaya Kumar Nath, V. Sreejith, Nikhil Francis Giji, Ramesh, A., & Meera Nair M. (2021). Detection of Melanoma using Deep Learning Techniques: A Review. 2021 International Conference on Communication, Control and Information Sciences (ICCISc). https://doi.org/10.1109/iccisc52257.2021.9484861

Wee, E., Wolfe, R., Mclean, C., Kelly, J. W., & Pan, Y. (2019). The anatomic distribution of cutaneous melanoma: A detailed study of 5141 lesions. Australasian Journal of Dermatology, 61(2), 125–133. https://doi.org/10.1111/ajd.13223

Ye, Z., Zhang, D., Zhao, Y., Chen, M., Wang, H., Seery, S., Qu, Y., Xue, P., & Jiang, Y. (2024). Deep learning algorithms for melanoma detection using dermoscopic imaging: A systematic review and meta-analysis. Artificial Intelligence in Medicine, 102934–102934. https://doi.org/10.1016/j.artmed.2024.102934

Schadendorf, D., Fisher, D. E., Garbe, C., Gershenwald, J. E., Grob, J., Halpern, A., Herlyn, M., Marchetti, M. A., McArthur, G., Ribas, A., Roesch, A., & Hauschild, A. (2015). Melanoma. Nature Reviews Disease Primers, 1(1). https://doi.org/10.1038/nrdp.2015.3

Schadendorf, D., Van Akkooi, A. C. J., Berking, C., Griewank, K. G., Gutzmer, R., Hauschild, A., Stang, A., Roesch, A., & Ugurel, S. (n.d.). Melanoma. The Lancet, 392(10151), 971–984. https://doi.org/10.1016/s0140-6736(18)31559-9

Matthews, N. H., Li, W., Qureshi, A. A., Weinstock, M. A., & Cho, E. (2017). Epidemiology of melanoma. In Codon Publications eBooks (pp. 3–22). https://doi.org/10.15586/codon.cutaneousmelanoma.2017.ch1

Australia, C. (2019, December 18). What are the risk factors for melanoma? Www.canceraustralia.gov.au. https://www.canceraustralia.gov.au/cancer-types/melanoma/awareness

Safdar, K., Akbar, S., & Gull, S. (2021, December 1). An Automated Deep Learning based Ensemble Approach for Malignant Melanoma Detection using Dermoscopy Images. IEEE Xplore. https://doi.org/10.1109/FIT53504.2021.00046

Likhitha, S., & Baskar, R. (2022, December 1). Skin Cancer Segmentation Using R-CNN Comparing with Inception V3 for Better Accuracy. IEEE Xplore. https://doi.org/10.1109/SMART55829.2022.10047686

Ahmed. (2022, December 11). Skin Tone Analysis with HAM10000 - Ahmed - Medium. Medium. https://medium.com/@nahmed3536/skin-tone-analysis-with-ham10000-3c40bd20bf6c

Kaplan, S. (2024, April 15). Mastering The Art Of Combining Data In Data Science [Boost Your Analysis Skills]» EML. EML. https://enjoymachinelearning.com/blog/combining-data-from-different-sources-in-data-science/